مقدمة في
الإحصاء التطبيقي

إعداد:
دكتور/ أبوبكر عبد الرحمن
عبد المتعال

الجزء الأول

2022 - 2023

كلية التجارة
جامعة جنوب الوادي
قسم الأساليب الكمية

مقدمة في

الإحصاء التطبيقي

(الجزء الأول)

إعداد:

دكتور/ أبوبكر عبد الرحمن عبد المتعال

كلية التجارة – جامعة جنوب الوادي

قسم الأساليب الكمية

العام الجامعي

٢٠٢٢ / ٢٠٢٣

مقدَّمــــة

يستمد الإحصاء التطبيقي أهميته كعلم من حيث قدرته على حل المشكلات المختلفة التي يمكن أن تواجه الباحث في جميع المجالات. وبدون هذا العلم يصبح الإحصاء مجرد نظريات وقوانين جوفاء لا ترتبط بواقعنا العملي بما يعاني من مشكلات شتى. وأحياناً ما نجد أن هناك الكثير من الدارسين الذين يُلِمُّون بالنظريات والقوانين المختلفة في علم الإحصاء، ومع ذلك نجدهم عاجزين عن كيفية استخدام تلك النظريات والقوانين في حل ما يواجههم من مشكلات في حياتهم العملية. ذلك لأنه - وللأسف الشديد - أصبح تدريس علم الإحصاء في كثير من الأحوال وكأنه مجرد سرد لعمليات حسابية لا تخرج عن كونها جمعاً وطرحاً وضرباً وقسمةً، مبتعداً في ذلك عن مدلولات النتائج التي نصل إليها والتي من شأنها أن ترسو بنا عند حقيقة الأمور.

ومن هنا كان تركيز هذا الكتاب على بيان مدلولات النتائج التي نصل إليها مستعينين في ذلك بالأمثلة والتطبيقات العملية والتي تُرسِّخ في ذهن القارئ نهجاً عملياً يفيده في توظيف ما يعرف من نظريات وقوانين في حل مشكلاته المتعددة والمتنوعة. ولعل أبرز دليل على ذلك هو ما يحتويه هذا الكتاب من على عدد كبير من الأمثلة والتطبيقات المحلولة، بالإضافة إلى التمارين والتطبيقات غير المحلولة. وقد كان هناك حرص على أن تشمل تلك الأمثلة والتطبيقات المجالات الاجتماعية والاقتصادية والسياسية والتعليمية، إلى غير ذلك من المجالات.

ويتناول هذا الكتاب بعض جوانب الإحصاء التطبيقي متمثِّلاً ذلك في خمسة فصول تتناول على الترتيب الموضوعات التالية: الاحتمالات والتوزيعات الاحتمالية، نظرية التقدير الإحصائي، اختبار الفرضيات الإحصائية، اختبارات مربع كاي، وأخيراً الإحصاءات الحيوية والصحية. ولعلّي لا أدَّعي باطلاً بقولي إن هذا الكتاب قد شمل فصلاً عن الإحصاءات الصحية والحيوية لا يتوافر في أمثاله من الكتب الأخرى.

ومع كل ذلك، فإنني ألتمس منك عزيزي القارئ أن تغفر لي زلَّتي عن كل خطأ أو هفوة وقعت فيهما بسبب حرصي على الإسراع بالانتهاء من إعداد هذا الكتاب حتى يكون بين يديك في الوقت المناسب. وكم سيكون عظيم تقديري لو أنك أسديْت لي النصح وأبديْت ملاحظاتك حول هذا الكتاب وذلك حتى تكون الطبعة المقبلة إن شاء الله أكثر دقةً وأقل أخطاءً.

والله الموفِّق،،،

دكتور
أبو بكر عبد الرحمن عبد المتعال

القاهرة في يناير ٢٠٢٣

بيانات الكتاب

الكلية: التجارة

الفرقة : الثانية (شعبة الدراسة باللغة العربية)

تاريخ النشر: ٢٠٢٣/٢٠٢٢

عدد الصفحات: ٤٤٦ صفحة

المؤلف: دكتور/ أبوبكر عبد الرحمن عبد المتعال

محتويات الكتاب

الفصل الأول
الاحتمالات والتوزيعات الاحتمالية
(١ – ١٠٧)

الفصل الثاني
التقدير الإحصائي
(١٠٨ – ١٨٨)

الفصل الثالث
اختبار الفرضيات الإحصائية
(١٨٩ ــ ٢٤٠)

الفصل الرابع
اختبارات مربع كاي
(٢٤١ ــ ٢٧٤)

الفصل الخامس
الإحصاءات الحيوية والصحية
(٢٧٥ ــ ٣١٣)

تطبيقات إحصائية محلولة
(٣١٤ ــ ٣٨٧)

امتحانات
(٣٨٨ ــ ٤٤٠)

الملاحق (الجداول الإحصائية)
(٤٤١ ــ ٤٤٩)

الأشكال

الفصل الأول
الاحتمالات والتوزيعات الاحتمالية

الفصل الثالث
اختبار الفرضيات الإحصائية

الفصل الرابع
اختبارات مربع كاي

الفصل الأول

الاحتمالات والتوزيعات الاحتمالية

مقدمـــــة:

الاحتمال والفرصة، كلمتان أصبحتا تشكلان جزءاً من حياتنا اليومية حيث يندر أن يمر يوم دون أن نسأل أو نتساءل فيما بيننا عن الكثير من الأمور والتي يعتمد الوصول إلى حقيقتها على الاحتمال. فعلى سبيل المثال:

- ما هي إمكانية وجود حياة على سطح كوكب المريخ؟
- هل يُحتمل أن تقع زلازل أخرى – لا قدر الله - في مصر خلال العام الحالي؟
- على ضوء نتائج فريق النادي الاهلي حتى الآن، هل هناك فرصة لأن يفوز ببطولة الدوري العام؟
- ما هو احتمال أن تحقق الشركة المالية والصناعية المصرية أرباحاً في نهاية هذا العام وذلك على ضوء وضعها المالي الحالي؟
- ما هو مقدار الفرصة المتاحة لطالب معين للحصول على تقدير عام جيد في نهاية مرحلة البكالوريوس وذلك بالأخذ في الاعتبار نتائجه في السنوات السابقة؟
- ما هي فرصة كل من المرشحين للفوز بمنصب رئيس الولايات المتحدة الأمريكية، وذلك على ضوء استطلاعات الرأي حول انتخابات الرئاسة الأمريكية؟
- ما هو احتمال أن تستطيع مصر الوفاء بجميع ديونها الخارجية خلال العشر سنوات القادمة وذلك على ضوء خططها الاقتصادية الحالية والمستقبلية. وإذا لم تستطع ذلك، ما هي القيمة المتوقَّعة لديونها في نهاية تلك الفترة؟

لو نظرنا إلى تلك الأسئلة لوجدنا أنها تتعلق بالعديد من مجالات حياتنا، وأن الوصول إلى حقيقة الإجابات عليها إنما يتطلب من المهتمين - كلٌ حسب اختصاصه - العلم والدراية بنظرية الاحتمالات والتي تعتبر أساساً للنظرية الإحصائية. فقد بات جلياً في ظل التقدم الحضاري الذي وصلت إليه جميع الأمم أنه لا يمكن اتخاذ قرار مبنى على أسس علمية ما لم يكن هناك تطبيق للأساليب الإحصائية تطبيقاً صحيحاً يصل بنا إلى حقيقة الأمور آخذين بجوهرها لا منساقين وراء مظهرها.

وفي دراستنا للاحتمالات سوف نبدأ بعرض سريع لبعض مفاهيم وتعريفات نظرية الاحتمالات وبعض أساسيات هذه النظرية وذلك بالقدر الذي يمهد لدراسة التوزيعات الاحتمالية وأساليب الاستدلال الإحصائي، وهما الموضوعان اللذان يدخلان ضمن نطاق الفرع الثاني لعلم الإحصاء وهو الإحصاء الاستدلالي.

وفي هذا الفصل سوف نقوم بدراسة الاحتمالات والتوزيعات الاحتمالية، وأما الاستدلال الإحصائي فسوف يكون موضوع الفصول التالية من هذا الكتاب.

مبادئ الاحتمالات

كما أشرنا سابقاً فإنه كثيراً ما نستخدم في حياتنا اليومية العادية التعبيرات: يحتمل أو يمكن أو يجوز، وذلك للدلالة على شيء غير مؤكد الحدوث. فقد يقول قائل: "من المحتمل أن أنجح هذا العام"، ويقول آخر: "إنني شبه متأكد من أن السماء سوف تمطر غداً"، بينما يقول ثالث: "إنه لمن المؤكد أن غداً سوف يكون إجازة رسمية لجميع موظفي المصالح الحكومية". إلى غير ذلك من التعبيرات التي تعكس درجات مختلفة من عدم التأكد أو التأكـــد وذلك فيما يتعلق بتحقق حدث معين. ولو تأملنا هذه التعبيرات لوجدنا أن اختلاف درجة التأكد من شخص لآخر إنما تأتى اعتماداً على مدى توافق الظروف الحالية مع ظروف سابقة مشابهة والذي ينعكس بدوره على تباين درجات التأكد، ليس فقط من شخص إلى آخر ولكن أيضاً من حدث لآخر ومن هنا ظهرت الحاجة إلى وضع أسس علمية لقياس الاحتمال (درجة عدم التأكد) وذلك بدلاً من الاعتماد في قياسه على مجرد الخبرة الشخصية والتي تختلف درجة الثقة فيها من شخص لآخر.

هذا ويسمى العلم الذي يبحث في هذا الموضوع بعلم الاحتمالات. ومن هذا المنطلق يمكننا تعريف علم الاحتمالات بأنه أحد فروع الرياضيات التطبيقية والذي يهتم بدراسة أثر عامل الصدفة على تحقق الظواهر العلمية وكيفية قياس هذا الأثر رياضياً.

ولعله من الملائم أن نبدأ دراستنا لنظرية الاحتمالات بتقديم بعض التعريفات والمفاهيم الأساسية التي تُستخدم في تعريف الاحتمال وذلك على النحو الموضح فيما يلي:

أولاً: بعض المفاهيم والتعريفات في علم الاحتمالات

سوف نتناول في هذا الشأن بعض التعريفات والمفاهيم التي تساعدنا على فهم المقصود بكلمة "احتمال" وذلك كما يلي:

١- التجربة العشوائية والحدث العشوائي:

يُقصد **بالتجربة العشوائية** أية عملية تكون لها نتائج معينة يمكن مشاهدتها ولا يمكن التنبؤ بنتيجة محددة لها قبل الانتهاء من إجراء التجربة.

مثال: رمي قطعة عملة (غير متحيزة) هو تجربة عشوائية تتمثل نتائجها في حالتيْن هما ظهور الوجه الذي عليه الشعار أو الوجه الذي عليه الكتابة. ومن الواضح أنه لا يمكننا التنبؤ مسبقاً بأي من الوجهيْن سوف يظهر عند رمي قطعة العملة. وكذلك رمي زهرة نرد هو أيضاً تجربة عشوائية نتائجها المتوقَّعة هي ظهور أحد الأوجه الستة والتي تحمل الأرقام من ١ إلى ٦.

وأما **الحدث العشوائي** فيُقصد به في نظرية الاحتمالات تلك الحالة أو الحالات التي تكون موضع الاهتمام. فقد يكون الحدث عند رمي قطعة عملة هو ظهور الوجه الذي عليه الشعار، أو قد يكون ظهور أحد الأوجه التي تحمل الأرقام ٤، ٥، ٦ وذلك عند رمي زهرة نرد. وهذا يعنى أن الحدث قد يشير إلى حالة واحدة (نتيجة واحدة ممكنة) أو إلى عدة حالات (أكثر من نتيجة).

٢- الحالات الكلية الممكنة والحالات المُواتية:

الحالات الكلية الممكنة لتجربة ما هي النتائج الكلية (الحالات الكلية) المتوقَّع حدوثها نتيجة لإجراء هذه التجربة.

مثال: عند رمي قطعة عملة نجد أن عدد الحالات الكلية الممكنة هو ٢ ويتمثل ذلك في ظهور الشعار أو ظهور الكتابة. وفي حالة رمي زهرة نرد فإن عدد الحالات الكلية الممكنة يساوي ٦، ويتمثل في ظهور الأوجه الستة المختلفة لزهرة نرد والتي تحمل الأرقام من ١ إلى ٦.

وأما **الحالات المواتية** لحدث معين فيُقصد بها الحالات التي يمكن أن يتحقق فيها الحدث موضع الاهتمام. فعند رمي قطعة العملة على سبيل المثال نجد أن عدد الحالات المواتية لظهور الوجه الذي عليه الشعار هو حالة واحدة فقط. وأما في حالة زهرة النرد نجد أن عدد الحالات المواتية لظهور رقم يزيد عن ٤ هو حالتان وهما ظهور الرقم ٥ وظهور الرقم ٦.

٣- الحوادث البسيطة والحوادث المركبة:

الحدث البسيط هو ذلك الحدث الذي يترتب عليه نتيجة واحدة.

مثال: ظهور الوجه الذي عليه الشعار عند رمي قطعة عملة، أو ظهور الرقم ٣ عند رمي زهرة فرد هما حدثان بسيطان.

وأما **الحدث المركب** فهو الذي يترتب عليه أكثر من نتيجة، ومثال ذلك ظهور الأوجه التي تحمل الأرقام ٤ أو ٦ عند رمي زهرة نرد. وكذلك أن تكون الورقة المسحوبة من مجموعة أوراق اللعب ولد أو سوداء اللون.

٤- الحوادث المتنافية وغير المتنافية:

يُقال إن الحدثيْن أ و ب هما **حدثان متنافيان** بالتبادل وذلك إذا كان حدوث أحدهما ينفي أو يمنع حدوث الآخر. بمعنى أنه إذا وقع الحدث أ فإن هذا يؤدى إلى عدم وقوع الحدث ب والعكس صحيح. ويمكن التعبير عن ذلك في صورة احتمال كما يلي:

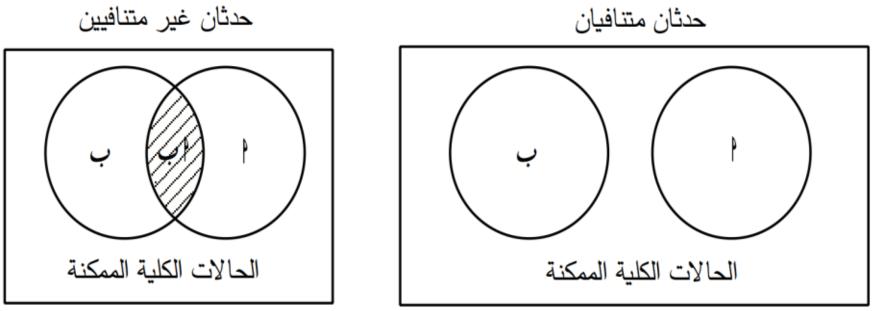
ح(أب) = صفر

حيث ح(أب) هو احتمال تحقق الحدثيْن أ و ب معاً في نفس الوقت.

مثال: ظهور الوجه الذي عليه الشعار وظهور الوجه الذي عليه الكتابة عند رمي قطعة عملة هما حدثان متنافيان، حيث أن ظهور أحدهما يمنـــع ظهور الآخر وذلك نظراً لاستحالة ظهور الوجهيْن معاً في نفس الوقت. وكذلك ظهور الورقة التي تحمل الرقم ١٠ والورقة التي تحمل الرقم ٨ عند سحب ورقة من مجموعة أوراق اللعب هما أيضا حدثان متنافيان، إذ أن ظهور ورقة تحمل الرقميْن معاً هو أمر يستحيل حدوثه.

وأما إذا كان وقوع الحدث أ لا ينفي أو لا يمنع وقوع الحدث الآخر ب أو العكس، فيُقال حينئذ أن الحدثيْن أ، ب غير متنافييْن. ومثال ذلك أنه عند سحب ورقة من مجموعة أوراق اللعب فإن الحدث الممثل في ظهور الورقة ٩ والحدث الممثل في ظهور الورقة من نوع القلب هما حدثان غير متنافييْن. بمعنى أنه من الممكن أن تكون الورقة المسحوبة تحمل الرقم ٩ وفي نفس الوقت تكون من نوع القلب. أي أن: ح(أب) $\neq$ صفر.

هذا والشكل (١ - ١) يوضح فكرة الحوادث المتنافية والحوادث غير المتنافية.



الشكل (١ – ١)
الأحداث المتنافية وغير المتنافية

٥- الحوادث المستقلة والحوادث غير المستقلة:

يقال للحدثيْن أ، ب أنهما **حدثان مستقلان** إذا كان احتمال تحقق أحدهما لا يتأثر بتحقق أو عدم تحقق الحدث الآخر.

مثال: إذا قمنا بسحب كرتيْن متتاليتيْن من صندوق به عدد من الكرات ذات لون ما وعدد آخر من الكرات ذات لون مختلف. فإذا كنا نعيد الكرة الأولى إلى الصندوق بعد سحبها ومشاهدتها فإن نتيجة سحب الكرة الثانية لا تتأثر بنتيجة سحب الكرة الأولى. وعليه يمكننا أن نقول إن نتيجتي سحب الكرة الأولى وسحب الكرة الثانية هما حدثان مستقلان.

وأما الحدثان أ، ب فإنهما يصبحان حدثيْن غير مستقليْن وذلك إذا كان السحب يتم مع إعادة الكرة المشار إليها إلى الصندوق. بمعنى أنه إذا لم نقم بإعادة الكرة المسحوبة أولاً إلى الصندوق فإن نتيجة السحبة الثانية سوف تتوقف بكل تأكيد على نتيجة السحبة الأولى.

٦- الحوادث المؤكدة والحوادث المستحيلة:

الحدث المؤكد هو ذلك الحدث الذي لابد من وقوعه.

مثال: إذا كان لدينا صندوق به عدد من الكرات البيضاء وعدد من الكرات السوداء فإن الحدث المؤكد عند سحب كرة من الصندوق هو ظهور كرة بيضاء أو كرة سوداء.

وأما **الحدث المستحيل** فهو ذلك الحدث الذي لا يمكن وقوعه. ومثال ذلك أن تكون الكرة المسحوبة من الصندوق ذات لون أحمر، حيث أن الصندوق لا يحتوي أساساً على أية كرات حمراء.

٧- الحوادث الشاملة:

يقصد **بالحوادث الشاملة** تلك الحوادث التي لابد أن يحدث إحداها كنتيجة لإجراء التجربة وبمعنى آخر فإن الحوادث الشاملة يقصد بها الحوادث الكلية الممكنة الحدوث عند إجراء تجربة ما.

مثال: عند رمي قطعة عملة فإنه لابد وأن يظهر الشعار أو الكتابة. حينئذ يمكننا القول إن الحدث الممثَّل في ظهور الشعار والحدث الممثَّل في ظهور الكتابة هما حدثان شاملان. وكذلك أوجه زهرة النرد الستة والتي تحمل الأرقام من ١ إلى ٦ هي أيضاً حوادث شاملة. ويلاحظ أن مجموع احتمالات حدوث الحوادث الشاملة لابد وأن يساوي الواحد الصحيح.

ثانيــاً: تعريف الاحتمال

هناك تعريفان أساسيان للاحتمال أحدهما يسمَّى بالتعريف الكلاسيكي، بينما يسمَّى الآخر بالتكرار النسبي لتعريف الاحتمال (الاحتمال التجريبي).

١- التعريف الكلاسيكي للاحتمال:

إذا افترضنا أن عدد الحالات المواتية لوقوع حدث معين (أ) في تجربة ما هو م، وإن عدد الحالات الكلية الممكنة لهذه التجربة هو ن فإن احتمال تحقق الحدث أ (نجاحه) يأخذ الصورة التالية وفقاً للتعريف الكلاسيكي للاحتمال:

$$ع(أ) = \frac{\text{عدد الحالات المُواتِية}}{\text{عدد الحالات الكلية}} = \frac{م}{ن} \qquad (١ - ١)$$

كما أن احتمال عدم تحقق الحدث أ (فشله) يأخذ الصورة التالية:

$$ع(\bar{أ}) = \frac{\text{عدد الحالات غير المواتية}}{\text{عدد الحالات الكلية}} = \frac{ن - م}{ن}$$

$$= ١ - \frac{م}{ن} = ١ - ع(أ)$$

وهذا التعريف للاحتمال يقوم على فرض أساسي وهو أن جميع الحوادث لها نفس الفرصة في الحدوث. وبلغة الاحتمالات فإنه يمكننا القول بأن هذا التعريف للاحتمال يَفترض أن احتمالات حدوث الحوادث المختلفة متساوية. وهنا نجد أن هذا التعريف للاحتمال يعتبر تعريفاً متداخلاً أو دائرياً، بمعنى أننا نعرف الاحتمال بالاعتماد على الاحتمال ذاته. لذلك اقترح البعض تعريفاً آخر للاحتمال وهو ما يسمى بالتعريف الإحصائي للاحتمال أو الاحتمال التجريبي Empirical Probability والذي يعتمد على التكرار النسبي. وهو ما سنوضحه فيما يلي:

٢- التكرار النسبي كتعريف للاحتمال:

وفقاً لهذا التعريف فإن الاحتمال التجريبي لوقوع حدث معين يساوي التكرار النسبي لحدوث الحدث وذلك عند تكرار التجربة العشوائية عدداً كبيراً من المرات. أي أن احتمال حدوث الحدث أ يأخذ الصورة التالية:

$$ح(أ) = \frac{\text{عدد مرات حدوث الحدث (النجاح)}}{\text{عدد مرات تكرار التجربة}} \qquad (١ - ٢)$$

هذا ويلاحظ أنه بزيادة عدد مرات التجربة ليقترب من ما لا نهاية ($ن \longrightarrow \infty$) فإن التكرار النسبي (الاحتمال التجريبي) يقترب في قيمته من الاحتمال المسبق الكلاسيكي. فعلى سبيل المثال إذا قمنا برمي قطعة عملة عدداً كبيراً جداً من المرات فإنه من المتوقَّع أن تقترب نسبة عدد المرات التي يظهر فيها الشعار إلى عدد مرات تكرار التجربة من $\frac{١}{٢}$ وهي قيمة الاحتمال الكلاسيكي. أي أن:

$$\text{الاحتمال التجريبي} = \text{الاحتمال الكلاسيكي} = \frac{١}{٢} \quad \text{عندما } ن \longrightarrow \infty$$

ثالثاً: خصائص الاحتمال

١- من أهم خصائص الاحتمال أن قيمته دائماً موجبة وأقل من أو تساوي الواحد الصحيح. أي أن:

$$\text{صفر} \leq ح(أ) \leq ١$$

أي أن قيمة الاحتمال تنحصر ما بين الصفر والواحد الصحيح. وقيمة الاحتمال تساوي الصفر في حالة الحدث المستحيل بينما تساوي الواحد الصحيح في حالة الحدث المؤكد.

٢- مجموع احتمالات الحدوث للأحداث الكلية الممكنة لتجربة ما يساوي الواحد الصحيح. فعلى سبيل المثال:

عند رمي قطعة عملة: ح(ظهور الشعار) + ح(ظهور الكتابة) = ١

عند رمي زهرة نرد: ح(١) + ح(٢) +.... + ح(٦) = ١

٣- إذا كان احتمال حدوث الحدث أ هو ع(أ) فإن احتمال عدم حدوثه هو:

ع($\overline{\text{أ}}$) = ١ – ع(أ)

أي أن الحدث $\overline{\text{أ}}$ هو الحدث (أو الأحداث) المكمِّلة للحدث أ.

مثال (١):

إذا ألقيت زهرة نرد (غير متحيزة)، احسب احتمال الحصول على:

أ- الرقم ٣ ب- عدد زوجي.

الحـــل:

أ- حيث أن لزهرة النرد ستة أوجه تحمل الأرقام من ١ إلى ٦ فإننا نجـــد أن:

ع(ظهور الرقم ٣) = $\frac{١}{٦}$

ب- الحالات المواتية لظهور عدد زوجي هي الحالات التي يظهر فيهـــا الأرقام ٢ ، ٤ ، ٦.

∴ ع(ظهور عدد زوجي) = $\frac{٣}{٦}$ = $\frac{١}{٢}$

مثال (٢):

إذا سحبت ورقة من مجموعة أوراق اللعب مخلوطة خلطاً جيداً، احسب احتمال أن تكون الورقة المسحوبة:

أ- عشـــرة ب- صـــــورة جـ- ذات لون أسود د- من نوع القلب

الحـــل:

أ- حيث أن مجموعة أوراق اللعب تحتوي على أربع ورقات تحمل كل منها الرقم ١٠، فإننا نحصل على:

ع(الحصول على عشرة) = $\frac{٤}{٥٢}$ = $\frac{١}{١٣}$

ب- من المعروف أن عدد الصور في مجموعة أوراق اللعب = ١٢

∴ ع(الحصول على صورة) = $\frac{١٢}{٥٢}$ = $\frac{٣}{١٣}$

جـ - من المعروف أيضا أن مجموعة أوراق اللعب تتكون من ٢٦ ورقة ذات لون أسود و ٢٦ ورقة ذات لون أحمر.

∴ ع(الحصول على لون أسود) = $\frac{٢٦}{٥٢}$ = $\frac{١}{٢}$

د- حيث أن مجموعة أوراق اللعب بها ١٣ ورقة من نوع القلب، فإننا نحصل على:

ع(ظهور ورقة من نوع القلب) = $\frac{١٣}{٥٢}$ = $\frac{١}{٤}$

مثال (٣):

صندوق به ١٠ كرات صغيرة متماثلة منها ٣ كرات بيضاء، ٥ كرات حمراء وكرتان ذات لون أسود. فإذا تم خلط الكرات داخل الصندوق خلطاً جيداً ثم سحبت منه كرة عشوائياً، احسب احتمال أن تكون الكرة المسحوبة:

أ- سـوداء ب- بيضاء أو سوداء جـ - ليست بيضاء

الحـل:

لو سمينا ظهور الكرات البيضاء والسوداء والحمراء على التوالي بالأحداث ب، س، ع. فإنه يمكننا الحصول على الاحتمالات المطلوبة كما يلي:

أ- $ح(س) = \frac{\text{عدد الكرات السوداء}}{\text{عدد الكرات بالصندوق}} = \frac{٢}{١٠} = ٠,٢$

ب- $ح(ب \text{ أو } س) = \frac{٣ + ٢}{١٠} = \frac{٥}{١٠} = ٠,٥$

جـ- احتمال أن الكرة ليست بيضاء = $ح(\bar{ب})$

$= ١ - ح(ب) = ١ - \frac{٣}{١٠} = ٠,٧$

مثال (٤):

ألقيت زهرتا نرد، احسب احتمال الحصول على:

أ- رقميْن متشابهيْن ب- رقميْن مجموعهما يساوي ٨

الحـل:

عدد النتائج الكلية الممكنة عند إلقاء زهرتي نرد = $(٦)^{٢}$ = ٣٦ حالة

أ- الحالات المواتية لظهور رقميْن متشابهيْن هي:

(١ ، ١)، (٢ ، ٢)، .....، (٦ ، ٦)، أي ست حالات.

∴ ح(الحصول على رقميْن متشابهيْن) = $\frac{٦}{٣٦} = \frac{١}{٦}$

ب- الحالات المواتية لظهور رقميْن مجموعهما يساوي ٨ هي كالتالي:

(رقم الزهرة الأولى، رقم الزهرة الثانية): (٢ ، ٦) ، (٣ ، ٥) ، (٤ ، ٤) ، (٥ ، ٣)، (٦ ، ٢).

∴ الاحتمال المطلوب = $\frac{٥}{٣٦}$

رابعاً: قانون جمع الاحتمالات

يستخدم قانون جمع الاحتمالات لحساب احتمال وقوع الحدث (أ) أو وقوع الحدث (ب)، ويأخذ الصورة التالية:

ع(أ أو ب) = ع(أ) + ع(ب) – ع(أ ب) **(١ – ٣)**

أو: **ع(أ أو ب) = ع(أ) + ع(ب)** **(١ – ٤)**

حيث: ع(أ) = احتمال تحقق الحدث أ ، ع(ب) = احتمال تحقق الحدث ب،

ع(أ ب) = احتمال تحقق الحدثيْن أ، ب معاً في نفس الوقت.

وتستخدم الصورة (١ – ٣) إذا كان الحدثان أ، ب غير متنافييْن، أي أن:

ع(أ ب) $\neq$ صفر. وأما الصورة (١ – ٤) فتستخدم إذا كان الحدثان أ، ب متنافييْن،

أي أن: ع(أ ب) = صفر.

هذا ويمكن تعميم قانون جمع الاحتمالات في حالة ثلاث حوادث على النحو التالي:

ع(أ أو ب أو جـ) = ع(أ) + ع(ب) + ع(جـ) – ع(أ ب)

– ع(أ جـ) – ع(ب جـ) + ع(أ ب جـ)

وذلك في حالة الحوادث غير المتنافية.

أو: ع(أ أو ب أو جـ) = ع(أ) + ع(ب) + ع(جـ) وذلك في حالة الحوادث المتنافية.

مثال (٥):

إذا سحبت ورقة من مجموعة أوراق اللعب، احسب احتمال أن تكون الورقة المسحوبة:

أ- عشرة أو بنت ب- خمسة أو ورقة من نوع البستوني

الحـل:

أ- لو رمزنا إلى ظهور العشرة بالحدث (أ) وإلى ظهور البنت بالحدث (ب)، وحيث أنه لا يمكن أن تكون الورقة المسحوبة عشرة وبنت في نفس الوقت، فإن هذيْن الحدثيْن متنافيان.

أي أن: ع(أ أو ب) = ع(أ) + ع(ب)

$$= \frac{٤}{٥٢} + \frac{٤}{٥٢} = \frac{٢}{١٣}$$

ب- سوف نرمز إلى ظهور الخمسة بالحدث (أ) وإلى ظهور ورقة من نوع البستوتي بالحدث (ب). وحيث أنه من الممكن أن تكون الورقة المسحوبة خمسة وفي نفس الوقت من نوع البستوتي (هناك ورقة واحدة تجمع بين هاتيْن الصفتيْن) فإن هذيْن الحدثيْن غير متنافييْن.

أي أن: ع(أ أو ب) = ع(أ) + ع(ب) – ع(أ ب)

$$= \frac{٤}{٥٢} + \frac{١٣}{٥٢} - \frac{١}{٥٢} = \frac{١٦}{٥٢} = \frac{٤}{١٣}$$

مثال (٦):

في مثال (١)، احسب احتمال ظهور الرقم ٤ أو الرقم ٦.

الحــل:

لو رمزنا لظهور الرقم ٤ بالحدث (أ) ولظهور الرقم ٦ بالحدث (ب). وحيث أنه لا يمكن أن يظهر كلا الرقميْن على وجه واحد فإنهما حدثان متنافيان.

أي أن: ع(أ أو ب) = ع(أ) + ع(ب) = $\frac{١}{٦} + \frac{١}{٦} = \frac{١}{٣}$

خامساً: الاحتمال الشرطي: Conditional Probability

في كثير من الأحيان يكون المطلوب هو حساب احتمال تحقق حدث معين وذلك بافتراض حدوث حدث آخر مرتبط به. وهذا الاحتمال هو ما يعرف بالاحتمال الشرطي، ويتم حسابه باستخدام الصورة التالية:

$$\text{ع}(\text{أ} \mid \text{ب}) = \frac{\text{ع}(\text{أب})}{\text{ع}(\text{ب})} \qquad (١ - ٥)$$

$$\text{ع}(\text{ب} \mid \text{أ}) = \frac{\text{ع}(\text{أب})}{\text{ع}(\text{أ})} \qquad (١ - ٦)$$

حيث: ع(أ) احتمال تحقق الحدث أ، ع (ب) احتمال تحقق الحدث ب،

ع (أ | ب) احتمال تحقق الحدث أ بافتراض تحقق الحدث ب،

ع (ب | أ) احتمال تحقق الحدث ب باقتراض تحقق الحدث أ،

ع (أ ب) احتمال تحقق الحدثيْن معاً في نفس الوقت.

مثال (٧):

إذا سحبت ورقة من مجموعة أوراق اللعب وكانت حمراء اللون، احسب احتمال أن تكون الورقة المسحوبة ولداً.

الحـل:

سوف نسمي ظهور ورقة حمراء اللون بالحدث أ، وظهور ورقة ولد بالحدث ب، وعليه يكون المطلوب هو: ع(ب | أ)، وحيث أن:

$$ع(ب \mid أ) = \frac{ع(أ ب)}{ع(أ)}$$

$$ع(أ) = \frac{٢٦}{٥٢} = \frac{١}{٢}$$

ع(أ ب) = احتمال أن تكون الورق المسحوبة حمراء اللون وولد في نفـــس الوقت

$$\therefore ع(أ ب) = \frac{٢}{٥٢} = \frac{١}{٢٦}$$

$$\therefore \text{الاحتمال المطلوب} = \frac{ع(أ ب)}{ع(أ)} = \frac{١}{٢٦} \div \frac{١}{٢} = \frac{١}{١٣}$$

مثال (٨):

إذا ألقيت زهرتا نرد، احسب احتمال الحصول على رقميْن مجموعهما ٧ على أن يكون الرقم على الزهرة الأولى ٣.

الحـــل:

سوف نسمي الأحداث على النحو التالي:

الرقمان مجموعهما ٧ الحدث (أ)

الرقم على الزهرة الأولى ٣ الحدث (ب)

فيكون المطلوب هو إيجاد الاحتمال: ع(أ | ب)

وحيث أن: $ع(أ \mid ب) = \frac{ع(أ ب)}{ع(ب)}$

ع(أ ب) = احتمال أن الرقمين مجموعهما ٧ وفي نفس الوقت يكون الرقم على الزهرة الأولى ٣ [وهذه يمثلها حالة واحدة هي (٣ ، ٤)].

∴ ع(أ ب) = $\frac{١}{٣٦}$

ع(ب) = احتمال أن الرقم على الزهرة الأولى ٣، ويمثل ذلك ست حالات هي: (٣ ، ١)، (٣ ، ٢)،......، (٣ ، ٦).

أي أن: ع(ب) = $\frac{٦}{٣٦} = \frac{١}{٦}$

∴ الاحتمال المطلوب = $\frac{ع(أ ب)}{ع(ب)} = \frac{١}{٣٦} \div \frac{١}{٦} = \frac{١}{٦}$

مثال (٩):

فيما يلي نتيجة امتحان مادة مبادئ الإحصاء لطلاب كلية التجارة بإحدى الجامعات وذلك في دور يناير ٢٠٢٠م:

نتيجة الامتحان / نظام الدراسة	راسب	ناجح	المجموع
انتظام	١٩٢	٧١٦	٩٠٨
انتساب موجه	١٤٤	٢٠٨	٣٥٢
المجموع	٣٣٦	٩٢٤	١٢٦٠

المطلوب:

أولاً: إذا أُختير أحد الطلاب الناجحين بطريقة عشوائية، احسب احتمال أن يكون الطالب:

أ- منتظماً ب- منتسباً

ثانياً: إذا أُختير أحد الطلاب المنتظمين بطريقة عشوائية، احسب احتمال أن يكون الطالب:

أ- ناجحاً ب- راسباً

الحل:

سوف نسمي الأحداث كما يلي:

طالب ناجح: الحدث (أ) ، طالب راسب: الحدث (ب)

طالب منتظم: الحدث (جـ) ، طالب راسب: الحدث (د)

أولاً:

أ- المطلوب هو: ع(منتظم | ناجح) = ع(جـ | أ)

وحيث أن:

$$ع(جـ \mid أ) = \frac{ع(جـ أ)}{ع(أ)}$$

ع(جـ أ) = احتمال أن يكون الطالب منتظماً ناجحاً

$$= \frac{٧١٦}{١٢٦٠} = ٠,٥٦٨ ،$$

ع(أ) = احتمال أن يكون الطالب ناجحاً

$$= \frac{٩٢٤}{١٢٦٠} = ٠,٧٣٣$$

∴ الاحتمال المطلوب = ع(منتظم | ناجح) = $\frac{ع(جـ أ)}{ع(أ)}$

$$= \frac{٠,٥٦٨}{٠,٧٣٣} = ٠,٧٧٥$$

ب- المطلوب هو: ع(منتسب | ناجح) = ع(د | أ)

وحيث أن:

$$ع(د \mid أ) = \frac{ع(د أ)}{ع(أ)}$$

ع(د أ) = احتمال أن يكون الطالب منتسباً ناجحاً

$$= \frac{٢٠٨}{١٢٦٠} = ٠,١٦٥ ،$$

ع(أ) = احتمال أن يكون الطالب ناجحاً = ٠,٧٣٣ تم حسابه من قبل.

∴ الاحتمال المطلوب = ع(منتسب | ناجح) = $\frac{ع(د أ)}{ع(أ)}$

$$= \frac{٠,١٦٥}{٠,٧٣٣} = ٠,٢٢٥$$

ويمكن حساب هذا الاحتمال بطريقة أخرى كما يلي:

ع(د | أ) = ١ – ع(جـ | أ)

حيث أن ع(جـ | أ) هو المكمل للاحتمال ع(د | أ)

∴ ع(د | أ) = ١ – ٠,٧٧٥ = ٠,٢٢٥

وهي نفس النتيجة السابقة.

ثانياً:

أ- الاحتمال المطلوب هو: ع(ناجح | منتظم) = ع(أ | جـ)

وحيث أن:

$$ع(أ \mid جـ) = \frac{ع(أ\,جـ)}{ع(جـ)}$$

ع(أ جـ) = ع(جـ أ) = ٠,٥٦٨ تم حسابه في أولاً

ع(جـ) = احتمال أن يكون الطالب منتظماً = $\frac{٩٠٨}{١٢٦٠}$ = ٠,٧٢١

∴ الاحتمال المطلوب = ع(أ | جـ) = $\frac{٠,٥٦٨}{٠,٧٢١}$ = ٠,٧٨٨

ب- الاحتمال المطلوب هو: ع(راسب | منتظم) = ع(ب | جـ)

وحيث أن:

$$ع(ب \mid جـ) = \frac{ع(ب\,جـ)}{ع(جـ)}$$

ع(ب جـ) = احتمال أن يكون الطالب منتظماً راسباً

= $\frac{١٩٢}{١٢٦٠}$ = ٠,١٥٢

ع(جـ) = ٠,٧٢١ تم حسابه من قبل

∴ الاحتمال المطلوب = ع(راسب | منتظم) = $\frac{ع(ب\,جـ)}{ع(جـ)}$

= $\frac{٠,١٥٢}{٠,٧٢١}$ = ٠,٢١١

ويمكن أيضاً حساب قيمة هذا الاحتمال باستخدام فكرة الاحتمال المكمِّل وذلك كما يلي:

ح(ب | جـ) = ١ – ح(أ | جـ) = ١ – ٠,٧٨٨ = ٠,٢١٢

ملحوظة: الفرق بين النتيجتيْن وهو ٠,٠٠١ حدث نتيجة التقريب.

سادساً: قانون ضرب الاحتمالات

في الأمثلة السابقة حسبنا احتمال وقوع حدثيْن معاً في نفس الوقت وذلك بحصر عدد الحالات المواتية وقسمتها على عدد الحالات الكلية الممكنة. وفي كثير من الأحيان لا تكون عملية حصر الحالات المواتية بالأمر الهيِّن أو المتاح. حينئذ يمكننا استخدام قانون ضرب الاحتمالات والموضح فيما يلي:

يمكن كتابة قانون ضرب الاحتمالات على الصورة التالية:

ح(أب) = ح(أ)ح(ب) إذا كان الحدثان مستقليْن (١ – ٧)

ح(أب) = ح(أ)ح(ب | أ)

= ح(ب)ح(أ | ب) إذا كان الحدثان غير مستقليْن (١ – ٨)

وهنا يمكننا استنتاج أنه إذا كان الحدثان مستقليْن فإن:

ح(أ) = ح(أ | ب) ، ح(ب) = ح(ب | أ)

أي أن ظهور أحد الحدثيْن لا يؤثر على احتمال ظهور الحدث الآخر.

مثال (١٠):

إذا كان احتمال نجاح الطالب في مادة الإحصاء هو ٠,٧ واحتمال نجاحه في مادة المحاسبة هو ٠,٩ احسب احتمال نجاح الطالب في مادتي الإحصاء والمحاسبة معاً.

الحـل:

حيث أن نجاح الطالب في مادة الإحصاء لا يعتمد على نجاحه أو رسوبه في مادة المحاسبة والعكس صحيح، فإن الحدثيْن مستقلان. أي أننا لو سمينا نجاح الطالب في مادة الإحصاء بالحدث (أ) ونجاحه في مادة المحاسبة بالحدث (ب)، فإن الاحتمال المطلوب هو:

ح(أب) = ح(أ)ح(ب) = ٠,٧ × ٠,٩ = ٠,٦٣

مثال (١١):

صندوق به ١٠ كرات منها ٥ كرات بيضاء، ٢ سوداء، ٣ خضراء. سُحبت كرتان متتاليتان من الصندوق. احسب احتمال أن تكون الكرتان المسحوبتان على النحو التالي:

أ- الأولى بيضاء والثانية بيضاء ب- الأولى سوداء والثانية سوداء

جـ- الأولى بيضاء والثانية سوداء د- الأولى سوداء والثانية بيضاء

وذلك بافتراض:
أولاً: إعادة الكرة الأولى المسحوبة إلى الصندوق قبل سحب الكرة الثانية.
ثانياً: عدم إعادة الكرة الأولى المسحوبة إلى الصندوق قبل سحب الكـرة الثانية.

الحــل:

أولاً: حيث أنه تم إعادة الكرة المسحوبة الأولى فإن احتمال أن تكون الكرة الثانية بيضاء أو سوداء لا يتأثر بنتيجة سحب الكرة الأولـــى، أي أن الحدثيْن في هذه الحالة مستقلان. فلو سمينا ظهور كرة بيضاء بالحدث (أ) وظهور كرة سوداء بالحدث (ب) فإنه يمكن حساب الاحتمالات المطلوبة على النحو التالي:

أ- الأولى بيضاء والثانية بيضاء:

الاحتمال المطلوب = ع(أ أ) = ع(أ)ع(أ)

$$= \frac{٥}{١٠} \times \frac{٥}{١٠} = ٠,٢٥$$

ب- الأولى سوداء والثانية سوداء:

الاحتمال المطلوب = ع(ب ب) = ع(ب)ع(ب)

$$= \frac{٢}{١٠} \times \frac{٢}{١٠} = ٠,٠٤$$

جـ - الأولى بيضاء والثانية سوداء:

الاحتمال المطلوب = ع(أ ب) = ع(أ)ع(ب)

$$= \frac{٥}{١٠} \times \frac{٢}{١٠} = ٠,١$$

د- الأولى سوداء والثانية بيضاء:

الاحتمال المطلوب = ع(ب أ) = ع(ب)ع(أ)

$$= \frac{٢}{١٠} \times \frac{٥}{١٠} = ٠,١$$

ثانياً: في هذه الحالة وحيث أنه لا يتم إعادة الكرة المسحوبة الأولى إلى الصندوق قبل سحب الكرة الثانية فإن احتمال أن تكون الكرة الثانيـــة بيضاء أو سوداء سوف يتوقف على نتيجة سحب الكرة الأولى، أي أن الحدثيْن غير مستقليْن وعليه يمكن حساب الاحتمالات المطلوبة كما يلي:

أ- الأولى بيضاء والثانية بيضاء:

الاحتمال المطلوب = ع(أ أ) = ع(أ)ع(أ | أ)

$$= \frac{٥}{١٠} \times \frac{٤}{٩} = ٠,٢٢٢$$

ب- الأولى سوداء والثانية سوداء:

الاحتمال المطلوب = ع(ب ب) = ع(ب)ع(ب | ب)

$$= \frac{٢}{١٠} \times \frac{١}{٩} = ٠,٠٢٢$$

جـ - الأولى بيضاء والثانية سوداء:

الاحتمال المطلوب = ع(أ ب) = ع(أ)ع(ب | أ)

$$= \frac{٥}{١٠} \times \frac{٢}{٩} = ٠,١١١$$

د- الأولى سوداء والثانية بيضاء:

الاحتمال المطلوب = ع(ب أ) = ع(ب)ع(أ | ب)

$$= \frac{٢}{١٠} \times \frac{٥}{٩} = ٠,١١١$$

سابعاً: تطبيق قانوني جمع وضرب الاحتمالات معاً:

فيما يلي بعض الأمثلة التي يمكن فيها استخدام كل من قانون جمع الاحتمالات وقانون ضرب الاحتمالات.

مثال (١٢):

في مثال (١١)، لو سحبنا كرتيْن متتاليتيْن من الصندوق مع عدم إرجاع الكرة الأولى المسحوبة. احسب احتمال أن تكون الكرة الثانية خضراء.

الحـل:

الحالة التي تكون فيها الكرة الثانية خضراء هي إحدى الحالات التالية: (الأولى، الثانية): (بيضاء، خضراء) أو (سوداء، خضراء) أو (خضراء، خضراء). فإذا سمينا ظهور كرة بيضاء وكرة سوداء وكرة خضراء بالأحداث أ، ب، جـ على التوالي، فإن الاحتمال المطلوب يمكن حسابه كما يلي:

ع(الثانية خضراء) = ع(أ)ع(جـ | أ) + ع(ب)ع(جـ | ب) + ع(جـ)ع(جـ | جـ)

$$= \frac{٥}{١٠} \times \frac{٣}{٩} + \frac{٢}{١٠} \times \frac{٣}{٩} + \frac{٣}{١٠} \times \frac{٢}{٩} = ٠,٣$$

مثال (١٣):

باستخدام بيانات مثال (١٠)، احسب احتمال نجاح الطالب في مادة الإحصاء أو نجاحه في مادة المحاسبة.

الحـــل:

المطلوب هو حساب الاحتمال: ع(أ أو ب)، وحيث أن الحدثيْن غير متنافييْن ومستقلان فإنه يمكن حساب الاحتمال المطلوب كما يلي:

ع(أ أو ب) = ع(أ) + ع(ب) – ع(أ ب)

= ع(أ) + ع(ب) – ع(أ)ع(ب)

= ٠,٧ + ٠,٩ – ٠,٧ × ٠,٩ = ١,٦ – ٠,٦٣ = ٠,٩٧

مثال (١٤):

إذا كان احتمال أن يصيب الشخص (أ) هدفاً معيناً هو ٠,٨، واحتمال أن يصيب شخصاً آخر (ب) نفس الهدف هو ٠,٩. فإذا صوب الشخصـــــان (أ،ب) على الهدف في نفس الوقت، احسب الاحتمالات التالية:

أ- إصابة الهدف.

ب- إصابة الهدف بواسطة شخص واحد فقط.

جـ- عدم إصابة الهدف بواسطة شخص واحد على الأقل.

د- احتمال عدم إصابة الهدف.

الحـل:

سوف نسمي الأحداث المختلفة على النحو التالي:

إصابة الهدف بواسطة الشخص (أ)	الحدث (أ)
إصابة الهدف بواسطة الشخص (ب)	الحدث (ب)
عدم إصابة الشخص (أ) للهـــــدف	الحدث ($\overline{أ}$)
عدم إصابة الشخص (ب) للهـــدف	الحدث ($\overline{ب}$)

هـــــــذا مـــــــــع ملاحظة أن كل حدثيْن ممثليْن في أزواج الأحداث: (أ،ب)، (أ،$\overline{ب}$)، ($\overline{أ}$،ب)، ($\overline{أ}$،$\overline{ب}$) إنما يمثلان حدثيْن غير متنافييْن ومســـتقليْن. وعليه يمكن حســـاب الاحتمالات المختلفة على النحو الموضح فيما يلي:

أ- الحالة التي يمكن أن يصاب فيها الهدف هي إحدى الحالات التالية:

- إصابة الهدف بواسطة الشخصيْن: وهي حالة واحدة فقط (أ ب)

• إصابة الهدف بواسطة شخص واحد فقط: وتتمثل في حالتيْن هما:
(ا $\overline{\text{ب}}$) أو ($\overline{\text{أ}}$ ب).

∴ الاحتمال المطلوب = ع(إصابة الهدف)

= ع(ا ب) + ع(ا $\overline{\text{ب}}$) + ع($\overline{\text{أ}}$ ب)

= ع(ا)ع(ب) + ع(ا)ع($\overline{\text{ب}}$) + ع($\overline{\text{أ}}$)ع(ب)

= ٠,٨ × ٠,٩ + ٠,٨ × ٠,١ + ٠,٢ × ٠,٩

= ٠,٧٢ + ٠,٠٨ + ٠,١٨ = ٠,٩٨

ويمكن حساب قيمة هذا الاحتمال باستخدام فكرة الاحتمال المكمِّل وذلك كما يلي:
احتمال إصابة الهدف = ١ – احتمال عدم إصابة الهدف (وهي حالة واحده فقط)

= ١ – ع($\overline{\text{أ}}$ $\overline{\text{ب}}$) = ١ – ٠,٢ × ٠,١ = ١ – ٠,٠٢ = ٠,٩٨

وهي نفس النتيجة التي حصلنا عليها من قبل.

ب - كما سبق أن أشرنا فإن هناك حالتيْن تفيد كل منهما بإصابة الهدف بواسطة شخص واحد فقط، وهاتان الحالتان هما: (ا $\overline{\text{ب}}$) أو ($\overline{\text{أ}}$ ب). أي أن:
احتمال إصابة الهدف بواسطة شخص واحد فقط

= ع(ا $\overline{\text{ب}}$) + ع($\overline{\text{أ}}$ ب)

= ٠,٨ × ٠,١ + ٠,٢ × ٠,٩ = ٠,٠٨ + ٠,١٨ = ٠,٢٦

جـ - عدم إصابة الهدف بواسطة شخص واحد على الأقل هي إحدى الحالات التالية:
($\overline{\text{أ}}$ ب)، (ا $\overline{\text{ب}}$)، ($\overline{\text{أ}}$ $\overline{\text{ب}}$)

أي أن:

الاحتمال المطلوب = ع($\overline{\text{أ}}$ ب) + ع(ا $\overline{\text{ب}}$) + ع($\overline{\text{أ}}$ $\overline{\text{ب}}$)

= ع($\overline{\text{أ}}$)ع(ب) + ع(ا)ع($\overline{\text{ب}}$) + ع($\overline{\text{أ}}$)ع($\overline{\text{ب}}$)

= ٠,٢ × ٠,٩ + ٠,٨ × ٠,١ + ٠,٢ × ٠,١

= ٠,١٨ + ٠,٠٨ + ٠,٠٢ = ٠,٢٨

ويمكن إيجاد قيمة هذا الاحتمال أيضاً باستخدام فكرة الاحتمال المكمل وذلك كما يلي:
احتمال عدم إصابة الهدف بشخص واحد على الأقل
= ١ – احتمال إصابة الهدف
= ١ – ٠,٨ × ٠,٩ = ١ – ٠,٧٢ = ٠,٢٨

د- عدم إصابة الهدف تتمثل في حالة واحدة فقط هي عدم إصابة كل من الشخصيْن أ،ب للهدف، أي الحالة ($\bar{أ}$ $\bar{ب}$):

∴ ع(عدم إصابة الهدف) = ع($\bar{أ}$ $\bar{ب}$)

= ٠,٢ × ٠,١ = ٠,٠٢

ويلاحَظ أن هذا الاحتمال هو المكمل للاحتمال المحسوب في أ.

مثال (١٥):

إذا كان احتمال فوز فريق النادي الاهلي لكرة القدم على فريق نادي الزمالك في يوم حار هو ٠,٩، وفي يوم معتدل الحرارة هو ٠,٧. فإذا كان من المقرر أن تقام مباراة بين الفريقيْن خلال شهر سبتمبر، وإذا علمـــت أن احتمال وجود يوم حار في شهر سبتمبر هو ٠,٦ وأن احتمال وجود يوم معتدل الحرارة هو ٠,٤ احسب الاحتمالات التالية:

أ- فوز فريق النادي الاهلي على فريق نادي الزمالك في هذه المباراة.

ب- إذا فاز الاهلي في هذه المباراة، احسب احتمال أن يكون اليوم الذي أقيمت فيه المباراة كان طقسه حارا.

الحــل:

سوف نسمي الأحداث المختلفة على النحو التالي:

فوز فريق النادي الاهلي	الحدث (أ)
وجـود يــوم حـــــــار	الحدث (ب)
وجود يوم معتدل الحرارة	الحدث (جـ)

وبذلك يمكن حساب الاحتمالات المختلفة على النحو التالي:

أ- المطلوب هو حساب الاحتمال: ع(أ)، ومن الواضح أن فوز النادي الاهلي يمكن أن يتأتى في إحدى الحالتيْن التاليتيْن.

- أن يكون اليوم حاراً ويفوز فيه النادي الاهلي. أو:
- أن يكون اليوم معتدل الحرارة ويفوز فيه النادي الاهلي أيضاً.

ويمكن ترجمة ذلك في صورة احتمالات على النحو التالي:

ع(أ) = ع(ب)ع(أ | ب) + ع(جـ)ع(أ | جـ)

= ٠,٦ × ٠,٩ + ٠,٤ × ٠,٧

= ٠,٥٤ + ٠,٢٨ = ٠,٨٢

ب- الاحتمال المطلوب هنا هو: ع(اليوم حار | فوز الاهلي)،أي حساب الاحتمال ع(ب | أ).

وحيث أن: $ع(ب \mid أ) = \frac{ع(ب أ)}{ع(أ)}$

$$= \frac{ع(ب)ع(أ|ب)}{ع(ب)ع(أ|ب) + ع(جـ)ع(أ|جـ)} \qquad (١ - ٩)$$

$$= \frac{٠,٦ \times ٠,٩}{٠,٦ \times ٠,٩ + ٠,٤ \times ٠,٧} = \frac{٠,٥٤}{٠,٨٢} = ٠,٦٥٩$$

هذا ويلاحظ أن الصورة (١ – ٩) هي ما يعرف بصيغة بيز Bayes Formula.

مثال (١٦):

صندوقان ($ص_١$، $ص_٢$)، الأول به ٤ كرات بيضاء و ٦ سوداء والثاني به ٣ كرات بيضاء، ٢ سوداء، ١٠ خضراء. تم اختيار أحد الصندوقيْن عشوائياً ثم سحبت منه كرة، احسب الاحتمالات التالية:

١- أن تكون الكرة المسحوبة سوداء.

٢- إذا كانت الكرة المسحوبة بيضاء، احسب احتمال أن تكون مسحوبة من الصندوق $ص_٢$ علماً بأن الصندوقيْن $ص_١$، $ص_٢$ لهما نفس الفرصة في الاختيار.

الحــل:

سوف نسمي الأحداث المختلفة على النحو التالي:

اختيار الصندوق $ص_١$	الحدث ($ص_١$)
اختيار الصندوق $ص_٢$	الحدث ($ص_٢$)
الكـــرة بيضــــــاء	الحدث (ب)
الكـــرة ســـــــوداء	الحدث (س)
الكـــرة خضــــــراء	الحدث (غ)

وبذلك يمكن حساب الاحتمالات المختلفة على النحو التالي:

١- الاحتمال المطلوب هو: ع(س)، ويمكن أن تكون الكرة المسحوبة سوداء في إحدى الحالتيْن التاليتيْن:

- اختيار الصندوق $ص_١$ وتكون الكرة المسحوبة منه سوداء.
- اختيار الصندوق $ص_٢$ وتكون الكرة المسحوبة منه سوداء.

ويمكن التعبير عن ذلك في صورة احتمالات كما يلي:

$$ع(س) = ع(ص_١)ع(س \mid ص_١) + ع(ص_٢)ع(س \mid ص_٢)$$

$$= \frac{١}{٢} \times \frac{٦}{١٠} + \frac{١}{٢} \times \frac{٢}{١٥} = ٠,٣٦٧$$

٢- الاحتمال المطلوب هو: ع(اختيار الصنـدوق $ص_١$ | الكرة المسحوبة بيضاء)، أي الاحتمال:

$ع(ص_٢ \mid ب)$. ويمكن حسابه كما يلي:

$$ع(ص_٢ \mid ب) = \frac{ع(ص_٢ ب)}{ع(ب)}$$

$$= \frac{ع(ص_٢)ع(ب \mid ص_٢)}{ع(ب)}$$

وحيث أن:

$$ع(ب) = ع(ص_١)ع(ب \mid ص_١) + ع(ص_٢)ع(ب \mid ص_٢)$$

$$\therefore ع(ص_٢ \mid ب) = \frac{ع(ص_٢)ع(ب \mid ص_٢)}{ع(ص_١)ع(ب \mid ص_١) + ع(ص_٢)ع(ب \mid ص_٢)}$$

$$= \frac{\frac{١}{٢} \times \frac{٣}{١٥}}{\frac{١}{٢} \times \frac{٤}{١٠} + \frac{١}{٢} \times \frac{٣}{١٥}} = \frac{٠,١}{٠,٣} = ٠,٣٣٣$$

وبعد أن انتهينا من تناولنا لموضوع الاحتمالات سوف ننتقل الآن إلى دراسة المتغيرات العشوائية والتوزيعات الاحتمالية وذلك كتمهيد لدراستنا للاستدلال الإحصائي في الفصل الثاني من هذا الكتاب.

المتغيرات العشوائية والتوزيعات الاحتمالية

تناولنا بإيجاز شديد في مستهل هذا الفصل بعض مبادئ نظرية الاحتمالات. ومع التطور السريع في استخدام هذه النظرية واتساع مجالات تطبيقاتها، وبالإضافة إلى ظهور الحاجة الملحة إلى الاحتمالات كأسلوب علمي في حل الكثير من المشكلات المعقدة، برز الاتجاه إلى تطوير نظرية الاحتمالات كفرع من فروع الرياضيات وكجانب مهم من جوانب النظرية الإحصائية وتطبيقاتها في شتى المجالات.

وتقوم نظرية الاحتمالات أساساً على المتغيرات العشوائية وذلك في محاولة لتحديد التوزيعات الرياضية لها ثم استخدام هذه التوزيعات في تقدير المقاييس الإحصائية المختلفة لهذه المتغيرات سواء كانت هذه المقاييس متعلقة بالنزعة المركزية أو التشتت مطلقاً كان أم نسبياً، هذا بالإضافة إلى الاحتمالات المتعلقة بالقيم المختلفة التي يمكن أن تأخذها تلك المتغيرات. هذا واستخدام نظرية الاحتمالات يمكِّننا من التوصل إلى الشكل الرياضي للتوزيع الذي يحكم الظاهرة موضوع الدراسة والذي يكون أكثر دقةً وأيسر تطبيقاً من التوزيعات التكرارية - بسيطة كانت أم متجمعة - والتي سبق الإشارة اليها والتي تسمى بالتوزيعات التجريبية Empirical Distributions. هذا ويُطلق على التوزيعات في شكلها الرياضي التوزيعات الاحتمالية Probability Distributions للمتغيرات العشوائية Random Variables. وبصورة أكثر تحديداً يمكننا القول إن التوزيعات التجريبية تتعلق بالإحصاء الوصفي، وأما التوزيعات الرياضية فإنها تتعلق بالإحصاء الاستدلالي.

وللوصول إلى تعريف شامل وواضح للمتغير العشوائي وكيفية قياس الاحتمالات المناظرة للقيم المختلفة التي يأخذها المتغير. دعنا نفترض أننا نجري تجربة عشوائية متمثلة في رمي زهرة نرد عدداً من المرات ثم تسجيل الرقم الذي يظهر على الوجه في كل مرة. من الواضح أن نتائج هذه التجربة تتمثل في ظهور الأرقام ١، ٢، ٣، ٤، ٥، ٦. وهذا يعنى أنه يمكن القول إن نتيجة التجربة تعتبر متغيراً (س) يأخذ وبشكل عشوائي قيمة من القيم ١ إلى ٦. ولذلك تسمى س بالمتغير العشوائي، وحيث أن س تأخذ القيم الصحيحة من ١ إلى ٦ فإن س هنا تمثل متغيراً عشوائياً منفصلاً. ولو كان الأمر متعلقاً برمي زهرة نرد مرة واحدة لما كانت هناك مشكلة في إيجاد احتمال ظهور رقم معين، حيث أنه من المعروف أن احتمال ظهور رقم معين في هذه الحالة يساوي $\frac{١}{٦}$ دائماً. ولكن لو أن الأمر يتعلق برمي زهرة نرد عدداً من المرات قد يزيد على المائة لأصبحت عملية إيجاد ظهور رقم معين عدداً محدداً من المرات - متبعين في ذلك الطرق التقليدية والمفاهيم البسيطة لنظرية الاحتمالات - أمراً غايةً في التعقيد. ويمكننا التغلب على تلك المشكلة بتحديد التوزيع الاحتمالي الذي يتبعه المتغير موضع الاعتبار (من المعروف أن التوزيع الاحتمالي هنا هو توزيع ثنائي الحدّيْن) والذي سنتعرض له فيما بعد بشيء من التفصيل وذلك في موضع لاحق من هذا الفصل.

والآن لعله من المناسب أن نتعرض في هذا الشأن لبعض التعريفات والمفاهيم وذلك على النحو الموضح فيما يلي:

أولاً: بعض التعريفات والمفاهيم

سوف نتناول فيما يلي بعض التعريفات والمفاهيم المتعلقة بالمتغيرات العشوائية والتوزيعات الاحتمالية:

١- المتغير العشوائي: Random Variable

هو المتغير الذي يأخذ وبطريقة عشوائية قيماً حقيقية، أي هو المتغير الذي تتحكم الصدفة في تحديد قيمته.

٢- المتغير العشوائي المنفصل: Discrete Random Variable

هو المتغير الذي يأخذ عدداً محدوداً أو متميزاً من القيم، وهذه القيم تكون صحيحة وليست كسرية. وكمثال على ذلك عدد أفراد الأسرة والذي لا يكون إلا عدد صحيحاً.

٣- المتغير العشوائي المتصل: Continuous Random Variable

هو المتغير الذي يأخذ عدداً لا نهائياً من القيم داخل أي فترة أو مدى معلوم. وهذا يعني أنه يمكن أن يأخذ أية قيمة صحيحة كانت أم كسرية. ومثال ذلك الأوزان والأطوال لمجموعة من الأشخاص.

٤- التوزيع التكراري التجريبي: Empirical Distribution

يقصد به القيم المختلفة التي يمكن أن يأخذها المتغير العشوائي والاحتمالات المناظرة لها والمحسوبة من واقع نتائج إجراء تجربة معينة عدداً من المرات. والاحتمالات هنا تشير إلى النسبة بين عدد المرات التي يحدث فيها ناتج معين إلى إجمالي عدد المحاولات أو المشاهدات الفعلية.

٥- التوزيع الاحتمالي: Probability Distribution

ويقصد به الاحتمالات الكلاسيكية المناظرة لكل القيم التي يمكـــن أن يأخذها المتغير العشوائي. وهذه الاحتمالات يتم تحديدها مقدماً وبدون إجراء أية تجارب، ولذلك فإنه عادة ما يسمى هذا التوزيع بالتوزيع التكراري النظري النسبي. وفي هذا فإنه يختلف عن التوزيع التكراري التجريبي والذي لا يمكن في حالته تحديد الاحتمالات المختلفة إلا بعد إجراء التجربة بالفعل.

هذا ولتوضيح الفارق بين التوزيع التكراري التجريبي والتوزيع الاحتمالي لمتغير عشوائي دعنا نضرب المثال التالي:

لنفترض أننا قمنا برمي زهرة نرد ستون مرة وكانت نتائج التجربة كالتالي:

الرقم الظاهر على الوجه:	١	٢	٣	٤	٥	٦
عدد مرات الظهـــــور:	١٢	٩	١١	٨	٦	١٤

وهنا تجدر الإشارة إلى أنه وفقاً للتوزيع الاحتمالي فإن عدد مرات ظهور رقم معين يجب أن يساوي بالضبط $\frac{١}{٦}$ إجمالي عدد مرات الرمي وهو عشر مرات لكل رقم. حيث أن احتمال ظهور رقم معين في كل رمية يساوي $\frac{١}{٦}$ دائماً. وهذا أمر غير متوقَّع حدوثه في حالة التوزيع

التكراري التجريبي وذلك كما تظهره النتائج المشار اليها حيث يختلف عدد مرات الظهور من رقم لآخر. ومع ذلك فإنـــه مع تزايد عدد مرات الرمي فإن التوزيع التكراري التجريبي النسبي يقترب أكثر فأكثر من التوزيع الاحتمالي (التوزيع النظري النسبي) والذي يساوي دائماً القيمة $\frac{١}{٦}$.

٦- التوزيع الاحتمالي المتجمع: Cumulative Probability Distribution

وهو التوزيع الاحتمالي الذي يوضح الاحتمالات المناظرة للقيم المختلفة التي يقل عنها أو يساويهـا المتغير العشوائي، أي: ح(س ≤ $س_ر$). وهذا التوزيع يناظره في الحالة التجريبية التوزيع التكراري المتجمع الصاعد النسبي.

٧- دالة كثافة الاحتمال: Probability Density Function

وهي الدالة ذات الشكل الرياضي والتي يتم عن طريقها تحديد الاحتمالات المناظرة للقيم المختلفة التي يأخذها المتغير العشوائي. أي هي الدالة الرياضية التي تحدد التوزيع الاحتمالي للمتغير العشوائي. وسوف نرمز لهذه الدالة بالرمز ف(س) في حالة المتغير المتصل وبالرمز ح(س) في حالة المتغير المنفصـل، حيث تشير ح($س_ر$) إلى احتمال أن يأخذ المتغير المنفصل القيمة $س_ر$، أي أن: ح(س = $س_ر$) = ح($س_ر$).

٨- دالة التوزيع الاحتمالي: Distribution Function

وهي الدالة الرياضية والتي يمكن بواسطتها تحديد التوزيع الاحتمالي المتجمع للمتغير العشوائي، أي التي تحدد احتمال أن يأخذ المتغير العشوائي قيمة أقل من أو تساوي مقداراً معيناً. وسوف نرمز لهذه الدالة بالرمز د(س) سواء في حالة المتغيرات المنفصلة أو المتغيرات المتصلة، أي أن:

د($س_ر$) = ح(س ≤ $س_ر$)

= ح(س = $س_١$) + ح(س = $س_٢$) +....... + ح(س = $س_ر$)

وفيما يلي بعض الأمثلة التي تلقي مزيداً من الضوء على التعريفات والمفاهيم التي أشرنا إليها وذلك في كل من حالتي المتغيرات المنفصلة والمتغيرات المتصلة.

أ- المتغيرات العشوائية المنفصلة:

مثال (١٧):

اوجد التوزيع الاحتمالي والتوزيع الاحتمالي التجميعي لعدد مرات ظهور الوجه الذي عليه الشعار (الصورة) وذلك عند رمي قطعة عملة ثلاث مرات (يلاحظ أن رمي قطعة عُملة

ثلاث مرات أو رمي ثلاث قطع عُملة مرة واحدة هما نفس الشيء وذلك من حيث النتائج المتحصَّل عليها).

الحـل:

عدد مرات ظهور الوجه الذي عليه الشعار عند رمي قطعة عملة ثلاث مرات هو متغير عشوائي منفصل سنرمز له بالرمز س. وباستعراض جميع الحالات الممكن ظهورها، وبافتراض أن ص ترمز إلى ظهور الوجه الذي عليـه الشعـار (الصورة)، ك ترمز إلى ظهور الوجه الذي عليه الكتابة سوف نحصل على ما يلي:

الحالـــــة	١	٢	٣	٤
نتيجة التجربـــة	ص ص ص	ص ص ك	ص ك ص	ك ص ص
قيمة المتغير ($س_{ر}$)	٣	٢	٢	٢
الحالـــــة	٥	٦	٧	٨
نتيجة التجربـــة	ك ص ك	ك ك ص	ص ك ك	ك ك ك
قيمة المتغير ($س_{ر}$)	١	١	١	صفر

من هذا يتضح أن التوزيع الاحتمالي والتوزيع الاحتمالي التجميعي لعدد مرات ظهور الوجه الذي عليه الشعار يمكن تحديدهما على النحو التالي:

$س_{ر}$	$س_{١}$	$س_{٢}$	$س_{٣}$	$س_{٤}$
قيمة المتغير ($س_{ر}$)	صفر	١	٢	٣
ع($س_{ر}$)	$\frac{١}{٨}$	$\frac{٣}{٨}$	$\frac{٣}{٨}$	$\frac{١}{٨}$
د($س_{ر}$)	$\frac{١}{٨}$	$\frac{٤}{٨}$	$\frac{٧}{٨}$	١

وكما هو واضح فإن مجموع الاحتمالات في التوزيع الاحتمالي يساوي الواحد الصحيح. كما يتضح لنا أيضاً ما يلي:

ع(س = صفر) = ع(س = ٣) = $\frac{١}{٨}$

ع(س = ١) = ع(س= ٢) = $\frac{٣}{٨}$

د($س_{٢}$) = ع(س ≤ $س_{٢}$) = ع(س ≤ ١) = د (١) = $\frac{٤}{٨}$

د($س_{٣}$) = ع(س ≤ $س_{٣}$) = ع(س ≤ ٢) = د (٢) = $\frac{٧}{٨}$

مثال (١٨):

الجدول التالي يوضح التوزيع التكراري لعشرين أسرة حسب أحجامها المختلفة:

حجم الأسرة ($س_ر$)	١	٢	٣	٤	٥	٦
عـــدد الأســـر	٢	٣	٧	٤	٣	١

أوجد التوزيع الاحتمالي والتوزيع الاحتمالي المتجمع ثم حدد قيمة كل من الاحتمالات التالية:

- احتمال أن حجم الأسرة يساوي أربعة أفراد.
- احتمال يبلغ حجم الأسرة ثلاثة أفراد على الأكثر.
- احتمال أن يزيد حجم الأسرة على أربعة أفراد.

الحـــل:

حجم الأسرة (س) هو متغير عشوائي منفصل يأخذ القيم الصحيحة ابتداءً من واحد إلى ستة أفراد. ويمكن تكوين التوزيع الاحتمالي والتوزيع الاحتمالي المتجمع وذلك على النحو الموضح فيما يلي:

حجم الأسرة ($س_ر$)	١	٢	٣	٤	٥	٦
التكـــــرارات	٢	٣	٧	٤	٣	١
ع($س = س_ر$)	٠,١٠	٠,١٥	٠,٣٥	٠,٢٠	٠,١٥	٠,٠٥
ع($س \leq س_ر$)	٠,١٠	٠,٢٥	٠,٦٠	٠,٨٠	٠,٩٥	١,٠٠

ومن التوزيع الاحتمالي والتوزيع الاحتمالي التجميعي يمكننا الحصول على ما يلي:

$$ع(س = ٤) = ٠,٢٠$$

$$ع(س \leq ٣) = د(٣) = ٠,٦٠$$

$$ع(س > ٤) = ع(س = ٥) + ع(س = ٦) = ٠,١٥ + ٠,٠٥ = ٠,٢٠$$

أو: $$ع(س > ٤) = ١ - ع(س \leq ٤) = ١ - ٠,٨٠ = ٠,٢٠$$

ب- المتغيرات العشوائية المتصلة:

من المعروف أنه في حالة المتغيرات المتصلة لا يمكن حساب احتمال أن يأخذ المتغير قيمة واحدة محددة وذلك لأن المتغير المتصل يمكن أن يأخذ عدداً غير محدود (لا نهائي) من القيم داخل مدى معين مهما كان هذا المدى صغيراً في قيمته العددية. هذا ومن الممكن حساب

احتمال أن يقع المتغير المتصل داخل فترة معينة، أي احتمال أن يأخذ المتغير المتصل قيمة تتراوح بين حد أدنى وحد أعلى مرغوبٌ فيهما. فإذا أردنا حساب احتمال أن يأخذ المتغير س قيماً بين القيمتيْن $\text{س}_{١}$، $\text{س}_{٢}$ فإن ذلك يتم بحساب التكامل المحدود التالي:

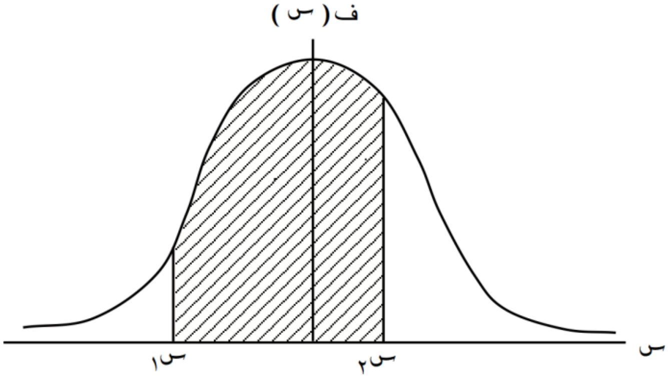
$$\text{ع}(\text{س}_{١} \leq \text{س} \leq \text{س}_{٢}) = \int_{\text{س}_{١}}^{\text{س}_{٢}} \text{ف}(\text{س})\, \text{ءس}$$

أي أن الاحتمال في هذه الحالة يساوي قيمة المساحة تحت منحنى التوزيع والمحصورة بين القيمتيْن أ، ب ويمثل ذلك بيانياً بالشكل (١ – ٢).

ولو افترضنا أن أصغر قيمة يمكن أن يأخذها المتغير هي ل وأن أكبر قيمة يمكن أن يأخذها هي ك فإن قيمة المساحة تحت منحنى التوزيع والمحصورة بين القيمتين ل، ك يجب أن تساوي الواحد الصحيح وذلك لأن مجموع الاحتمالات في التوزيع الاحتمالي يساوي دائماً الواحد الصحيح.

وبصفة عامة فإن المدى الذي يأخذه المتغير المتصل عادة ما يعرَّف بالقيمتيْن - ∞ ، ∞.

أي أن:



الشكل (١ – ٢)

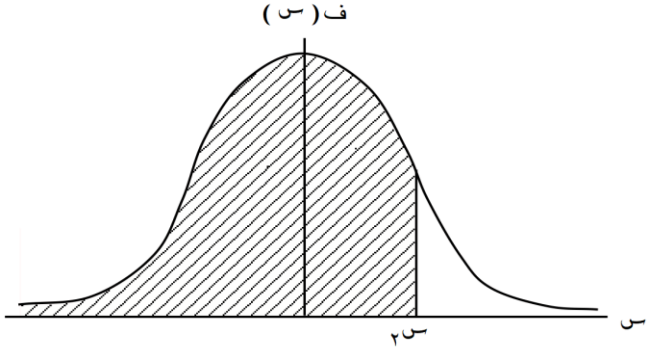
$$\int_{-\infty}^{\infty} \text{ف}(\text{س})\, \text{ءس} = ١$$

وهي من خصائص دالة كثافة الاحتمال.

هذا وبمعلومية دالة كثافة الاحتمال يمكن إيجاد دالة التوزيع الاحتمالي وذلك على النحو التالي:

$$\text{د}(\text{س}_{١}) = \int_{-\infty}^{\text{س}_{١}} \text{ف}(\text{س})\, \text{ءس}$$

ويمثِّل ذلك بيانيا الشكل (١ – ٣).



الشكل (١ – ٣)

ويلاحَظ في هذا الشكل ان $س_١$ هي قيمة محددة للمتغير س ، د($س_١$) هي المساحة تحت المنحنى على يسار العمود المقام عند النقطة $س_١$.

وباستخدام دالة التوزيع الاحتمالي يمكن حساب احتمال أن يأخذ المتغير قيماً تتراوح بين حدّيْن أدنى وأعلى مرغوب فيهما ($س_١$، $س_٢$ على سبيل المثال) وذلك كما يلي:

$$ح(س_١ \le س \le س_٢) = \int_{س_١}^{س_٢} ف(س)\, ءس$$

$$= \int_{-\infty}^{س_٢} ف(س)\, ءس - \int_{-\infty}^{س_١} ف(س)\, ءس$$

$$= د(س_٢) - د(س_١)$$

ويمثِّل ذلك بيانياً المساحة المظللة تحت المنحنى والمحصورة بين العموديْن المقاميْن عند النقطتيْن $س_١$، $س_٢$ (راجع الشكل [١ – ٢])

ملحوظة: في حالة المتغير المتصل (س) ، الاحتمال ح(س= أ) يساوي الصفر دائما لأي قيمة لـ س.

مثال (١٩):

إذا كانت س متغير عشوائي يمكن أن يأخذ أي قيمة داخل المدى من ١ إلى ١١ باحتمالات متساوية، المطلوب:

أ- إيجاد دالة كثافة الاحتمال ب- إيجاد دالة التوزيع الاحتمالي

جـ- ح(س ≤ ٩) د- ح(٤ ≤ س ≤ ٧)

الحــل:

أ- حيث ان المتغير س يمكن أن يأخذ أية قيمة داخل مدى محدود فإن س هي متغير متصل.

وحيث أن: $\int_{١}^{١١}$ ف(س) ءس = ١

حيث: ف(س) مقدار ثابت في هذه الحالة ولنفترض أنه يساوي أ

∴ $\int_{١}^{١١}$ أ ءس = أ $\int_{١}^{١١}$ ءس = أ $[\text{س}]_{١}^{١١}$ = ١

أي أن: أ(١١ - ١) = ١

ومنها نحصل على: أ = $\frac{١}{١٠}$

أي أن: ف(س) = $\frac{١}{١٠}$ وهي دالة كثافة الاحتمال للمتغير س

ب - حيث أن: د(س) = $\int_{-\infty}^{\text{س}}$ ف(س) ءس = $\int_{١}^{\text{س}}$ $\frac{١}{١٠}$ ءس

= $\frac{١}{١٠}$ $\int_{١}^{\text{س}}$ ءس = $\frac{١}{١٠}$ $[\text{س}]_{١}^{\text{س}}$ = $\frac{١}{١٠}$(س - ١)

أي أن: د(س) = $\frac{\text{س} - ١}{١٠}$ وهي دالة التوزيع الاحتمالي للمتغير س.

جـ - ع(س ≤ ٩) = د(٩)

وحيث أن: د(س) = $\frac{\text{س} - ١}{١٠}$

∴ ع(س ≤ ٩) = د(٩) = $\frac{٩ - ١}{١٠}$ = ٠,٨

د- حيث أن: ع($\text{س}_١$ ≤ س ≤ $\text{س}_٢$) = د($\text{س}_٢$) – د($\text{س}_١$)

∴ ع(٤ ≤ س ≤ ٧) = د (٧) – د (٤)

= $\frac{٧ - ١}{١٠}$ – $\frac{٤ - ١}{١٠}$

= ٠,٦ - ٠,٣ = ٠,٣

هذا وتجدر الإشارة إلى أنه يمكننا إيجاد التوزيع الاحتمالي للمتغير المتصل وذلك من واقع التوزيع التكراري للمتغير متَّبعين في ذلك نفس الأسلوب المستخدم في حالة المتغير المنفصل، مع مراعاة أن التكرار النسبي في هذه الحالة يعطى احتمال أن يقع المتغير داخل فئة معينة (أي احتمال أن يأخذ المتغير قيمة تقع بين الحد الأدنى والحد الأعلى لهذه الفئة).

والمثال التالي يوضح ذلك:

مثال (٢٠):

الجدول التالي يوضح توزيع ٥٠ طالباً في إحدى كليات التجارة موزَّعين حسب درجاتهم في مادة مبادئ الإحصاء.

الدرجـــــات	٤٠-	٥٠-	٦٠-	٧٠-	٨٠-	٩٠-١٠٠	المجموع
عدد الطلاب	٦	٩	١٢	١٠	٨	٥	٥٠

المطلــوب:

أ- ع(٧٠ ≤ س < ٨٠)　　　　ب- ع(٥٠ ≤ س < ٧٠)

جـ- هل يمكن تحديد الشكل الرياضي لدالة كثافة الاحتمال؟

الحــل:

في هذا المثال نجد أن درجة الطالب في مادة مبادئ الإحصاء تمثل متغيراً عشوائياً متصلاً، ويمكن تحديد التوزيع الاحتمالي لهذا المتغير بإيجاد التوزيع التكراري النسبي له والذي يوضحه الجدول (١ - ١).

الجدول (١ - ١)

التوزيع التكراري النسبي (التوزيع الاحتمالي) والتوزيع الاحتمالي التجميعي لدرجات ٥٠ طالباً في مادة الإحصاء

ر	١	٢	٣	٤	٥	٦
الدرجة (س)	٤٠ -	٥٠ -	٦٠ -	٧٠ -	٨٠ -	٩٠-١٠٠
ع(سر ≤ س < سر + ١)	٠,١٢	٠,١٨	٠,٢٤	٠,٢٠	٠,١٦	٠,١٠
ع(س < سر + ١)	٠,١٢	٠,٣٠	٠,٥٤	٠,٧٤	٠,٩٠	١,٠٠

من الجدول (١ - ١) يتضح ما يلي:

أ- ع(٧٠ ≤ س < ٨٠) = ٠,٢

ب- ع(٥٠ ≤ س < ٧٠) = ع(س < ٧٠) - ع(س < ٥٠)

= ٠,٥٤ - ٠,١٢ = ٠,٤٢

جـ- دالة كثافة الاحتمال للمتغير س في هذا المثال بالرغم من عدم تعريفها رياضياً بشكل محدد إلا انها يجب أن تحقق العلاقات التالية:

$$\int_{٤٠}^{٥٠} ف(س)\,ءس = ٠,١٢ \quad ، \quad \int_{٥٠}^{٦٠} ف(س)\,ءس = ٠,١٨$$

وهكذا حتى نصل إلى: $\int_{٩٠}^{١٠٠} ف(س)\,ءس = ٠,١$

كما أن: $\int_{٤٠}^{١٠٠} ف(س)\,ءس = ١$

ثانياً: بعض المقاييس الإحصائية

كما سبق وأن ذكرنا فإن الأساليب الإحصائية تعتمد أساساً في عمليات التحليل على بيانات العينات وذلك في محاولة للتعرف على خصائص المجتمعات التي سُحبت منها هذه العينات. هذا وتهدف عمليات التحليل في المقام الأول إلى تقدير ما يسمى بمعلمات المجتمع مستخدمة في ذلك البيانات المتاحة عن العينات. وهذه المعلمات هي المؤشرات التي تحدد خصائص المجتمع المراد التعرف عليها. ومن أهم هذه المعلمات مقاييس المتوسطات ومقاييس التشتت والتي سنتعرض لتوضيح مفهوم كل منها على النحو التالي:

١- مقاييس المتوسطات:

أ- القيمة المتوقَّعة: Expected Value

أوضحنا فيما سبق أن المتغير العشوائي يمكن أن يأخذ قيما متعددة باحتمالات متناظرة. وفي كثير من الأحيان قد يكون المراد هو معرفة القيمة التي يأخذها المتغير في المتوسط والتي تسمى بالقيمة المتوقعة للمتغير (الوسط الحسابي). ويمكن تعريف القيمة المتوقَّعة للمتغير س والتي سنرمز لها بالرمز: ت(س) وذلك كما يلي:

في حالة المتغير المنفصل:

$$ت(س) = \bar{س} = \sum_{ر=١}^{ن} س_ر \times ع(س_ر)$$

حيث تُجرى عملية الجمع بالنسبة لجميع القيم التي يمكن أن يأخذها المتغير العشوائي مضروبة في احتمالاتها المناظرة.

في حالة المتغير المتصل:

$$ت(س) = \bar{س} = \int س ف(س)\,ءس$$

ويُجرى التكامل على المدى الذي يأخذه المتغير العشوائي أي على جميع القيم التي يمكن أن يأخذها.

وفيما يلي بعض الأمثلة التي توضح كيفية إيجاد القيمة المتوقَّعة للمتغير العشوائي منفصلاً كان أم متصلاً:

مثال (٢١):

احسب القيمة المتوقَّعة لعدد مرات ظهور الوجه الذي عليه الشعار (الصورة) وذلك عند رمي قطعة عملة ثلاث مرات.

الحـــل:

لنفترض أن س هي المتغير الذي يمثل عدد المرات التي يظهر فيها الوجه الذي عليه الشعار. وحيث أن القيم المختلفة للمتغير س والاحتمالات المناظرة لها قد تم حسابها في مثال (١٧) وكانت كما يلي:

قيمة المتغير $(س_ر)$	صفر	١	٢	٣
$ع(س_ر)$	$\frac{١}{٨}$	$\frac{٣}{٨}$	$\frac{٣}{٨}$	$\frac{١}{٨}$

فإنه يمكن حساب القيمة المتوقعة كما يلي:

$$ت(س) = \sum س_ر \; ع(س_ر)$$

$$= \text{صفر} \times \frac{١}{٨} + ١ \times \frac{٣}{٨} + ٢ \times \frac{٣}{٨} + ٣ \times \frac{١}{٨}$$

$$= ١,٥$$

مثال (٢٢):

احسب القيمة المتوقَّعة للرقم الذي يظهر عند رمي زهرة نرد.

الحــل:

المتغير هنا يمثل الرقم الذي يظهر عند رمي زهرة نرد وهو متغير عشوائي منفصل، وحيث أن كلاً من الأرقام ١، ٢،...، ٦ يظهر باحتمال $\frac{١}{٦}$ فإنه يمكن حساب القيمة المتوقعة كما يلي:

$$ت(س) = \sum س_ر \times ع(س_ر)$$

$$= \text{صفر} \times \frac{١}{٦} + ١ \times \frac{١}{٦} + ...+ ٦ \times \frac{١}{٦}$$

$$= \frac{١}{٦}(١ + ٢ + ٣ + ...+ ٦) = ٣,٥$$

مثال (٢٣):

إذا كانت دالة كثافة الاحتمال للمتغير العشوائي س هي:

$$ف(س) = \begin{cases} \frac{٢}{٩}(٣س - س^٢) & \text{صفر} \leq س \leq ٣ \\ \text{صفر} & \text{خلاف ذلك} \end{cases}$$

احسب القيمة المتوقَّعة للمتغير س.

الحـــل:

المتغير س هنا هو متغير عشوائي متصل يمكن حساب توقُّعه كالآتي:

$$ت(س) = \int_{٠}^{٣} س \times \frac{٢}{٩}(٣س - س^٢)\, ءس$$

$$= \frac{٢}{٩}\int_{٠}^{٣}(٣س^٢ - س^٣)\, ءس$$

$$= \frac{٢}{٩}\left[\frac{٣س^٣}{٣} - \frac{س^٤}{٤}\right]_{٠}^{٣}$$

$$= \frac{٢}{٩}\left(\frac{٣ \times ٢٧}{٣} - \frac{٨١}{٤}\right) = ١,٥$$

خصائص القيمة المتوقَّعة:

فيما يلي نتناول خصائص القيمة المتوقَّعة وذلك على النحو التالي:

١- توقع المقدار الثابت = المقدار الثابت

أي أنه إذا كانت أ مقدار ثابت فإن:

$$ت(أ) = أ$$

وهذا أمر طبيعي إذ أننا في هذه الحالة لا نكون أمام متغير عشوائي. وبالتالي فإنه عند اجراء التجربة نتوقع أن نحصل على نفس القيمة الثابتة أياً كان عدد مرات التجربة.

٢- توقُّع حاصل جمع (متغير + مقدار ثابت) = توقع المتغير + المقـدار الثابت
أي أن:
ت(س + أ) = ت(س) + أ حيث أ مقدار ثابت

٣- توقُّع حاصل ضرب مقدار ثابت × متغير = المقدار الثابت × توقـــع المتغير
أي أن: ت(أ س) = أ × ت(س)

٤- توقُّع حاصل جمع (أو طرح) متغيريْن عشوائييْن = حاصل جمع (أو طرح) توقعهما،
أي أن: ت(س ± ص) = ت(س) ± ت(ص)
حيث س، ص متغيران عشوائيان.

٥- إذا كانت س، ص متغيران عشوائيان مستقلان فإن:
ت(س ص) = ت(س) × ت(ص)
أي أن توقُّع حاصل ضرب متغيريْن عشوائييْن مستقليْن يساوي حاصل ضرب توقعهما.

مثال (٢٤):

باستخدام دالة كثافة الاحتمال المعطاة في المثال السابق رقم (٢٣)، احسب ما يلي:
أ- ت(س + ١) ب- ت(٢س)

الحـــل:

أ- حيـث أن: ت(س + ١) = ت(س) + ١
وحيث أن: ت(س) = ١,٥ من المثال السابق
∴ ت(س + ١) = ت(س) + ١
= ١,٥ + ١ = ٢,٥
ويمكن حساب هذا التوقع بطريقة أخرى كما يلي:

$$\text{ت}(\text{س} + ١) = \int_{٠}^{٣} \frac{٢}{٩}(\text{س} + ١)(٣\text{س} - \text{س}^{٢})\, ء\text{س}$$

$$= \frac{٢}{٩}\int_{٠}^{٣} (٣\text{س} + ٢\text{س}^{٢} - \text{س}^{٣})\, ء\text{س}$$

$$= \frac{٢}{٩}\left[\frac{٣\text{س}^{٢}}{٢} + \frac{٢\text{س}^{٣}}{٣} - \frac{\text{س}^{٤}}{٤}\right]_{٠}^{٣}$$

$$= \frac{٢}{٩}\left(\frac{٣ \times ٩}{٢} + \frac{٢ \times ٢٧}{٣} - \frac{٨١}{٤}\right) = ٢,٥$$

ب- حيث أن: ت(أ س) = أ× ت(س)

∴ ت(٢س) = ٢ت(س) = ٢ × ١,٥ = ٣

ويمكن أيضاً حساب قيمة هذا التوقُّع على النحو التالي:

$$\text{ت}(\text{٢س}) = \int_{\text{٠}}^{\text{٣}} \text{٢س} \times \frac{\text{٢}}{\text{٩}}(\text{٣س} - \text{س}^{\text{٢}})\, \text{ءس}$$

$$= \frac{\text{٤}}{\text{٩}} \int_{\text{٠}}^{\text{٣}} (\text{٣س}^{\text{٢}} - \text{س}^{\text{٣}})\, \text{ءس}$$

$$= \frac{\text{٤}}{\text{٩}} \left[\frac{\text{٣س}^{\text{٣}}}{\text{٣}} - \frac{\text{س}^{\text{٤}}}{\text{٤}}\right]_{\text{٠}}^{\text{٣}}$$

$$= \frac{\text{٤}}{\text{٩}} \left(\frac{\text{٣} \times \text{٢٧}}{\text{٣}} - \frac{\text{٨١}}{\text{٤}}\right) = \text{٣}$$

وهي نفس النتيجة التي حصلنا عليها من قبل.

مثال (٢٥):

إذا كان التوزيع الاحتمالي للمتغير س هو كما يلي:

س	١	٢	٣
ع(س)	$\frac{\text{١}}{\text{٤}}$	$\frac{\text{١}}{\text{٤}}$	$\frac{\text{١}}{\text{٢}}$

احسب ما يلي:

أ- ت(س)   ب- ت(س - ٢)

الحـل:

أ- المتغير س في هذا المثال هو متغير منفصل يمكن حساب توقعه كما يلي:

$$\text{ت}(\text{س}) = \sum \text{س} \times \text{ع}(\text{س})$$

$$= \text{١} \times \frac{\text{١}}{\text{٤}} + \text{٢} \times \frac{\text{١}}{\text{٤}} + \text{٣} \times \frac{\text{١}}{\text{٢}} = \text{٢,٢٥}$$

ب- ت(س - ١) = ت(س) - ١

= ٢,٢٥ - ١ = ١,٢٥

ويمكن حساب ذلك بطريقة أخرى كما يلي:

س - ١	صفر	١	٢
ع(س)	$\frac{١}{٤}$	$\frac{١}{٤}$	$\frac{١}{٢}$

∴ ت(س - ١) = صفر × $\frac{١}{٤}$ + ١ × $\frac{١}{٤}$ + ٢ × $\frac{١}{٢}$ = ١,٢٥

وهي نفس النتيجة التي حصلنا عليها من قبل.

ب- الوسيط: Median

سبق وأن عرَفنا في مرحلة دراسية سابقة أن الوسيط هو القيمة التي يقل عنها نصف القيم ويزيد عليها النصف الآخر. وعليه فإنه يمكن تحديد قيمة الوسيط كمتغير عشوائي باستخدام دالة كثافة الاحتمال وذلك على النحو التالي:

قيمة الوسيط (و) هي التي تحقق العلاقة التالية:

$$\int_{-\infty}^{و} ف(س)\, ءس = \frac{١}{٢}$$

وفقاً لذلك نجد أن:

احتمال أن يقل المتغير عن قيمة الوسيط = احتمال أن يزيد المتغير على قيمة الوسيط = ٠,٥ .

أي أن الوسيط هو القيمة التي تتوسط جميع القيم الممكنة للمتغير العشوائي بعد ترتيبها تصاعدياً أو تنازلياً.

مثال (٢٦):

احسب قيمة الوسيط للمتغير العشوائي الموضحة دالة كثافة احتماله في المثال (٢٣).

الحـل:

لنفترض أن **و** ترمز إلى قيمة الوسيط، وحيث أن س متغير عشوائي متصل فإنه يمكن تحديد قيمة الوسيط كما يلي:

$$\int_{٠}^{و} \frac{٢}{٩}(٣س - س^٢)\, ءس = \frac{٢}{٩}\left[\frac{٣س^٢}{٢} - \frac{س^٣}{٣}\right]_{٠}^{و} = \frac{١}{٢}$$

$$∴ \frac{٢}{٩}\left[\frac{٣و^٢}{٢} - \frac{و^٣}{٣}\right] = \frac{١}{٢}$$

ومن هذه العلاقة نحصل على:

$$٤و^٣ - ١٨و^٢ + ٢٧ = صفر$$

وبحل هذه المعادلة نجد أن:

$$(٢و - ٣)(٢و^{٢} - ٦و - ٩) = صفر.$$

وهذا يعنى أن أحد جذور تلك المعادلة هو: $و = \frac{٣}{٢} = ١,٥$

أي أن: الوسيط = ١,٥

ويمكننا التحقق من أن الجذريْن الآخريْن (و = -١,١ و ٤,١) يقعان خارج المدى المعروف وهو (صفر و ٣) للمتغير العشوائي وبالتالي يتم رفضهما.

جـ- المنوال: Mode

المنوال هو القيمة الأكثر تكراراً أو شيوعاً. وعليه يمكن تعريف المنوال مستخدمين في ذلك دالة كثافة الاحتمال بأنه النقطة والتي إن وجدت تكون عندها دالة كثافة الاحتمال نهاية عظمى. والمثال التالي يوضح كيفية إيجاد المنوال باستخدام دالة كثافة الاحتمال:

مثال (٢٧)

أوجد المنوال للمتغير العشوائي المعرَّف في مثال (٢٣)، ثم علِّق على النتيجة التي تحصل عليها وذلك في ضوء نتائجك في المثاليْن (٢٣)، (٢٦).

الحــل:

حيث أن دالة كثافة الاحتمال للمتغير العشوائي هي:

$$ف(س) = \begin{cases} \frac{٢}{٩}(٣س - س^{٢}) & صفر < س < ٣ \\ صفر & خلاف ذلك \end{cases}$$

فإن المنوال يمكن تحديده بالنقطة التي تجعل هذه الدالة نهاية عظمى وذلك على النحو التالي:

$$\frac{د ف(س)}{د س} = \frac{٢}{٩}(٣ - ٢س)$$

$$= (\frac{٢}{٣} - \frac{٤}{٩}س) = صفر$$

بمساواة المشتقة التفاضلية الأولى بالصفر، نحصل على:

$$س = \frac{٢}{٣} \times \frac{٩}{٤} = ١,٥$$

$$\frac{د^{٢}ف(س)}{د س^{٢}} = -\frac{٤}{٩} < صفر$$

أي أنه عدد $س = ١,٥$ يوجد نهاية عظمة للدالة $\frac{٢}{٩}(٣س - س^٢)$
وعليه يمكن استنتاج أن: قيمة المنوال (م) = ١,٥
وبمقارنة هذه النتيجة بالنتائج المتحصل عليها في المثاليْن (٢٣) (٢٦)، والمتعلقة بالقيمة المتوقَّعة والوسيط يمكننا ملاحظة أن:

القيمة المتوقَّعة (الوسط الحسابي) = الوسيط = المنوال = ١,٥

وهذا يدل على أن دالة كثافة الاحتمال لهذا المتغير هي دالة متماثلة. وفي الواقع تمثل هذه الدالة قطعاً مكافئاً فتحته إلى أسفل ورأسه هي النقطة (١,٥ ، ١,٥)، كما يقطع المحور السيني في نقطتيْن هما (٠ ، ٠) ، (٣ ، ٠). وهو شكل متماثل حول الخط المستقيم: $س = ١,٥$.

٢- مقاييس التشتت:

سوف نتناول هنا مقياسان للتشتت هما التباين والانحراف المتوسط (متوسط الانحرافات المطلقة) وذلك على النحو الموضح فيما يلي:

أ- التبايـــــن: Variance

من المعروف أن التباين يقيس درجة التشتت، أي أنه يعتبر مقياساً لدرجة تركز القيم التي يأخذها المتغير العشوائي حول القيمة المتوسطة لهذا المتغير. ولذلك يمكن إعادة تعريف التباين على النحو التالي:

$$تبا(س) = ت[س - ت(س)]^٢$$
$$= ت(س^٢) - [ت(س)]^٢$$

هذا وفي حالة المتغير العشوائي المنفصل يعطَى التباين في الصورة التالية:

$$تبا(س) = \sum ع(س_ر)[س_ر - \bar{س}]^٢$$
$$= ت(س^٢) - [ت(س)]^٢$$
$$= \sum س_ر^٢ ع(س_ر) - [\sum س_ر ع(س_ر)]^٢$$

وأما إذا كان المتغير متصلاً فإن التباين يعطى كما يلي:

$$تبا(س) = \int [س - \bar{س}]^٢ ءس$$
$$= ت(س^٢) - [ت(س)]^٢$$
$$= \int س^٢ ف(س) ءس - [\int س ف(س) ءس]^٢$$

مثال (٢٨):

احسب التباين للمتغير العشوائي المعرَّف في المثال (٢٥).

الحـل:

حيث أن التوزيع الاحتمالي للمتغير س يأخذ الصورة التالية:

س	١	٢	٣
ع(س)	$\frac{١}{٤}$	$\frac{١}{٤}$	$\frac{١}{٢}$

فإنه يمكن حساب التباين لهذا المتغير كما يلي:

تبا(س) = ت (س^{٢}) - [ت(س)]^{٢}

ويمكن حساب ذلك بطريقة أخرى كما يلي:

س^{٢}	١	٤	٩
ع(س)	$\frac{١}{٤}$	$\frac{١}{٤}$	$\frac{١}{٢}$

∴ ت(س^{٢}) = ١ × $\frac{١}{٤}$ + ٤ × $\frac{١}{٤}$ + ٩ × $\frac{١}{٢}$ = ٥,٧٥

وحيث أنه تم حساب ت(س) في مثال (٢٥) ووجدنا أنه يساوي ٢,٢٥ ، فإنه يمكن ايجاد تبا(س) كما يلي:

تبا(س) = ت (س^{٢}) - [ت(س)]^{٢}

= ٥,٧٥ - (٢,٢٥)^{٢} = ٠,٦٨٧٥

ويمكن حساب التباين لهذا المتغير بطريقة أخرى كما يلي:

حيث أن:

تبا(س) = ت[س - ت(س)]^{٢} = ت[س - ٢,٢٥]^{٢}

= ت[س^{٢} - ٤,٥ س + ٥,٠٦٢٥]

= ت(س^{٢}) - ٤,٥ ت(س) + ٥,٠٦٢٥

= ٥,٧٥ - ٤,٥ × ٢,٢٥ + ٥,٠٦٢٥ = ٠,٦٨٧٥

وهي نفس النتيجة التي حصلنا عليها من قبل.

مثال (٢٩):

باستخدام دالة كثافة الاحتمال للمتغير العشوائي المُعطاة في مثال (٢٣)، احسب التباين لهذا المتغير.

الحــل:

يمكن حساب التباين لهذا المتغير بطريقتيْن مختلفتيْن كما يلي:

الطريقة الأولى:

حيث أن: $\text{تبا}(\text{س}) = \text{ت}(\text{س}^{٢}) - [\text{ت}(\text{س})]^{٢}$

$$\text{ت}(\text{س}^{٢}) = \int_{٠}^{٣} \text{س}^{٢} \times \frac{٢}{٩}(٣\text{س} - \text{س}^{٢})\, \text{ءس}$$

$$= \frac{٢}{٩} \int_{٠}^{٣} (٣\text{س}^{٣} - \text{س}^{٤})\, \text{ءس}$$

$$= \frac{٢}{٩} \left[\frac{٣\text{س}^{٤}}{٤} - \frac{\text{س}^{٥}}{٥} \right]_{٠}^{٣}$$

$$= \frac{٢}{٩} \left(\frac{٣ \times ٨١}{٤} - \frac{٢٤٣}{٥} \right) = ٢,٧$$

هذا وقد تم حساب $\text{ت}(\text{س})$ في مثال (٢٣) ووجد أنه يساوي ١,٥. وعليه يمكن حساب $\text{تبا}(\text{س})$ كما يلي:

$\text{تبا}(\text{س}) = \text{ت}(\text{س}^{٢}) - [\text{ت}(\text{س})]^{٢}$

$= ٢,٧ - (١,٥)^{٢} = ٠,٤٥$

الطريقة الثانية:

حيث أن: $\text{تبا}(\text{س}) = \text{ت}(\text{س}^{٢}) - [\text{ت}(\text{س})]^{٢}$ ، $\text{ت}(\text{س}) = ١,٥$

$\therefore \text{تبا}(\text{س}) = \text{ت}(\text{س} - ١,٥)^{٢}$

$$= \int_{٠}^{٣} (\text{س} - ١,٥)^{٢} \times \frac{٢}{٩}(٣\text{س} - \text{س}^{٢})\, \text{ءس}$$

$$= \frac{٢}{٩} \int_{٠}^{٣} (\text{س}^{٢} - ٣\text{س} + ٢,٢٥)(٣\text{س} - \text{س}^{٢})\, \text{ءس}$$

$$= \frac{٢}{٩} \int_{٠}^{٣} (-س^{٤} + ٦س^{٣} - ١١,٢٥ س^{٢} + ٦,٧٥ س) ءس$$

$$= \frac{٢}{٩} \left[\frac{-س^{٥}}{٥} + \frac{٦س^{٤}}{٤} - \frac{١١,٢٥س^{٣}}{٣} + \frac{٦,٧٥س^{٢}}{٢} \right]_{٠}^{٣}$$

$$= \frac{٢}{٩} \left[\frac{-٢٤٣}{٥} + \frac{٤٨٦}{٤} - \frac{٣٠٣,٧٥}{٣} + \frac{٦٠,٧٥}{٢} \right] = ٠,٤٥$$

وهي نفس النتيجة السابقة.

خصائص التباين:

تجدر الإشارة إلى أن التباين يتميز بالخصائص التالية:

١- تباين المقدار الثابت = صفر

أي أن: تبا(أ) = صفر حيث أ مقدار ثابت

وهذا أمر منطقي إذا أن المقدار الثابت ليس متغيراً عشوائياً، أي لا تتغير قيمته من تجربة لأخرى وإنما يأخذ قيمة ثابتة دائماً.

٢- تبا(مقدار ثابت × متغير عشوائي) = مربع المقدار الثابت × تبايـن المتغير العشوائي

أي أن: تبا(أ س) = $أ^{٢}$تبا(س) حيث أ مقدار ثابت

٣- إذا كانت س، ص متغيريْن عشوائييْن مستقليْن فإن:

تبا(س ± ص) = تبا(س) + تبا(ص)

مثال (٣٠):

احسب تبا(٢س) للمتغير العشوائي المعطى في مثال (٢٣).

الحــل:

حيــث أن: تبا(أ س) = $أ^{٢}$تبا(س)

∴ تبا(٢س) = ٤تبا(س)

= ٤ × ٠,٤٥ = ١,٨

ب- الانحراف المتوسط: Mean Deviation

الانحراف المتوسط هو من أهم مقاييس التشتت التي تم دراستها في مرحلة دراسية سابقة ويمكن إعادة تعريف هذا المقياس على النحو التالي:

الانحراف المتوسط = $\sum$ |سـر - $\bar{س}$| ع(سـر)

في حالة المتغير المنفصل

= $\int$ |سـر - $\bar{س}$| ف(س) ءس

في حالة المتغير المتصل

مثال (٣١):

احسب الانحراف المتوسط للمتغير العشوائي المُعطَى توزيعه الاحتمالي في المثال (٢٥).

الحـل:

حيث أن: الانحراف المتوسط = $\sum$ |سـر - $\bar{س}$| ع(سـر)

$\bar{س}$ = ٢,٢٥ تم حسابه في مثال (٢٥)، فإنه يمكن حساب الانحراف المتوسط لهذا المتغير على النحو التالي:

سـر	١	٢	٣
|سـر - $\bar{س}$|	١,٢٥	٠,٢٥	٠,٧٥
ع(س)	$\frac{١}{٤}$	$\frac{١}{٤}$	$\frac{١}{٢}$

∴ الانحراف المتوسط =١,٢٥ × $\frac{١}{٤}$ +٠,٢٥ × $\frac{١}{٤}$ +٠,٧٥ × $\frac{١}{٢}$ =٠,٧٥

مثال (٣٢):

أوجد الانحراف المتوسط للمتغير العشوائي المعرَّف في المثال (٢٣).

الحــل:

حيث أن دالة كثافة الاحتمال لهذا المتغير تأخذ الصورة التالية:

$$ف(س) = \begin{cases} \frac{٢}{٩}(٣س - س^{٢}) & صفر \leq س \leq ٣ \\ صفر & خلاف ذلك \end{cases}$$

وحيث أن: $\bar{س}$ = ١,٥ تم حسابه في المثال (٢٣)، فإنه يمكن حساب الانحراف المتوسط لهذا المتغير وذلك على النحو الموضح فيما يلي:

$$\text{الانحراف المتوسط} = \int_{٠}^{٣} |\text{س} - ١,٥| \frac{٢}{٩} (٣\text{س} - \text{س}^{٢})\, \text{ءس}$$

$$= \frac{٢}{٩} \int_{٠}^{١,٥} (١,٥ - \text{س})(٣\text{س} - \text{س}^{٢})\, \text{ءس}$$

$$+ \frac{٢}{٩} \int_{١,٥}^{٣} (\text{س} - ١,٥)(٣\text{س} - \text{س}^{٢})\, \text{ءس}$$

$$= \frac{٢}{٩} \int_{٠}^{١,٥} (\text{س}^{٣} - ٤,٥\text{س}^{٢} - ٤,٥\text{س})\, \text{ءس}$$

$$+ \frac{٢}{٩} \int_{١,٥}^{٣} (-\text{س}^{٣} + ٤,٥\text{س}^{٢} - ٤,٥\text{س})\, \text{ءس}$$

$$= \frac{٢}{٩} \left[\frac{\text{س}^{٤}}{٤} + \frac{٤,٥\text{س}^{٣}}{٣} - \frac{٤,٥\text{س}^{٢}}{٢} \right]_{٠}^{١,٥}$$

$$+ \frac{٢}{٩} \left[\frac{-\text{س}^{٤}}{٤} - \frac{٤,٥\text{س}^{٣}}{٣} + \frac{٤,٥\text{س}^{٢}}{٢} \right]_{١,٥}^{٣}$$

$$= \frac{٢}{٩} [١,٢٧] + \frac{٢}{٩} [\text{صفر} - (-١,٢٧)]$$

$$= \frac{٢}{٩} \times ٢,٥٤ = ٠,٥٦$$

وبعد أن انتهينا من تناول مفاهيم المتغير العشوائي ودالة كثافة الاحتمال وبعض المقاييس الإحصائية المهمة، إلى غير ذلك من المفاهيم، فإنه من المناسب الآن التطرق إلى تلك المفاهيم والمقاييس وذلك فيما يتعلق ببعض التوزيعات الاحتمالية الهامة.

بعض التوزيعات الاحتمالية الهامة

تناولنا بالشرح في الجزء الثاني من هذا الفصل كيفية إيجاد التوزيع الاحتمالي للمتغير وذلك عن طريق حصر جميع الأحداث الممكنة الوقوع لهذا المتغير ثم حساب احتمال تحقق كل حدث من هذه الأحداث وذلك بقسمة عدد الحالات المواتية لكل حدث على حدة على عدد الأحداث الكلية الممكنة. ومع ذلك فإنه في كثير من الحالات يصعب علينا - إن لم يكن مستحيلاً - إيجاد التوزيع الاحتمالي بهذه الطريقة. وعلى سبيل المثال إذا أردنا تكوين التوزيع الاحتمالي للمتغير العشوائي الذي يمثل عدد مرات ظهور الشعار عند رمي ١٠ قطع نقود مرة

واحدة (أو رمي قطعة نقود ١٠ مرات) لكان لزاماً علينا حصر جميع الحالات الممكنة لهذه التجربة والتي تساوي في هـــذه الحالـــة (٢)^١٠ = ١٠٢٤ حالة. وكذلك حصر عدد المرات التي يظهر فيها الشعار في كل محاولة. ومن هنا يتضح لنا مدى صعوبة تكوين التوزيع الاحتمالي في مثل تلك الحالات خاصة إذا كان عدد قطع النقود المستخدمة في التجربة كبيراً.
لذلك كان من الضروري الاستعانة بالتوزيعات الاحتمالية النظرية والتي تتناولها نظرية الاحتمالات كفرع من فروع الرياضيات وكجانب مهم من جوانب النظرية الإحصائية وتطبيقاتها في المجالات المختلفة. إذ أن استخدام مبادئ الاحتمالات بمفاهيمها البسيطة وطرقها التقليدية في كثير من الأمور المعقدة لا يعدو عن كونه ضرباً من ضروب المستحيل.

هذا ويوجد العديد من التوزيعات الاحتمالية المعروفة والتي تُستخدم في عمليات التحليل الإحصائي لكثير من الظواهر. وسوف نتناول هنا وباختصار بعض التوزيعات الاحتمالية الهامة والشائعة الاستخدام وذلك على النحو الموضح فيما يلي:

أولاً: التوزيعات الاحتمالية المنفصلة Discrete Distributions

في دراستنا لبعض التوزيعات الاحتمالية المنفصلة سوف يقتصر اهتمامنا على التوزيع ثنائي الحديْن وتوزيع بواسون والتوزيع فوق الهندسي وذلك كما يلي:

١- التوزيع ثنائي الحديْن: Binomial Distribution

هو من أهم التوزيعات الاحتمالية المنفصلة. ويُستخدم هذا التوزيع لإيجاد احتمال وقوع حدث معين عدداً من المرات مقداره (ر) من بين (ن) من المحاولات المستقلة لتجربة ما. ويُرمز لهذا الاحتمال بالرمز ح(ر). هذا وفي حالة إجراء التجربة فإن تحقق الحدث موضع الاهتمام يُطلق عليه حالة "نجاح،" وأما عدم تحققه فيُطلق عليه حالة "فشل". ولعله من المناسب أن نوضح في هذا الشأن المقصود باستقلالية المحاولات وهو أنه لو كررنا التجربة عدداً من المرات يساوي (ن) فإن احتمال النجاح في كل مرة ثابت ومستقل، أي أنه لا يتأثر - في قيمته - بنتيجة التجربة في المحاولة السابقة والمحاولة اللاحقة.
وزيادة في التوضيح، لنفترض أن لدينا تجربة عشوائية وأننا كررنا إجراء التجربة - تحت نفس الظروف - (ن) من المرات. فإذا كانت نتيجة هذه التجربة ممثلة في حالتيْن فقط متنافيتيْن، وهاتان الحالتان هما النجاح أو الفشل (مرة أخرى ليس المقصود بالنجاح والفشل المعنى الحرفي لكل من الكلمتيْن، وإنما يُقصد بالنجاح تحقق حدث أو نتيجة معينة بينما يقصد بالفشل عدم تحقق هذا الحدث أو هذه النتيجة). فلو كنا بصدد حساب احتمال تحقق الحدث عدداً من المرات قدره (ر) فإننا نلاحظ أن عدد مرات تحقـــق (أو ظهور) الحدث يعتبر متغيراً عشوائياً يأخذ القيم: صفر، ١، ٢،...، ن أي أن: ر = صفر، ١، ٢،...، ن. حيث أن ر = صفر يعنى عدم تحقق الحدث بالمرة عند إجراء التجربة ن من المرات،

وأما ر = ١ فإنها تفيد تحقق الحدث مرة واحدة من بين ن من المرات، وهكذا حتى نصل إلى الحالة ر = ن والتي يتحقق فيها الحدث في كل مرة نجري فيها التجربة.

فإذا اعتبرنا أن احتمال تحقق الحدث (أي احتمال النجاح) هو مقدار ثابت يساوي ل، وأن احتمال عدم تحققه هو أيضاً مقدار ثابت ويساوي ك.

وحيث أن النجاح والفشل هما حدثان مكمِّلان ومتنافيان فإننا نحصل على:

احتمال النجاح + احتمال الفشل = ١

أي أن: ل + ك = ١

وبناءً على ما تقدم فإنه يمكن إيجاد احتمال تحقق الحدث ر من المرات عند إجراء التجربة ن من المحاولات المستقلة وذلك باستخدام الصورة التالية:

$$ع(س = ر) = ع(ر)$$

$$ع(ر) = ق^{ن}_{ر}\, ل^{ر}\, ك^{ن-ر} \qquad (١ - ١٠)$$

حيث: ر = صفر، ١، ٢، ....، ن

هذا والصورة (١ – ١٠) هي ما يعرف بدالة كثافة الاحتمال لمتغير عشوائي يتبع التوزيع ثنائي الحديْن. وقد سُمِّي هذا التوزيع بهذا الاسم وذلك لأن الصورة (١ – ١٠) ما هي إلا المفكوك العام لذات الحديْن $(ل + ك)^{ن}$. ومن خصائص دالة كثافة الاحتمال ما يلي:

- صفر ≤ ع(ر) ≤ ١ لجميع قيم ر، وهذا أمر واضح
- $\sum_{ر=٠}^{ن} ع(ر) = ١$

ويمكن إثبات الخاصية الثانية على النحو التالي:

$$\sum_{ر=٠}^{ن} ع(ر) = \sum_{ر=٠}^{ن} ق^{ن}_{ر}\, ل^{ر}\, ك^{ن-ر}$$

$$= ق^{ن}_{٠}\, ك^{ن} + ق^{ن}_{١}\, ل\, ك^{ن-١} + ... + ق^{ن}_{ن}\, ل^{ن}$$

$$= (ل + ك)^{ن} = \{ل + (١ - ل)\}^{ن} = ١$$

وفيما يلي بعض الأمثلة التي توضح كيفية استخدام التوزيع ثنائي الحديْن:

مثال (٣٣):

بالرجوع إلى مثال (١٧) والمتعلق برمي قطعة عملة ثلاث مرات، احسب الاحتمالات التالية:

أ- الحصول على الوجه الذي عليه الشعار مرتان.

ب- الحصول على الوجه الذي عليه الشعار مرة واحدة فقط.

الحـــل:

الأحداث هنا مستقلة، وحيث أن احتمال ظهور الوجه الذي عليه الشعار يأخذ قيمة ثابتة تساوي $\frac{١}{٢}$ في كل مرة فإنه يمكن استخدام التوزيع ثنائي الحديْن لحساب الاحتمالات المطلوبة وذلك كما يلي:

أ- ظهور الوجه الذي عليه الشعار يمثل حالة النجاح في هذه الحالة بينما حالة الفشل يمثلها ظهور الوجه الذي عليه الكتابة. وعليه يمكن حساب الاحتمال المطلوب على النحو التالي:

حيث أن: $ن = ٣،\ ر = ٢ ،\ ل = ك = \frac{١}{٢}$

∴ احتمال الحصول على الوجه الذي عليه الشعار مرتان

$$= ع(ر = ٢) = {}^{٣}ق_{٢}\left(\frac{١}{٢}\right)\left(\frac{١}{٢}\right)^{٢} = \frac{٣}{٨}$$

ب- في هذه الحالة نجد أن:

$$ن = ٣ ،\ ر = ١ ،\ ل = ك = \frac{١}{٢}$$

∴ احتمال الحصول على الشعار مرة واحدة

$$= ع(ر = ١) = {}^{٣}ق_{١}\left(\frac{١}{٢}\right)\left(\frac{١}{٢}\right)^{٢} = \frac{٣}{٨}$$

ويلاحَظ هنا تطابق النتائج مع ما حصلنا عليه من نتائج في المثال (١٧) والذي استُخدمت فيه المبادئ الأولية للاحتمالات ممثلةً في حصر جميع الحالات الممكنة وكذلك الحالات المواتية.

مثال (٣٤):

ألقيت زهرة نرد أربع مرات، احسب الاحتمالات التالية:

أ- الحصول على الرقم ٥ مرة واحدة.

ب- الحصول على الرقم ٣ مرتان.

جـ- الحصول على الرقم ٦ مرة واحدة على الأقل.

الحـل:

أ- في هذا الحالة نجد أن حالة النجاح تتمثل في ظهور الرقم ٥ بينما تتمثل حالة الفشل في ظهور أي رقم آخر. وظهور الرقم ٥ في هذه التجربة هـــو متغير عشوائي يأخذ القيم: صفر، ١، ٢، ٣، ٤. كما أن احتمال ظهور الرقم ٥ هو احتمال النجاح وهو احتمال ثابت ويساوي $\frac{١}{٦}$، واحتمال عدم ظهوره هو احتمال الفشل وهو ثابت أيضاً ويساوي $\frac{٥}{٦}$.

كما أن الأحداث هنا مستقلة. لذلك فإن المتغير العشوائي هنا له توزيع ثنائي الحديْن يمكن استخدامه لحساب الاحتمال المطلوب كما يلي:

حيث أن: $\text{ع}(\text{ر}) = {}^{\text{ن}}\text{ق}_{\text{ر}}\,\text{ل}^{\text{ر}}\,\text{ك}^{\text{ن}-\text{ر}}$

$$\therefore\ \text{ع}(\text{ر} = ١) = {}^{٤}\text{ق}_{١}\left(\frac{١}{٦}\right)\left(\frac{٥}{٦}\right)^{٣}$$

$$= ٤\left(\frac{١}{٦}\right)\left(\frac{١٢٥}{٢١٦}\right) = ٠,٣٨٦$$

ب- حالة النجاح هنا يمثلها ظهور الرقم ٣ بينما حالة الفشل يمثلها عدم ظهور هذا الرقم، وبالتالي نجد أن:

$$\text{ن} = ٤ \quad ، \quad \text{ر} = ٢ \quad ، \quad \text{ل} = \frac{١}{٦} \quad ، \quad \text{ك} = \frac{٥}{٦}$$

$$\therefore\ \text{الاحتمال المطلوب} = {}^{٤}\text{ق}_{٢}\left(\frac{١}{٦}\right)^{٢}\left(\frac{٥}{٦}\right)^{٢}$$

$$= ٦\left(\frac{١}{٣٦}\right)\left(\frac{٢٥}{٣٦}\right) = ٠,١١٦$$

جـ- في هذه الحالة نجد أن ظهور الرقم ٦ هو حالة النجاح وأن عدم ظهوره هي حالة الفشل. وعليه يمكن حساب احتمال ظهور الرقم ٦ مرة واحدة على الأقل باستخدام إحدى الطريقتيْن التاليتيْن:

الطريقة الأولى:

الحصول على الرقم ٦ مرة واحدة على الأقل يعنى الحصول على هذا الرقم إما مرة واحدة أو مرتيْن أو ثلاث مرات أو أربع مرات. وحيث أن هذه الحوادث متنافية فإنه يمكن تطبيق قانون جمع الاحتمالات لحساب الاحتمال المطلوب وذلك على النحو التالي:

$$\text{ع}(\text{ر} \geq ١) = \text{ع}(\text{ر} = ١) + \text{ع}(\text{ر} = ٢) + \text{ع}(\text{ر} = ٣) + \text{ع}(\text{ر} = ٤)$$

ويمكن استخدام التوزيع ثنائي الحديْن لحساب كل من هذه الاحتمالات كما يلي:

$$\text{ع}(\text{ر} = ١) = {}^{٤}\text{ق}_{١}\left(\frac{١}{٦}\right)\left(\frac{٥}{٦}\right)^{٣} = ٤\left(\frac{١}{٦}\right)\left(\frac{١٢٥}{٢١٦}\right) = ٠,٣٨٦$$

$$\text{ع}(\text{ر} = ٢) = {}^{٤}\text{ق}_{٢}\left(\frac{١}{٦}\right)^{٢}\left(\frac{٥}{٦}\right)^{٢} = ٦\left(\frac{١}{٣٦}\right)\left(\frac{١٥}{٣٦}\right) = ٠,١١٦$$

$$\text{ع}(\text{ر} = ٣) = {}^{٤}\text{ق}_{٣}\left(\frac{١}{٦}\right)^{٣}\left(\frac{٥}{٦}\right) = ٤\left(\frac{١}{٢١٦}\right)\left(\frac{٥}{٦}\right) = ٠,٠١٥$$

$$\text{ع}(\text{ر} = ٤) = {}^{٤}\text{ق}_{٤}\left(\frac{١}{٦}\right)^{٤}\left(\frac{٥}{٦}\right) = \left(\frac{١}{٦}\right)^{٤} = ٠,٠٠١$$

أي أن: احتمال ظهور الرقم ٦ مرة وأحدة على الأقل يساوي:

ع(ر ≥ ١) = ٠,٣٨٦ + ٠,١١٦ + ٠,٠١٥ + ٠,٠٠١ = ٠,٥١٨

الطريقة الثانية:

حيث أن حالة النجاح هنا يمثلها ظهور الرقم ٦ بينما حالة الفشل هي عدم ظهوره. فإننا لو استعرضنا جميع الحالات الممكنة عند رمي زهرة نرد أربع مرات وذلك فيما يتعلق بظهور الرقم ٦ أو عدم ظهوره لوجدناها كما يلي:

- عدم ظهور الرقم ٦ على الإطلاق
- ظهور الرقم ٦ مرة واحدة
- ظهور الرقم ٦ مرتين
- ظهور الرقم ٦ ثلاث مرات
- ظهور الرقم ٦ أربع مرات

وحيث أن مجموع احتمالات ظهور الحالات الخمسة المُشار إليها يجب أن يساوي الواحد الصحيح، فإنه يمكن إيجاد احتمال ظهور الرقم ٦ مرة واحدة على الأقل بحساب الاحتمال المكمِّل للحالة التي لا يظهر فيها الرقم ٦ على الإطلاق. أي أن:

ع(ر ≥ ١) = ١ - ع(ر = صفر)

$$= ١ - ق^{٤}_{.} \left(\frac{١}{٦}\right)^{صفر} \left(\frac{٥}{٦}\right)^{٤} = ١ - \frac{٦٢٥}{١٢٩٦} = ٠,٥١٨$$

وهي نفس النتيجة التي حصلنا عليها من قبل. ويلاحَظ هنا أن الطريقة الثانية هي الأيسر وبشكل ملحوظ من حيث إجراء العمليات الحسابية.

مثال (٣٥):

إذا كانت نسبة المعيب في إنتاج أحد المصانع هي ١٠٪. فإذا أخذنا عينة مكونة من خمس وحدات من إنتاج هذا المصنع، احسب الاحتمالات التالية:

أ- احتمال وجود وحدتيْن معيبتيْن.

ب- احتمال ألا يكون بالعينة أية وحدات معيبة.

جـ- احتمال وجود وحدة واحدة سليمة.

د- احتمال وجود ثلاث وحدات سليمة على الأكثر.

الحـل:

أ- حالة النجاح في هذه الحالة تتمثل في وجود وحدة معيبة بالعينة المختارة بينما تتمثل حالة الفشل في الحصول على وحدة سليمة بين مفردات العينة. وعليه يمكن حساب الاحتمال المطلوب كما يلي:

حيث أن: ن = ٥ ، ر = ٢ ، ل = ٠,١ ، ك = ٠,٩

∴ احتمال وجود وحدتيْن معيبتيْن = $ع(ر = ٢)$

$$= {}^{٥}ق_{٢}(٠,١)^{٢}(٠,٩)^{٣}$$

$$= ١٠(٠,٠١)(٠,٧٢٩) = ٠,٠٧٣$$

ب- هنا أيضا ظهور وحدة معيبة هو حالة النجاح وظهور وحدة سليمة هو حالة الفشل.

وحيث أن:

$ن = ٥$ ، $ر =$ صفر ، $ل = ٠,١$ ، $ك = ٠,٩$

∴ احتمال عدم وجود وحدة معيبة = $ع(ر =$ صفر$)$

$$= {}^{٥}ق_{٠}(٠,١)^{صفر}(٠,٩)^{٥} = ٠,٥٩$$

جـ- في هذه الحالة ظهور وحدة سليمة هي حالة النجاح بينما ظهور وحدة معيبة هي حالة الفشل. وعليه نجد أن:

$ن = ٥$ ، $ر = ١$ ، $ل = ٠,٩$ ، $ك = ٠,١$

∴ احتمال وجود وحدة واحدة سليمة = $ع(ر = ١)$

$$= {}^{٥}ق_{١}(٠,٩)(٠,١)^{٤} = ٠,٠٠٠٤٥$$

د- حالة النجاح هنا هي ظهور وحدة سليمة بينما تتمثل حالة الفشل في ظهور وحدة معيبة.

وحيث أن:

$ن = ٥$ ، $ر = ٣، ٢، ١،$ صفر ، $ل = ٠,٩$ ، $ك = ٠,١$

فإننا نحصل على:

احتمال وجود ثلاث وحدات سليمة على الأكثر = $ع(ر \geq ٣)$

$$= ع(ر = ٣) + ع(ر = ٢) + ع(ر = ١) + ع(ر = ٠)$$

$$= {}^{٥}ق_{٣}(٠,٩)^{٣}(٠,١)^{٢} + {}^{٥}ق_{٢}(٠,٩)^{٢}(٠,١)^{٣}$$

$$+ {}^{٥}ق_{١}(٠,٩)(٠,١)^{٤} + {}^{٥}ق_{٠}(٠,٩)^{٠}(٠,١)^{٥}$$

$$= ٠,٠٧٢٩ + ٠,٠٠٨١٤ + ٠,٠٠٠٤٥ + ٠,٠٠٠٠١ = ٠,٠٨١٥٦$$

هذا ويمكن حساب قيمة هذا الاحتمال باستخدام الاحتمالات المكملة وذلك على النحو التالي:

$$ع(ر \leq ٣) = ١ - ع(ر > ٣)$$

$$= ١ - [\, ع(ر = ٤) + ع(ر = ٥) \,]$$

$$= ١ - {}^{٥}ق_{٤}(٠,٩)^{٤}(٠,١) + {}^{٥}ق_{٥}(٠,٩)^{٥}(٠,١)^{٠}$$

$$= ١ - ٠,٣٢٨٠٥ - ٠,٥٩٠٤٩ = ٠,٠٨١٥$$

وهي نفس النتيجة التي حصلنا عليها من قبل.

مثال (٣٦):

في مصنع ما لصناعة الملابس الجاهزة أوضحت سجلات نوع معين من الآلات داخل المصنع أن عدد الأعطال خلال أيام العمل السنوية هو ٧٢ يوماً فإذا كان لدى المصنع ٦ آلات من هذا النوع، وأريد صناعة خمسين فستانــاً في يوم واحد. الأمر الذي يستلزم وجود أربع آلات على الأقل تعمل في هذا اليوم بدون أعطال. ما هو احتمال الانتهاء من صناعة الخمسين فستاناً في يوم واحد إذا كان عدد أيام العمل في السنة هو ٣٦٠ يوماً؟.

الحـــل:

احتمال تعطل الآلة = $\frac{٧٢}{٣٦٠}$ = ٠,٢

∴ احتمال عدم تعطل الآلة = ١ - ٠,٢ = ٠,٨

احتمال صناعة الخمسين فستاناً في يوم واحد = احتمال عدم تعطل أربع آلات على الأقل. أي أن المطلوب هو حساب احتمال عدم تعطل أربع آلات على الأقل. وهنا يمكننا اعتبار أن عدم تعطل الآلة هو حالة النجاح وأن تعطلها هو حالة الفشل.

وحيث أن:

ن = ٦ ، ر = ٤ ، ٥ ، ٦ ، ل = ٠,٨ ، ك = ٠,٢

∴ الاحتمال المطلوب = ح(ر ≥ ٤)

= ح(ر = ٤) + ح(ر = ٥) + ح(ر = ٦)

$= {}^{٦}ق_{٤}\,(٠,٨)^{٤}(٠,٢)^{٢} + {}^{٦}ق_{٥}\,(٠,٨)^{٥}(٠,٢) + {}^{٦}ق_{٦}\,(٠,٨)^{٦}(٠,٢)^{٠}$

= ٠,٢٤٦ + ٠,٣٩٣ + ٠,٢٦٢ = ٠,٩٠١

مثال (٣٧):

صندوق به أربع كرات بيضاء و٦ كرات سوداء. فإذا تم سحب كرة من الصندوق خمس مرات مع إعادة الكرة المسحوبة في كل مرة قبل سحب الكرة التالية احسب الاحتمالات الآتية:

أ- الحصول على ثلاث كرات بيضاء من بين الخمس كرات المسحوبة.

ب- الحصول على أربع كرات سوداء.

الحـــل:

حيث أن الكرة المسحوبة تعاد في كل مرة قبل إجراء عملية السحب التالية، فإن احتمال الحصول على كرة ذات لون معين هو مقدار ثابت في كل مرة وهو يساوي ٠,٤ في حالة الكرة البيضاء و ٠,٦ في حالة الكرة السوداء. وحيث أن الأحداث مستقلة فإنه يمكن استخدام التوزيع ثنائي الحديْن في إيجاد الاحتمالات المطلوبة وذلك على النحو الموضح فيما يلي:

أ- حالة النجاح في هذه الحالة يمثلها ظهور الكرة البيضاء، بينما يمثل ظهور الكرة السوداء حالة الفشل. وعليه نجد أن:

$$\text{ن} = ٥ \quad ، \quad \text{ر} = ٣ \quad ، \quad \text{ل} = ٠,٤ \quad ، \quad \text{ك} = ٠,٦$$

أي أن:

احتمال ظهور ٣ كرات بيضاء = $\text{ع}(\text{ر} = ٣)$

$$= {}^{٥}\text{ق}_{٣}\,(٠,٤)^{٣}(٠,٦)^{٢} = ٠,٢٣$$

ب- ظهور الكرة السوداء هنا يمثل حالة النجاح بينما يمثل ظهور الكرة البيضاء حالة الفشل. وعلى ذلك نجد أن:

$$\text{ن} = ٥ \quad ، \quad \text{ر} = ٤ \quad ، \quad \text{ل} = ٠,٦ \quad ، \quad \text{ك} = ٠,٤$$

أي أن:

احتمال ظهور أربع كرات سوداء = $\text{ع}(\text{ر} = ٤)$

$$= {}^{٥}\text{ق}_{٤}\,(٠,٦)^{٤}(٠,٤) = ٠,٢٥٩$$

بعض خصائص التوزيع ثنائي الحديْن:

يمكننا التعرف على خصائص التوزيع ثنائي الحديْن وذلك من خلال قياس التوقُّع والتباين والالتواء والتفلطح لهذا التوزيع باستخدام العزوم. وتيسيراً على القارئ سوف نكتفي هنا بعرض الوسط الحسابي (التوقُّع) والتباين للتوزيع ثنائي الحديْن وذلك دون إثبات:

حيث أن: $\text{ت}(\text{س}) = \sum \text{س}_{\text{ر}}\,\text{ع}(\text{س}_{\text{ر}})$

وعليه فإنه يمكن حساب القيمة المتوقَّعة لمتغير يتبع التوزيع ثنائي الحديْن حيث نجد أن:

$$\text{ت}(\text{س}) = \sum_{\text{ر}=٠}^{\text{ن}} \text{ر}\;{}^{\text{ن}}\text{ق}_{\text{ر}}\,\text{ل}^{\text{ر}}\,\text{ك}^{\text{ن}-\text{ر}} = \text{ن ل}$$

أي أن التوقع للتوزيع ثنائي الحديْن (الوسط الحسابي له) يساوي حاصل ضرب عدد مرات إجراء التجربة (ن) في احتمال النجاح (ل).

كما يمكن إثبات أن:

$$\text{ت}(\text{س}^{٢}) = \sum_{\text{ر}=٠}^{\text{ن}} \text{ر}^{٢}\,\text{ع}(\text{س} = \text{ر})$$

$$= \sum_{\text{ر}=٠}^{\text{ن}} \text{ر}^{٢}\;{}^{\text{ن}}\text{ق}_{\text{ر}}\,\text{ل}^{\text{ر}}\,\text{ك}^{\text{ن}-\text{ر}}$$

$$= \text{ن}(\text{ن} - ١)\text{ل}^{٢} + \text{ن ل}$$

وحيث أن:

$$\text{تبا}(\text{س}) = \text{ت}(\text{س}^{٢}) = [\text{ت}(\text{س})]^{٢}$$

فإن تباين التوزيع ثنائي الحديْن هو:

$$\text{تبا}(\text{س}) = \text{ن}(\text{ن} - ١)\text{ل}^{٢} + \text{ن ل} - (\text{ن ل})^{٢} = \text{ن ل ك}$$

أي أن الانحراف المعياري للتوزيع ثنائي الحديْن يحدد بالعلاقة:

$$\text{ع} = \sqrt{\text{ن ل ك}}$$

معامل الاختلاف للتوزيع ثنائي الحديْن:

سبق أن بيننا كلاً من التوقع (الوسط الحسابي) والانحراف المعياري للتوزيع ثنائي الحديْن، وحيث أن معامل الاختلاف يعطى بالعلاقة:

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times ١٠٠$$

فإنه يمكن ايجاد معامل الاختلاف للتوزيع ثنائي الحديْن وذلك كما يلي:

$$\text{معامل الاختلاف للتوزيع ثنائي الحديْن} = \frac{\text{الانحراف المعياري للتوزيع}}{\text{الوسط الحسابي للتوزيع}} \times ١٠٠$$

$$= \frac{\sqrt{\text{ن ل ك}}}{\text{ن ل}} \times ١٠٠ = \sqrt{\frac{\text{ك}}{\text{ن ل}}} \times ١٠٠$$

الالتواء والتفلطح للتوزيع ثنائي الحديْن:

بإيجاد العزوم الأربعة الأولى حول الوسط الحسابي يمكننا إثبات أن معاملي الالتواء والتفلطح للتوزيع ثنائي الحديْن يأخذان الصورتيْن التاليتيْن:

$$\text{معامل الالتواء} = \frac{\text{ك} - \text{ل}}{\sqrt{\text{ن ل ك}}} \quad ، \quad \text{معامل التفلطح} = ٣ + \frac{١ - ٦\text{ل ك}}{\text{ن ل ك}}$$

وهذا يعنى أنه كلما زادت ن واقتربت قيمتي ل، ك من بعضهما (أي اقتربت قيمة كل منهما من ٠,٥) كلما اقترب معامل الالتواء من الصفر واقترب كذلك معامل التفلطح من القيمة ٣. وفي هذه الحالة يقترب التوزيع ثنائي الحديْن من التماثل ويؤول بزيادة ن إلى التوزيع المعتاد (من المعروف أن معامل الالتواء للتوزيع المعتاد يساوي الصفر وأن معامل التفلطح له يبلغ ٣) وسوف يتم التعرض لذلك في موضع لاحق من هذا الفصل.

مثال (٣٨):

إذا كانت نسبة المعيب في إنتاج إحدى الآلات هي ٦٪. فإذا تم شراء - وبطريقة عشوائية - مائة وحدة من إنتاج هذه الآلة، أوجد ما يلي:

أ - الوسط الحسابي والانحراف المعياري لتوزيع الوحدات المعيبة في الوحدات المشتراة.

ب- معاملي الالتواء والتفلطح لتوزيع الوحدات المعيبة، ثم علق على ما تحصل عليه من نتائج.

الحـل:

أ- حيث أن الوحدات المعيبة تتبع التوزيع ثنائي الحديْن، وحيث أن:

$\text{ل} = ٠,٠٦ \quad ، \quad \text{ك} = ٠,٩٤ \quad ، \quad \text{ن} = ١٠٠$

فإنه يمكن حساب الوسط الحسابي والانحراف المعياري لتوزيع الوحدات المعيبة وذلك على النحو التالي:

$$\text{ت}(\text{س}) = \bar{\text{س}} = \text{ن ل}$$

$$= ١٠٠\ (٠,٠٦) = ٦ \text{ وحدات معيبة}$$

$$\text{الانحراف المعياري} = \sqrt{\text{ن ل ك}} = \sqrt{١٠٠(٠,٠٦)٠,٩٤} = \sqrt{٥,٦٤}$$

$$= ٢,٣٧ \text{ وحدة معيبة}$$

وهذا يعنى أننا لو قمنا بشراء ١٠٠ وحدة من انتاج هذه الآلة ولأكثر من مرة، بمعنى انه في كل مرة نشترى ١٠٠ وحدة، فإن عدد الوحدات المعيبة سوف يبلغ في المتوسط ٦ وحدات وبانحراف معياري مقداره ٢,٣٧ وحدة.

ب- معامل الالتواء $= \dfrac{\text{ك} - \text{ل}}{\sqrt{\text{ن ل ك}}}$

$$= \frac{٠,٩٤ - ٠,٠٦}{\sqrt{١٠٠(٠,٠٦)٠,٩٤}} = ٠,٣٧$$

معامل التفلطح $= ٣ + \dfrac{١ - ٦\text{ل ك}}{\text{ن ل ك}}$

$$= ٣ + \frac{١ - ٦(٠,٠٦)٠,٩٤}{١٠٠(٠,٠٦)٠,٩٤} = ٣ + ٠,١٢ = ٣,١٢$$

أي أن توزيع الوحدات المعيبة يقترب كثيراً من التوزيع المعتاد.

٢- توزيع بواسون: Poisson Distribution

توزيع بواسون من التوزيعات المنفصلة التي لها تطبيقات هامة في الحياة العملية، شأنه في ذلك شأن التوزيع ثنائي الحديْن. وينسب هذا التوزيع إلى عالم الرياضيات الفرنسي سيمون دنيس بواسون Simon Denis Poisson والذي قدمه عام ١٨٣٧م.

هذا ويمكن استخدام توزيع بواسون بدلاً من التوزيع ثنائي الحديْن وذلك في الحالات التي يكون فيها عدد المحاولات كبيراً (ن كبيرة) بينما تكون الحوادث نادرة الوقوع (ل صغيرة جداً). كما يرتبط حدوث الحوادث بوحدة قياس كالزمن أو المسافة أو الحجم... الخ. وبذلك

يتضح الفارق ما بيـــن التوزيع ثنائي الحديْن وتوزيع بواسون وهو أنه في الوقت الذي يستخدم فيه التوزيع ثنائي الحديْن في إيجاد احتمال تحقق حادث ما عدداً معيناً من المرات وذلك من بين عدد ن من المحاولات نجد أن توزيع بواسون يستخدم في إيجاد احتمال تحقق هذا الحدث عدداً معيناً من المرات خلال وحدة قياس معينة (وحدات الزمن على سبيل المثال). وفيما عدا ذلك يتفق التوزيعان فيما بينهما من حيث الشروط الواجبة لتطبيق كل منهما وهي:

- أن يكون هناك حادثان فقط متنافيان.
- أن تكون الأحداث مستقلة.
- أن يبقى متوسط عدد مرات تحقق الحدث ثابتاً لوحدة من الزمن.

هذا ويستخدم توزيع بواسون في حالات كثيرة من حياتنا العملية نذكر منها على سبيل المثال ما يلي:

⇦ عدد المكالمات التليفونية التي ترد إلى قسم البوليس في وحدة الزمن (دقيقة، ساعة، يوم... الخ).

⇦ عدد حوادث السيارات التي تحدث عند تقاطع ما أو في شارع ما وذلك خلال وحدة الزمن (ساعة، يوم... الخ).

⇦ عدد مرات تعطل آلة ما في وحدة الزمن.

⇦ عدد العملاء الذين يستخدمون طلمبة البنزين في محطة ما لخدمة السيارات وذلك خلال ساعة معينة.

⇦ عدد الأخطاء المطبعية في الصفحة الواحدة أو الملزمة الواحدة في كتاب ما.

⇦ عدد مخالفات المرور في الساعة عند تقاطع طرق معين.

⇦ عدد مرات حوادث القطارات والتي يخرج فيها القطار من على الشريط الحديدي المخصص له.

⇦ عدد الوفيات في سن معينة وذلك خلال وحدة زمنية محددة (شهر، سنة على سبيل المثال).

وعموما فإن تطبيق توزيع بواسون يصبح ملائماً في الحالات التي يتحقق فيما ما يلي:

- أن يكون احتمال النجاح - وبالتالي احتمال الفشل - ثابتاً في كل محاولة.
- أن يكون الاحتمال (ل) صغيراً جداً ويقترب في قيمته من الصفر، وبالتالي يكون احتمال الفشل كبيراً جداً ويقترب في قيمته من الواحد الصحيح.
- أن يكون عدد المحاولات (ن) كبيراً بحيث أن: ن ل = مقدار ثابت يرمز له بالرمز λ ، أي أن: λ = ن ل

وإذا ما تحققت هذه الشروط فإن التوزيع ثنائي الحديْن يؤول إلى توزيع بواسون والذي يأخذ الصورة التالية:

$$ع(س = ر) = ع(ر)$$

$$ع(ر) = \frac{هـ^{-\lambda}\lambda^{ر}}{ر!} \qquad (١ - ١١)$$

حيث:

ر = عدد النجاحات (عدد مرات تحقق حادث ما)، ر = ٠، ١، ٢، ....، ∞

λ = متوسط عدد النجاحات في وحدة الزمن.

هـ = أساس نظام اللوغاريتم الطبيعي، أي تساوي ٢,٧١٨٢٨.

هذا ويمكن إثبات أن الدالة (١ - ١١) هي دالة كثافة احتمال وذلك على النحو الموضح فيما يلي:

$$\sum_{ر=٠}^{\infty} ع(س = ر) = \sum_{ر=٠}^{\infty} \frac{هـ^{-\lambda}\lambda^{ر}}{ر!}$$

$$= هـ^{-\lambda} \sum_{ر=٠}^{\infty} \frac{\lambda^{ر}}{ر!}$$

وحيث أن: $\sum_{ر=٠}^{\infty} \frac{\lambda^{ر}}{ر!} = هـ^{\lambda}$

$$\therefore \sum_{ر=٠}^{\infty} ع(س = ر) = هـ^{-\lambda} هـ^{\lambda} = ١$$

وفيما يلي بعض الأمثلة التي توضح كيفية تطبيق توزيع بواسون:

مثال (٣٩):

من واقع السجلات الحديثة لأحد أقسام البوليس تبيَّن أن هذا القسم قد تلقى مكالمة تليفونية واحدة للتبليغ عن حادث تصادم سيارتيْن خلال الخمس والعشرين يوماً المنصرمة. احسب الاحتمالات التالية:

أ - ورود مكالمة واحدة للتبليغ عن حادث تصادم خلال أسبوع قادم.

ب- ورود مكالمتيْن خلال شهر قادم عدد أيامه ثلاثون يوماً.

جـ - ورود ثلاث مكالمات على الأكثر خلال شهرين (وذلك باعتبار أن عدد أيام الشهر هو ٣٠ يوماً).

د- ورود مكالمة واحدة على الأقل خلال أسبوع معين.

هـ- عدم ورود أية مكالمات على الإطلاق خلال عشرة أيام مقبلة.

الحـــل:

متوسط عدد المكالمات في اليوم الواحد = $\frac{١}{٢٥}$ = ٠,٠٤

أ - متوسط عدد المكالمات خلال أسبوع = ٠,٠٤ × ٧ = ٠,٢٨

أي أن: λ = ٠,٢٨

وحيث أن: $ر$ = ١

∴ احتمال ورود مكالمة واحدة خلال اسبوع = $\frac{هـ^{-٠,٢٨}\left(٠,٢٨\right)^{١}}{١!}$

= ٠,٢٨(٠,٧٦) = ٠,٢١٣

ب- متوسط عدد المكالمات خلال شهر = ٠,٠٤ × ٣٠ = ١,٢

أي أن: λ = ١,٢

وحيث أن: $ر$ = ٢

∴ احتمال ورود مكالمة واحدة خلال اسبوع = $\frac{هـ^{-١,٢}\left(١,٢\right)^{٢}}{٢!}$

= $\frac{١,٤٤ \times ٠,٣٠١}{٢}$ = ٠,٢١٧

جـ - متوسط عدد المكالمات خلال شهريْن = ٠,٠٤ × ٣٠ × ٢ = ٢,٤

أي أن: λ = ٢,٤

وحيث أن: $ر$ = صفر، ١، ٢، ٣

∴ احتمال ورود ثلاث مكالمات على الأكثر خلال شهريْن

$$= \sum_{ر=٠}^{٣} \frac{هـ^{-\lambda}\lambda^{ر}}{ر!} = \frac{هـ^{-٢,٤}\left(٢,٤\right)^{٠}}{٠!} + \frac{هـ^{-٢,٤}\left(٢,٤\right)^{١}}{١!}$$

$$+ \frac{هـ^{-٢,٤}\left(٢,٤\right)^{٢}}{٢!} + \frac{هـ^{-٢,٤}\left(٢,٤\right)^{٣}}{٣!}$$

$$= هـ^{-٢,٤}[١ + ٢,٤ + \frac{\left(٢,٤\right)^{٢}}{٢} + \frac{\left(٢,٤\right)^{٣}}{٣ \times ٢}]$$

= ٠,٠٩١[١ + ٢,٤ + ٢,٨٨ + ٢,٣٠٤]

= ٠,٠٩١ × ٨,٥٨٤ = ٠,٧٨١

د - متوسط عدد المكالمات خلال أسبوع = ٠,٠٤ × ٧ = ٠,٢٨

أي أن: λ = ٠,٢٨

∴ احتمال ورود مكالمة واحدة على الاقل خلال اسبوع

$$= \sum_{ر=1}^{\infty} \frac{هـ^{-0,28} (0,28)^{ر}}{ر!}$$

وحيث أن: $\sum_{ر=0}^{\infty} \frac{هـ^{-0,28} (0,28)^{ر}}{ر!} = 1$

∴ الاحتمال المطلوب = $1 - \frac{هـ^{-0,28} (0,28)^{0}}{0!}$

= ١ - ٠,٧٥٦ = ٠,٢٤٤

هـ- متوسط عدد المكالمات خلال عشرة أيام = ٠,٠٤ × ١٠= ٠,٤

أي أن: λ = ٠,٤، وحيث أن: ر = صفر

∴ احتمال عدم ورود مكالمات خلال عشرة ايام = $\frac{هـ^{-0,4} (0,4)^{0}}{0!}$ = ٠,٦٧

مثال (٤٠):

إذا كانت نسبة الوحدات المعيبة في إنتاج أحد المصانع هي ٣٪، وأن عدد الوحدات المعيبة يتبع توزيع بواسون. وبافتراض أننا سحبنا عينة عشوائية حجمها ٣٠ وحدة من إنتاج هذا المصنع، المطلوب:

أ- إيجاد دالة كثافة الاحتمال للوحدات المعيبة.

ب- احتمال الحصول على وحدة واحدة معيبة.

جـ- احتمال الحصول على وحدتيْن معيبتيْن على الأقل.

الحـل:

أ- حيث أن عدد الوحدات المعيبة يتبع توزيع بواسون فإن:

$$ع(س = ر) = \frac{هـ^{-\lambda} \lambda^{ر}}{ر!}$$

ل = احتمال وجود وحدة معيبة = ٠,٠٣ ، ن = ٣٠

∴ λ = نل = ٣٠ × ٠,٠٣ = ٠,٩

وبالتالي نجد أن دالة كثافة الاحتمال للوحدات المعيبة تأخذ الصورة التالية:

$$ع(س = ر) = \frac{هـ^{-0,9} (0,9)^{ر}}{ر!}$$

ب- حيث أن: $\lambda = ٠,٩$ ، $ر = ١$

∴ احتمال الحصول على وحدة معيبة = $\dfrac{هـ^{-٠,٩}(٠,٩)}{١!}$

$= ٠,٤٠٧ \times ٠,٩ = ٠,٣٦٦$

جـ- حيث أن: $\lambda = ٠,٩$ ، ر = ٢، ٣، ٤،........، ٣٠

∴ احتمال الحصول على وحدتيْن معيبتيْن على الأقل

$$= \sum_{ر=٢}^{٣٠} \frac{هـ^{-٠,٩}(٠,٩)^{ر}}{ر!}$$

وحيث أن:

$$= \sum_{ر=٠}^{\infty} \frac{هـ^{-٠,٩}(٠,٩)^{ر}}{ر!} = ١$$

∴ الاحتمال المطلوب = $١ - \sum_{ر=٠}^{١} \dfrac{هـ^{-٠,٩}(٠,٩)^{ر}}{ر!}$

$$= ١ - \left[\frac{هـ^{-٠,٩}(٠,٩)^{٠}}{٠!} + \frac{هـ^{-٠,٩}(٠,٩)}{١!}\right]$$

$$= ١ - [٠,٤٠٧ + ٠,٣٦٦] = ١ - ٠,٧٧٣ = ٠,٢٢٧$$

مثال (٤١):

إذا كان احتمال الإصابة بمرض معين بين سكان مدينة القاهرة يساوي ٠,٠٠٣ فإذا اختيرت عينة عشوائية حجمها ١٠٠٠ شخص من بين سكان مدينة القاهرة، احسب احتمال أن يوجد من بين أفراد العينة خمسة أشخاص مصابين بهذا المرض وذلك بافتراض أن عدد الأطفال المصابين يتبع توزيع بواسون.

الحـل:

حيث أن: $\lambda = ن ل = ١٠٠٠ \times ٠,٠٠٣ = ٣$

$$ع(س = ر) = \frac{هـ^{-\lambda}\lambda^{ر}}{ر!}$$

∴ الاحتمال المطلوب = $ع(س = ٥) = \dfrac{هـ^{-٣}(٣)^{٥}}{٥!}$

$$= \frac{٠,٠٥ \times ٢٤٣}{١٢٠} = ٠,١٠١$$

مثال (٤٢):

من سجلات إحدى محطات تزويد السيارات بالوقود وُجِد أنه تتوقف خمس سيارات في المتوسط كل عشر دقائق وذلك للتزود بالوقود. احسب الاحتمالات التالية:

أ- توقف أربع سيارات للتزود بالوقود خلال ١٥ دقيقة.

ب- توقف سيارة واحدة على الأقل للتزود بالوقود خلال خمس دقائق.

الحـــل:

متوسط عدد السيارات التي تتوقف خلال دقيقة واحدة = $\frac{٥}{١٠}$ = ٠,٥

أ- متوسط عدد السيارات خلال ربع ساعة = ٠,٥ × ١٥ = ٧,٥

أي أن: λ = ٧,٥

وحيث أن: ر = ٤

∴ الاحتمال المطلوب = $\frac{هـ^{-٧,٥}\,(٧,٥)^{٤}}{٤!}$ = ٠,٠٧٣

ب- متوسط عدد السيارات خلال خمس دقائق = ٠,٥ × ٥ = ٢,٥

أي أن: λ = ٢,٥

وحيث أن: ر = ١ ، ٢، ٣، ....

∴ الاحتمال المطلوب = $\sum_{ر=١}^{\infty} \frac{هـ^{-٢,٥}\,(٢,٥)^{ر}}{ر!}$

وحيث أن: $\sum_{ر=٠}^{\infty} \frac{هـ^{-٢,٥}\,(٢,٥)^{ر}}{ر!}$ = ١

∴ الاحتمال المطلوب = ١ - $\frac{هـ^{-٢,٥}\,(٢,٥)^{٠}}{٠!}$

= ١ - ٠,٠٨٢ = ٠,٩١٨

توزيع بواسون كتقريب للتوزيع ثنائي الحدّيْن:

يرى بعض الإحصائيين أنه يمكن استخدام توزيع بواسون كتقريب للتوزيع ثنائي الحدّيْن وذلك في الحالات التي تكون فيها ن كبيرة، ل أو ك صغيرة (أي أحداث نادرة). وكقاعدة عملية جيدة يستخدم توزيع بواسون في الحالات التي يتحقق فيها ما يلي:

ن ≤ ٣٠ و ن ل أو ن ك < ٥

إذ أنه عندما تكون ن كبيرة فإن استخدام التوزيع ثنائي الحدّيْن يستغرق وقتاً طويلاً، كما أن جداول الاحتمالات الخاصة به قد لا تكون متاحة للقيم الصغيرة للاحتمال (ل).

مثال (٤٣):

إذا كانت ل = ٠,٠١ ، ن = ١٥٠. وضِّح أن توزيع بواسون في هذه الحالة يعطى تقريبا جيداً للتوزيع ثنائي الحديْن.

الحـل:

لتوضيح ذلك سوف نوجد الاحتمالات المختلفة في الحالات التالية:

ر = صفر، ١، ٢، ٣ وذلك على سبيل المثال مستخدمين في ذلك كلاً من التوزيعيْن ثم نقارن بين القيمة المحسوبة للاحتمال الواحد في الحالتيْن.

أ- باستخدام التوزيع ثنائي الحديْن:

حيث أن:

ل = ٠,٠١ ، ك = ٠,٩٨ ، ن = ١٥٠

$$\therefore \text{ع(ر = صفر)} = {}^{\text{١٥٠}}\text{ق}_{\text{.}}(\text{٠,٠١})^{\text{صفر}}(\text{٠,٩٩})^{\text{١٥٠}} = \text{٠,٢٢١٥}$$

$$\text{ع(ر = ١)} = {}^{\text{١٥٠}}\text{ق}_{\text{١}}(\text{٠,٠١})(\text{٠,٩٩})^{\text{١٤٩}}$$

$$= \text{١٥٠} \times \text{٠,٠١} \times \text{٠,٢٢٣٧} = \text{٠,٣٣٥٦}$$

$$\text{ع(ر = ٢)} = {}^{\text{١٥٠}}\text{ق}_{\text{٢}}(\text{٠,٠١})^{\text{٢}}(\text{٠,٩٩})^{\text{١٤٨}}$$

$$= \text{١١١٧٥} \times \text{٠,٠٠٠١} \times \text{٠,٢٢٥٩} = \text{٠,٢٥٢٤}$$

$$\text{ع(ر = ٣)} = {}^{\text{١٥٠}}\text{ق}_{\text{٣}}(\text{٠,٠١})^{\text{٢}}(\text{٠,٩٩})^{\text{١٤٧}}$$

$$= \text{٥٥١٣٠٠} \times \text{٠,٠٠٠٠٠١} \times \text{٠,٢٢٨٢} = \text{٠,١٢٥٨}$$

ب- باستخدام توزيع بواسون:

حيث أن: ل = ٠,٠١ ، ن = ١٥٠

$$\therefore \lambda = \text{ن ل} = \text{١٥٠} \times \text{٠,٠١} = \text{١,٥}$$

$$\therefore \text{ع(ر = صفر)} = \frac{\text{هـ}^{-\text{١,٥}}(\text{١,٥})^{\text{٠}}}{\text{٠!}} = \text{٠,٢٢٣١}$$

$$\text{ع(ر = ١)} = \frac{\text{هـ}^{-\text{١,٥}}(\text{١,٥})}{\text{١!}} = \text{٠,٢٢٣١} \times \text{١,٥} = \text{٠,٣٣٤٧}$$

$$\text{ع(ر = ٢)} = \frac{\text{هـ}^{-\text{١,٥}}(\text{١,٥})^{\text{٢}}}{\text{٢!}} = \frac{\text{٠,٢٢٣١٠} \times \text{٢,٢٥}}{\text{٢!}} = \text{٠,٢٥١}$$

$$\text{ع(ر = ٣)} = \frac{\text{هـ}^{-\text{١,٥}}(\text{١,٥})^{\text{٣}}}{\text{٣!}} = \text{٠,١٢٥٥}$$

ويمكننا تلخيص النتائج التي حصلنا عليها باستخدام كل من التوزيعيْن على النحو الموضَّح في الجدول التالي:

الاحتمالات المختلفة باستخدام
التوزيع ثنائي الحديْن وتوزيع بواسون

ر	الاحتمال: ع(س = ر) = ع(ر)	
	التوزيع ثنائي الحديْن	توزيع بواسون
٠	٠,٢٢١٥	٠,٢٢٣١
١	٠,٣٣٥٦	٠,٣٣٤٧
٢	٠,٢٥٢٤	٠,٢٥١٠
٣	٠,١٢٥٨	٠,١٢٥٥

وبمقارنة قيمة كل احتمال في حالة التوزيع ثنائي الحديْن بقيمة نظيره في حالة توزيع بواسون يمكننا استنتاج أن هناك توافقاً بينهما في رقميْن أو ثلاثة أرقام عشرية. وهذا يدل على أن توزيع بواسون في هذه الحالة يعتبر تقريباً جيداً للتوزيع ثنائي الحديْن. ويرجع ذلك إلى أن قيمة ل صغيرة وقيمة ن كبيرة وأن **ن ل** = ١٥٠ × ٠,٠١ = ١,٥ وهو مقدار يقل عن الخمسة، مما يحقق الشرط المُشار إليه سابقاً والذي يتعلق بإمكانية استخدام توزيع بواسون كتقريب جيد للتوزيع ثنائي الحديْن.

بعض خصائص توزيع بواسون:

يمكننا التعرف على خصائص توزيع بواسون وذلك بإيجاد الوسط الحسابي والانحراف المعياري والالتواء والتفلطح لهذا التوزيع وذلك على النحو الموضح فيما يلي:

التوقُّع (الوسط الحسابي) لتوزيع بواسون:

حيث أن التوقع في حالة المتغير المنفصل يأخذ الصورة التالية:

$$\text{ت}(\text{س}) = \sum \text{س}_{\text{ر}}\ \text{ع}(\text{س} = \text{ر})$$

فإنه يمكن ايجاد توقع توزيع بواسون كما يلي:

$$\text{ت}(\text{س}) = \sum_{\text{ر}=٠}^{\infty} \text{ر}\ \text{ع}(\text{س} = \text{ر})$$

$$= \sum_{\text{ر}=٠}^{\infty} \text{ر}\ \frac{\text{هـ}^{-\lambda}\lambda^{\text{ر}}}{\text{ر}!}$$

والذي يمكن إثبات انه يساوي λ

أي أن: متوسط عدد مرات تحقق الحدث = λ = **ن ل**

التباين لتوزيع بواسون:

حيث أن: تبا(س) = ت(س^{٢}) - [ت(س)]^{٢}

فإنه يمكن إيجاد التباين لتوزيع بواسون على النحو التالي:

$$ت(س^{٢}) = \sum_{ر=٠}^{\infty} ر^{٢}\ ع(س = ر)$$

$$= \sum_{ر=٠}^{\infty} ر^{٢} \frac{هـ^{-\lambda}\lambda^{ر}}{ر!}$$

والذي يمكن إثبات انه يساوي $\lambda^{٢} + \lambda$.

أي أن: تباين(س) = $\lambda^{٢} + \lambda - \lambda^{٢} = \lambda$

وهذا يعنى أن: التوقع = التباين = λ

وهي خاصية هامة لتوزيع بواسون.

∴ الانحراف المعياري لتوزيع بواسون = $\sqrt{\lambda}$

معامل الاختلاف لتوزيع بواسون:

حيث أن: معامل الاختلاف = $\frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}}$ × ١٠٠

∴ معامل الاختلاف لتوزيع بواسون = $\frac{\sqrt{\lambda}}{\lambda}$ × ١٠٠ = $\frac{١٠٠}{\sqrt{\lambda}}$

الالتواء والتفلطح لتوزيع بواسون:

بإيجاد العزوم الأربعة حول الوسط الحسابي يمكننا إثبات أن معاملي الالتواء والتفلطح لتوزيع بواسون يأخذان الصورتيْن التاليتيْن:

معامل الالتواء = $\frac{١}{\sqrt{\lambda}}$ ، معامل التفلطح = ٣ + $\frac{١}{\lambda}$

ومن هذا يمكننا استنتاج أن توزيع بواسون هو توزيع موجب الالتواء، وأن مقدار هذا الالتواء يتناسب عكسياً مع الجذر التربيعي لقيمة المعلمة λ. كما أن توزيع بواسون أكثر تحدباً من التوزيع المعتاد، وأن مقدار هذا التحدب ينقص بزيادة λ والعكس صحيح.

مثال (٤٤):

باستخدام البيانات في مثال (٤٠)، المطلوب حساب ما يلي:

أ- العدد المتوقع للوحدات المعيبة.

ب- الانحراف المعياري ومعامل الاختلاف لعدد الوحدات المعيبة في العينة.
جـ- قياس الالتواء والتفلطح لتوزيع الوحدات المعيبة في العينة.

الحــل:

أ- حيث أن: ن = ٣٠ ، ل = ٠,٠٣

∴ العدد المتوقَّع للوحدات المعيبة = ت(س) = λ

أي أن: λ = نل = ٣٠ × ٠,٠٣ = ٠,٩ وحدة

ب- الانحراف المعياري للوحدات المعيبة = $\sqrt{\lambda}$ = ٠,٣ وحدة معيبة.

وهذا يعنى أننا لو قمنا بسحب ٣٠ وحدة عشوائية من إنتاج هذا المصنع ولأكثر من مرة فإن عدد الوحدات المعيبة سوف يبلغ في المتوسط ٠,٩ وحدة بانحراف معياري مقداره ٠,٣ وحدة.

معامل الاختلاف للوحدات المعيبة = $\frac{١٠٠}{\sqrt{\lambda}}$ = $\frac{١٠٠}{\sqrt{٠,٩}}$ = ١٠٥,٣٪

جـ- معامل الالتواء لتوزيع الوحدات المعيبة = $\frac{١}{\sqrt{\lambda}}$ = $\frac{١}{\sqrt{٠,٩}}$ = ١,٠٥٤

معامل التفلطح لتوزيع الوحدات المعيبة = ٣ + $\frac{١}{\lambda}$ = ٣ + $\frac{١}{٠,٩}$ = ٤,١١١

أي أن توزيع الوحدات المعيبة يقترب من التوزيع المعتاد.

٣- التوزيع فوق الهندسي: Hypergeometric Distribution

عند تناولنا لكل من التوزيع ثنائي الحديْن وتوزيع بواسون ذكرنا أنه عند تطبيق أي من هذيْن التوزيعيْن لابد وأن يتوافر شرط استقلالية الأحداث، بمعنى أن يكون احتمال النجاح ثابتاً من محاولة لأخرى. كما أوضحنا أنه في حالة استخدام توزيع بواسون يشترط كبر عدد المحاولات. والآن لعله من الملائم أن نستدرك أن هذا ليس هو الحال دائماً بشأن التطبيق العملي للتوزيعات الإحصائية. فإنه من الأمور المألوفة في كثير من الحالات أن يكون عدد المحاولات صغيراً، كمـا أن احتمال النجاح (تحقق حدث ما) لا يكون ثابتاً وإنما يتغير في قيمته من محاولة لأخرى. مما يعنى بدوره عدم استقلالية الأحداث بعضها عن بعض (يلاحَظ أن هذا هو حال مثال (٣٧) ولكن بعد افتراض عدم إعادة الكرة المسحوبة قبل عملية السحب التالية). وهنا فإن التوزيع فوق الهندسي هو التوزيع الملائم للتطبيق في مثل تلك الحالات والتي لا يصلح فيها استخدام أي من التوزيعيْن ثنائي الحديْن أو بواسون.

وللتعرف على دالة التوزيع فوق الهندسي، لنفترض أن حجم المجتمع محدود ونرمز له بالرمز (ن)، وأن هذا المجتمع يتكون من جزئيْن لكل منهما صفته المميزة وحجم كل منهما

$ن_١$، $ن_٢$. ولنفترض أننا أخذنا عينة عشوائية من المجتمع حجمها ں، وأنه تم اختيار مفردات هذه العينة بدون إرجاع. فإذا أُريد أن يكون من بين مفردات العينة ر مفردة تحقق صفة معينة (ولتكن حالة النجاح والممثَّلة في ظهور مفردة من الجزء الأول ذي الحجم $ن_١$ من المجتمع). في هذه الحالة فإن عدد حالات النجاح تتبع التوزيع فوق الهندسي والذي يأخذ الصورة التالية:

$$ع(س = ر) = \frac{{}^{ن_١}ق_{ر} \times {}^{ن_٢}ق_{ں - ر}}{{}^{ن}ق_{ں}} \quad ، \quad ر = ٠ ، ١ ، .... ن$$

حيث:

ن = $ن_١$ + $ن_٢$ وتمثل حجم المجتمع،

$ن_١$ = حجم الجزء الأول من المجتمع والذي يتميز بصفة معينة،

$ن_٢$ = حجم الجزء الثاني من المجتمع والذي يتميز بالصفة الأخرى،

ں = حجم العينة،

ر = عدد حالات النجاح (عدد المفردات داخل العينة والتي تتميز بإحدى الصفتيْن التي ينقسم إليهما المجتمع).

وبصفة عامة يمكننا تلخيص الشروط الواجبة لتطبيق هذا التوزيع وذلك على النحو التالي:

- أن ينقسم المجتمع إلى جزئيْن فقط لكل منهما صفته المميزة والتي لا توجد في الآخر.
- أن يكون السحب بدون إرجاع، بمعنى أن احتمال النجاح - وبالتالي احتمال الفشل - غير ثابت.
- أن يكون حجم المجتمع (ن) محدوداً، أي ليس كبيراً جداً.

وفيما يلي بعض الأمثلة التي توضح كيفية تطبيق التوزيع فوق الهندسي:

مثال (٤٥):

المطلوب هو حل مثال (٣٧) بعد افتراض عدم ارجاع الكرات المسحوبة.

الحــل:

حيث أن سحب الكرات تم بدون إعادة فإنه يمكننا استخدام التوزيع فوق الهندسي وذلك كما يلي:

أ- حيث أن:

ن = ١٠ ، $ن_١$ = ٤ ، $ن_٢$ = ٦ ، ں = ٥ ، ر = ٣

$$\therefore ع(س = ٣) = \frac{{}^{٤}ق_{٣} \times {}^{٦}ق_{٢}}{{}^{١٠}ق_{٥}}$$

$$= \frac{٤!\,٦!\,٥!\,٥!}{٣!\,١!\,٢!\,٤!\,١٠!} = ٠,٢٣٨$$

ب- $ع(س = ٤) = \frac{{}^{٦}\text{ق}_{٤} \times {}^{٤}\text{ق}_{١}}{{}^{١٠}\text{ق}_{٥}}$

$$= \frac{٤!\,٦!\,٥!\,٥!}{٣!\,١!\,٢!\,٤!\,١٠!} = ٠,٢٣٨$$

وهنا يلاحَظ اختلاف قيمة كل من هذيْن الاحتماليْن عن نظيره المحسوب في حالة استخدام التوزيع ثنائي الحديْن (راجع مثال (٣٧)).

مثال (٤٦):

مجموعة مكونة من ستة عشر طالباً منهم عشرة طلاب وست طالبات. فإذا اختيرت عينة مكونة من خمسة أفراد من بين مجموعة الطلاب وذلك بطريقة عشوائية وبدون إرجاع، احسب احتمالات أن تكون العينة المختارة مكونة من:

أ- ثلاث طلاب وطالبتان   ب- طالب واحد وأربع طالبات

جـ- خمسة طلاب

الحـل:

أ- حيث ان:

ن = ١٦ ، $\text{ن}_{١}$ = ١٠ ، $\text{ن}_{٢}$ = ٦ ، ه = ٥ ، ر = ٣

∴ الاحتمال المطلوب = $\frac{{}^{١٠}\text{ق}_{٣} \times {}^{٦}\text{ق}_{٢}}{{}^{١٦}\text{ق}_{٥}}$

$$= \frac{١٠!\,٦!\,٥!\,١١!}{٣!} = ٠,٤١٢$$

ب- حيث ان:

ن = ١٦ ، $\text{ن}_{١}$ = ١٠ ، $\text{ن}_{٢}$ = ٦ ، ه = ٥ ، ر = ١

∴ الاحتمال المطلوب = $\frac{{}^{١٠}\text{ق}_{١} \times {}^{٦}\text{ق}_{٤}}{{}^{١٦}\text{ق}_{٥}}$

$$= \frac{١٠!\,٦!\,٥!\,١١!}{١!\,٩!\,٤!\,٢!\,١٦!} = ٠,٠٣٤$$

جـ- حيث ان:

ن = ١٦ ، $\text{ن}_{١}$ = ١٠ ، $\text{ن}_{٢}$ = ٦ ، ه = ٥ ، ر = ٥

$$\therefore \text{الاحتمال المطلوب} = \frac{{}^{١٠}\text{ق}_{٥} \times {}^{٦}\text{ق}_{٠}}{{}^{١٦}\text{ق}_{٥}}$$

$$= \frac{١٠!\,٦!\,٥!\,١١!}{٥!\,٥!\,٠!\,٦!\,١٦!} = ٠,٠٥٨$$

مثال (٤٧):

من مجموعة أوراق اللعب تم سحب أربع ورقات، احسب احتمال الحصول على صورة واحدة والباقي أوراق ليست صورة علماً بأن السحب تم بدون إعادة.

الحـــل:

من المعلوم أن هناك ١٢ صورة من بين مجموعة أوراق اللعب، ٤٠ ورقة ليست صورة. وعليه نجد أن:

$$\text{ن} = ٥٢ \;،\; \text{ن}_{١} = ١٢ \;،\; \text{ن}_{٢} = ٤٠ \;،\; \text{هـ} = ٤ \;،\; \text{ر} = ١$$

$$\therefore \text{الاحتمال المطلوب} = \frac{{}^{١٢}\text{ق}_{١} \times {}^{٤٠}\text{ق}_{٣}}{{}^{٥٢}\text{ق}_{٤}}$$

$$= \frac{١٢!\,٤٠!\,٤!\,٤٨!}{١!\,١١!\,٣!\,٣٧!\,٥٢!} = ٠,٤٣٨$$

بعض خصائص التوزيع فوق الهندسي:

في تناولنا لخصائص التوزيع فوق الهندسي سوف نكتفي بعرض بعض خصائص هذا التوزيع دون إثبات وذلك تجنباً لما قد يواجه القارئ من صعوبات في هذا الشأن:

- الوسط الحسابي للتوزيع فوق الهندسي يأخذ الصورة التالية:

$$\text{ت}(\text{س}) = \bar{\text{س}} = \text{هـ}\left(\frac{\text{ن}_{١}}{\text{ن}}\right)$$

أي أنه يساوي حجم العينة مضروباً في احتمال النجاح في المجتمع.

- التباين للتوزيع فوق الهندسي يأخذ الصورة التالية:

$$\text{تبا}(\text{س}) = \text{هـ}\left(\frac{\text{ن}_{١}}{\text{ن}}\right)\left(\frac{\text{ن} - \text{ن}_{١}}{\text{ن}}\right)\left(\frac{\text{ن} - \text{هـ}}{\text{ن} - ١}\right)$$

حيث: $\frac{\text{ن}_{١}}{\text{ن}}$ هو احتمال النجاح في المجتمع.

$\frac{\text{ن} - \text{هـ}}{\text{ن} - ١}$ هو مقدار يساوي الواحد الصحيح عندما يتم السحب مع الإعادة، ويطلق على هذا المقدار معامل التصحيح للمجتمعات المحدودة.

مثال (٤٨):

احسب الوسط الحسابي والتباين مستخدما في ذلك بيانات مثال (٤٧).

الحــل:

حيث أن:

$\text{ن} = ٥٢ \quad ، \quad \text{ن}_١ = ١٢ \quad ، \quad \text{ن}_٢ = ٤٠ \quad ، \quad \text{ﻫ} = ٤ \quad ، \quad \text{ر} = ١$

فإنه يمكن حساب الوسط الحسابي والتباين على النحو التالي:

$$\bar{\text{س}} = \text{ﻫ}\left(\frac{\text{ن}_١}{\text{ن}}\right) = ٤\left(\frac{١٢}{٥١}\right) = ٠,٩٢$$

$$\text{تبا}(\text{س}) = \text{ﻫ}\left(\frac{\text{ن}_١}{\text{ن}}\right)\left(\frac{\text{ن} - \text{ن}_١}{\text{ن}}\right)\left(\frac{\text{ن} - \text{ﻫ}}{\text{ن} - ١}\right)$$

$$= ٤\left(\frac{١٢}{٥١}\right)\left(\frac{٥٢ - ١٢}{٥٢}\right)\left(\frac{٥٢ - ٤}{٥٢ - ١}\right)$$

$$= \frac{٤ \times ١٢ \times ٤٠ \times ٤٨}{٥٢ \times ٥٢ \times ٥١} = ٠,٦٧$$

وهذا يعنى أننا لو قمنا بسحب أربع ورقات عشوائياً من مجموعة أوراق اللعب ولأكثر من مرة (أي أننا نسحب أربع ورقات في كل مرة وبدون إعادة) فإن عدد الورقات التي تحمل الصورة سوف يبلغ في المتوسط ٠,٩٢ ورقة وبانحراف معياري مقداره ٠,٦٧ ورقة.

التوزيع ثنائي الحديْن كتقريب للتوزيع فوق الهندسي:

هنا لا يفوتنا أن نشير إلى أن التوزيع ثنائي الحديْن يمكن أن يستخدم كتقريب جيد للتوزيع فوق الهندسي وذلك في الحالات التي يكون فيها حجم المجتمع كبيراً بينما يكون حجم العينة صغيراً جداً بالنسبة لحجم المجتمع. ويرى بعض الإحصائيين أنه إذا كان حجم العينة أقل من ٥٪ من حجم المجتمـــع (أي أن: $\text{ﻫ} < ٠,٠٥\,\text{ن}$) فإن السحب بدون إرجاع يكون ذا تأثير صغير جداً على احتمال النجاح في كل محاولة. كما يعطى التوزيع ثنائي الحديْن - وهو الأيسر من حيث الاستخدام - تقريبا جيداً للتوزيع فوق الهندسي. والمثال التالي يوضح ذلك:

مثال (٤٩):

مجموعة مكونة من ١٠٠ طالب منهم ٨٠ طالباً حاصلين على تقدير نجاح جيد فأقل، ٢٠ طالباً حاصلين على تقدير نجاح جيد جداً فأكثر. فـإذا اختيرت عينة مكونة من ثلاثة طلاب، احسب احتمال أن تحتوي العينة المختارة على طالب واحـد تقدير نجاحه جيداً فأقل مستخدماً في ذلك التوزيع ثنائي الحديْن والتوزيع فوق الهندسي. ثم علق على ما تحصل عليه من نتائج.

الحــل:

أولاً: باستخدام التوزيع ثنائي الحدّين

حالة النجاح هنا يمثلها وجود طالب تقدير نجاحه جيد فأقل، وبالتالي يمكن حساب احتمال النجاح كما يلي:

$$\text{احتمال النجاح} = \frac{\text{عدد الطلاب الحاصلين على تقدير جيد فأقل}}{\text{إجمالي عدد الطلاب}}$$

$$= \frac{٨٠}{١٠٠} = ٠,٨$$

وحيث أن: $ع(س = ر) = {}^{ن}ق_{ر}\, ل^{ر}\, ك^{ن-ر}$،

$ن = ٣$ ، $ل = ٠,٨$ ، $ك = ٠,٢$ ، $ر = ١$

$\therefore$ الاحتمال المطلوب $= {}^{٣}ق_{١}\,(٠,٨)(٠,٢)^{٢}$

$$= ٣ \times ٠,٨ \times ٠,٠٤ = ٠,٠٩٦$$

ثانياً: باستخدام التوزيع فوق الهندسي:

حيث أن:

$$ع(س = ر) = \frac{{}^{ن_1}ق_{ر} \times {}^{ن_2}ق_{م-ر}}{{}^{ن}ق_{م}}،$$

$ن = ١٠٠$ ، $ن_1 = ٨٠$ ، $ن_2 = ٢٠$ ، $م = ٣$ ، $ر = ١$

$$\therefore \text{الاحتمال المطلوب} = \frac{{}^{٨٠}ق_{١} \times {}^{٢٠}ق_{٢}}{{}^{١٠٠}ق_{٣}}$$

$$= \frac{٨٠ \times ٢٠!٣!٩٧!}{٢!١٨!١٠٠!} = ٠,٠٩٤$$

وبذلك يمكننا استنتاج أن الافتراض بأن السحب قد تم مع الإعادة (باستخدامنا للتوزيع ثنائي الحدّين) أو بدون الإرجاع (باستخدامنا للتوزيع فوق الهندسي) ليس له تأثير بذكر على قيمة الاحتمال المطلوب مما يدل على أنه يمكننا استخدام التوزيع ثنائي الحدّين - وهو الأيسر من حيث إجراء العمليات الحسابية - كتقريب جيد للتوزيع فوق الهندسي. وفي واقع الأمـــر كان ذلك هو حال مثالنا هذا وذلك لكبر حجم $ن$ بالإضافة إلى أن حجم العينة كان صغيراً بالنسبة لحجم المجتمع (يلاحظ هنا أن: $م < ٠,٠٥\,ن$ أي أن: $٣ < ٠,٠٥ \times ١٠٠$). وهو الشرط الكافي لاستخدام التوزيع ثنائي الحدّين بدلاً من التوزيع فوق الهندسي.

ثانياً: التوزيعات الاحتمالية المتصلة
Continuous Probability Distributions

لقد عرفنا في مراحل دراسية سابقة بأن المتغير المتصل هو المتغير الذي يأخذ عدداً غير محدود من القيم وذلك داخل مدى معين أو فترة معينة. وعليه فإن السمة الأساسية للمتغير المتصل هي أن احتمال أن يأخذ هذا المتغير قيمة معينة يساوي الصفر دائماً، حيث أنه من الممكن أن يأخذ المتغير عدداً لانهائياً من القيم داخل هذا المدى أو تلك الفترة. ومن جهة أخرى فإنه يمكن حساب احتمال أن تقع قيمة المتغير المتصل داخل مدى معين أو فترة محددة مستخدمين في ذلك دالة الاحتمال لهذا المتغير. وهنا تجدر الإشارة إلى أن دالة الاحتمال للمتغير المتصل عادة ما يطلق عليها دالة كثافة الاحتمال وذلك تمييزاً لها عن دالة الاحتمال للمتغير المنفصل. وبالرغم من شيوع هذه التفرقة إلا أن البعض قد يستخدم مُسمَّى «دالة كثافة الاحتمال» سواء كان المتغير متصلاً أو منفصلاً.

وهناك العديد من التوزيعات الاحتمالية المتصلة والتي من أهمها التوزيع المعتاد وتوزيع مربع كاي وتوزيع ف وتوزيع ت. ويُعتبر التوزيع المعتاد أهم التوزيعات الاحتمالية على إطلاقها متصلة كانت أم منفصلة وذلك لاعتبارات عديدة. فالتوزيع المعتاد هو أكثر هذه التوزيعات استخداماً في التطبيقـــات الإحصائية وذلك لأن كثيراً من الظواهر (وبتعبير أدق قياسات هذه الظواهر) موزعة توزيعاً معتاداً أو قريباً جداً من التوزيع المعتاد، أو أنها على الأقل يمكن أن تؤول إلى التوزيع المعتاد تحت شروط يسهل تحقيقها. فالظواهـــــر مثل أطوال مجموعة من الأفراد أو أوزانهم أو درجات مجموعة من الطلاب أو درجة ذكائهم... الخ لها جميعا توزيعات معتادة. وبالإضافة إلى ذلك فإن كثيراً من التوزيعات الأخرى تقترب - تحت شروط معينة - من التوزيع المعتاد. وبصفة عامة فإنه كلما زاد عدد المفردات (أو القياسات) لظاهرة من الظواهر كلما اقترب توزيعها من التوزيع المعتاد.

وفيما يلي سوف نتعرض بشيء من التفصيل إلى دراسة كل من التوزيع المعتاد وتوزيع ت وهما من أكثر التوزيعات الاحتمالية المتصلة استخداماً في مجالات العينات واختبارات الفروض وعمليات الاستدلال الإحصائي.

١- التوزيع المعتاد: Normal Distribution

التوزيع المعتاد هو - كما ذكرنا - أكثر التوزيعات الاحتمالية استخداماً في عمليات التحليل الإحصائي، وعليه تُبنى النظرية الحديثة لعلم الإحصاء. وهو توزيع جرسي الشكل، أي أنه توزيع متماثل حول القيمة المتوسطة (وسطاً حسابياً كانت ام وسيطاً ام منوالاً). كما أنه لو تحركنا بعيداً عن الوسط الحسابي في كلا الاتجاهيْن فإن منحنى التوزيع المعتاد يقترب أكثر فأكثر من المحور الأفقي ولكنه لا يلامسه أبداً. ويرجع اكتشاف ذلك إلى مجموعة من العلماء منهم لابلاس Laplace وكذلك جاوس Gauss والذي ينسب إليه أحياناً هذا التوزيع فيقال توزيع جاوس.

وتأخذ دالة كثافة الاحتمال للمتغير العشوائي المعتاد الصورة التالية:

$$\text{د}(\text{س}) = \frac{1}{\sqrt{2\text{ط}}\,\text{ع}}\,\text{هـ}^{-\frac{1}{2}\left(\frac{\text{س}-\bar{\text{س}}}{\text{ع}}\right)^{2}} \qquad (1-12)$$

حيث: $-\infty \leq$ س $\leq \infty$ ،

ط النسبة التقريبية (مقدار ثابت) وتساوي ٣,١٤١٦ تقريباً،

هـ أساس اللوغاريتم الطبيعي وتساوي ٢,٧١٨٣ تقريبا،

$\bar{\text{س}}$ الوسط الحسابي للتوزيع،

ع الانحراف المعياري للتوزيع.

هذا ويُطلق على $\bar{\text{س}}$، ع معلمات التوزيع Parameters

وحيث أن الدالة (١ ــ ١٢) تمثل دالة كثافة احتمال فإن المساحة تحت منحنى هذه الدالة تساوي الواحد الصحيح. أي أن:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\text{ط}}\,\text{ع}}\,\text{هـ}^{-\frac{1}{2}\left(\frac{\text{س}-\bar{\text{س}}}{\text{ع}}\right)^{2}}\,\text{ءس} = 1$$

ومن الممكن إثبات ذلك رياضياً، وهو ما يخرج عن نطاق هذا الكتاب. كما يُلاحَظ هنا أننا استخدمنا علامة التكامل للمتغير المتصل وذلك بدلاً من علاقة المجموع $\sum$ والخاصة بالمتغيرات المنفصلة.

هذا ويلعب كل من الوسط الحسابي والانحراف المعياري دوراً هاماً في تحديد شكل المنحنى، حيث تمثل الدالة (١ ــ ١٢) عدداً كبيراً جداً من المنحنيات يتوقف شكل كل منها على قيمة كل من $\bar{\text{س}}$ و ع. فنجد أن موقع المنحنى يتحدد بقيمة الوسط الحسابي، وبتغير هذه القيمة يتغير مكان المنحنى على المحور الأفقي من مكان لآخر. وأما فيما يتعلق بالمعلمة الثانية للتوزيع وهي الانحراف المعياري فإننا نجد أنه كلما زادت قيمة ع كلما زاد اتساع المنحنى وكلما صغرت قيمة ع كان المنحنى أكثر تدبباً وارتفاعاً.

ويمكن حساب الوسط الحسابي والتباين ومعاملي الالتواء والتفلطح للتوزيع المعتاد وذلك باستخدام عمليات التكامل اللازمة لذلك. وهنا، وتجنباً لتفصيلات متعددة قد ترهق القارئ ذهنيا وتنحو به بعيداً عن الهدف من تناولنا لهذا التوزيع وذلك وفقاً لحدود فرضها محتوى هذا الكتاب، فإننا سوف نكتفي بعرض تلك المقاييس وذلك على النحو التالي:

$$\text{ت}(\text{س}) = \int_{-\infty}^{\infty} \frac{\text{س}}{\sqrt{2\text{ط}}\,\text{ع}}\,\text{هـ}^{-\frac{1}{2}\left(\frac{\text{س}-\bar{\text{س}}}{\text{ع}}\right)^{2}}\,\text{ءس} = \bar{\text{س}}$$

$$\text{تبا}(\text{س}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\text{ط}}\,\text{ع}}\,\text{هـ}^{-\frac{1}{2}\left(\frac{\text{س}-\bar{\text{س}}}{\text{ع}}\right)^{2}}\,[\text{س}-\bar{\text{س}}]\,\text{ءس} = \text{ع}^{2}$$

بالإضافة إلى ذلك، فإنه يكن حساب معاملي الالتواء والتفلطح للتوزيع المعتاد مستخدمين في ذلك العزوم حول الوسط الحسابي والتي تؤدى بنا في النهاية إلى أن:

معامل الالتواء = صفر ، معامل التفلطح = ٣

وفي هذا الشأن تجدر الإشارة إلى أن درجة الالتواء لأي توزيع تزداد حدة كلما بعدت قيمة معامل الالتواء عن الصفر. كما أن التوزيع يكون أقـــل تفلطحاً كلما زادت قيمة معامل التفلطح عن ٣ والعكس صحيح في الحالتيْن.

هذا ومن أهم خصائص التوزيع المعتاد أنه يمكن تقسيم المساحة تحت المنحنى الممثِّل له إلى قطاعات محددة وذلك بدلالة الوسط الحسابي والانحراف المعياري. فمن المعروف أنه لو أسقطنا عموداً من قمة المنحنى المعتاد على المحور الأفقي فإن هذا العمود يقسم المساحة تحت المنحنى إلى نصفيْن متطابقيْن تماماً. وأن هذا العمود سوف يقطع المحور الأفقي في نقطة تمثل الوسط الحسابي والوسيط والمنوال والتي تتساوي جميعها من حيث القيمة. بالإضافة إلى ذلك فإنه من الممكن تحديد نسبة المساحة تحت المنحنى المعتاد والمحصورة بين قيمتيْن محددتيْن من قيم الظاهرة التي يمثلها المنحنى وذلك إذا عرفنا انحراف كل من القيمتيْن عن الوسط الحسابي مقدَّراً بالانحراف المعياري. فعلى سبيل المثال لو كان انحراف القيمتيْن عن الوسط الحسابي للتوزيع يساوي +ع، –ع فإن المساحة المحصورة بين هاتيْن القيمتيْن تمثل ٦٨,٢٧٪ من المساحة الكلية تحت المنحنى. ويمكننا التعبير عن ذلك على النحو التالي:

المساحة المحصورة بين ($\bar{\text{س}}$ – ع، $\bar{\text{س}}$ + ع) = ٦٨,٢٧٪ من المساحة تحت المنحنى أي

أن: $\text{ع}(\bar{\text{س}} - \text{ع} \le \text{س} \le \bar{\text{س}} + \text{ع}) = ٠,٦٨٢٧$

وكذلك: $\text{ع}(\bar{\text{س}} - ١,٩٦\,\text{ع} \le \text{س} \le \bar{\text{س}} + ١,٩٦\,\text{ع}) = ٠,٩٥٤٥$

$\text{ع}(\bar{\text{س}} - ٢,٥٨\,\text{ع} \le \text{س} \le \bar{\text{س}} + ٢,٥٨\,\text{ع}) = ٠,٩٩$

كما يمكن حساب احتمال أن تقع قيمة المتغير العشوائي المتصل داخل مدى معين (أ، ب على سبيل المثال) وذلك باستخدام الصيغة التالية:

$$\text{ع}(\text{أ} \le \text{س} \le \text{ب}) = \int_{\text{أ}}^{\text{ب}} \text{ف}(\text{س})\, \text{ءس}$$

وفي حالة التوزيع المعتاد بحسب هذا الاحتمال كما يلي:

$$\text{ع}(\text{أ} \le \text{س} \le \text{ب}) = \int_{\text{أ}}^{\text{ب}} \frac{١}{\sqrt{٢\text{ط}}\,\text{ع}}\, \text{هـ}^{-\frac{١}{٢}\left(\frac{\text{س} - \bar{\text{س}}}{\text{ع}}\right)^{٢}}\, \text{ءس}$$

وبصفة عامة فإنه متى عرفنا الوسط الحسابي والانحراف المعياري للمتغير العشوائي المعتاد فإنه يمكننا حساب احتمال أن تقع قيمة المتغير داخل مدى معيــن.

ونظراً لاختلاف الوسط الحسابي والانحراف المعياري من حالة لأخرى وما يترتب على ذلك من اختلاف في شكل التوزيع فإن إجراء عملية التكامل في كل مرة يصبح أمراً على قدر كبير من الصعوبة والتعقيد. ومن هنا نشأت فكرة تكوين جداول لحساب المساحات تحت المنحنى وذلك عند مختلف النقط. وتعتمد هذه الجداول على ما يسمى بالتوزيع المعتاد المعياري والذي ستناوله فيما يلي:

التوزيع المعتاد المعياري: Standard Normal Distribution

التوزيع المعتاد المعياري هو أيضاً توزيع جرسي الشكل ويتماثل حول الوسط الحسابي. وقد نشأ هذا التوزيع باستخدام النظرية الإحصائية الهامة التي تقول إنه إذا كانت $س$ متغير عشوائي له توزيع معتاد وسطه الحسابي $\bar{س}$ وتباينه $ع^{٢}$ فإن المتغير العشوائي:

$$ي = \frac{س - \bar{س}}{ع} \qquad (١ - ١٣)$$

له أيضاً توزيع معتاد توقعه يساوي الصفر وتباينه الوحـــدة (أي أن: $\bar{س}$ = صفر، $ع^{٢}$ = ١). وهذا التوزيع يسمى بالتوزيع المعتاد المعياري (القياسي)، ودالة كثافة احتماله تأخذ الصورة التالية:

$$ف(ي) = \frac{١}{\sqrt{٢ط}} هـ^{-\frac{ي^{٢}}{٢}} \qquad (١ - ١٤)$$

حيث: $- \infty \leq ي \leq \infty$،

ط النسبة التقريبية (مقدار ثابت) وتساوي ٣,١٤١٦ تقريبا،

هـ أساس اللوغاريتم الطبيعي وتساوي ٢,٧١٨٣ تقريبا.

لذلك نجد أنه لكي نستخدم التوزيع المعتاد المعياري فإنه يلزم تحديد القيم المعيارية للمتغير والمعرَّفة في (١– ١٣).

وحيث أن الدالة (١ – ١٤) هي دالة كثافة احتمال فإنه يمكننا إثبات أن:

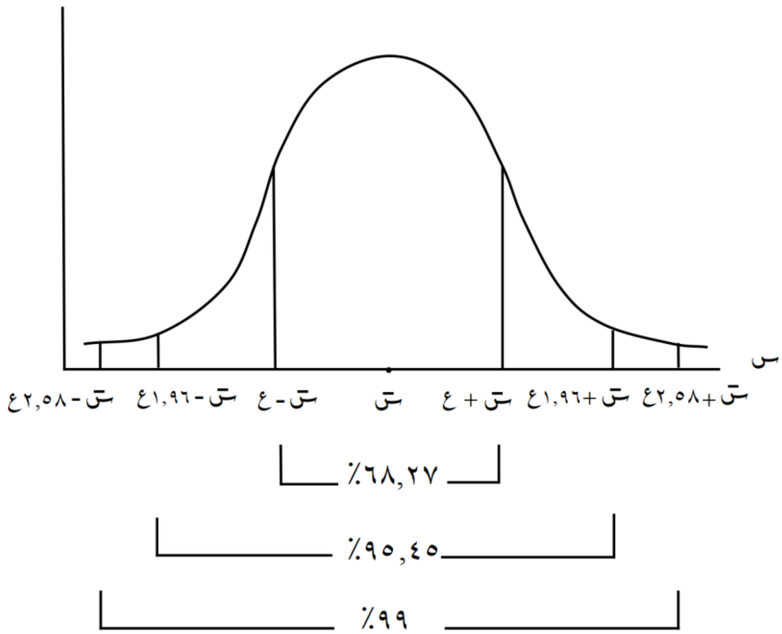
$$\int_{-\infty}^{\infty} \frac{١}{\sqrt{٢ط}} هـ^{-\frac{ي^{٢}}{٢}} ءي = ١$$

وعلى غرار ما تم في حالة التوزيع المعتاد يمكننا بسهولة استنتاج الاحتمالات التالية وذلك فيما يتعلق بالتوزيع المعتاد المعياري:

$ع(-١ \leq ي \leq ١) = ٠,٦٨$ ، $ع(-١,٩٦ \leq ي \leq ١,٩٦) = ٠,٩٥$

$ع(-٢,٥٨ \leq ي \leq ٢,٥٨) = ٠,٩٩$

وهذا يعنى أن ٩٥,٤٥٪ - على سبيل المثال - من قيم المتغير العشوائي المعتاد المعياري تقع داخل المدى -١,٩٦، ١,٩٦، وهكذا بالنسبة لبقية النسب.



الشكل (١ – ٤)
المنحنى المعتاد المعياري

جدول المساحات تحت المنحنى المعتاد المعياري:

سبق أن أوضحنا أنه يمكن إيجاد احتمال أن تقع قيمة المتغير العشوائي المعتاد داخل مدى معين وذلك بإجراء عملية تكامل محدود لدالة كثافة احتمال هذا المتغير مستخدمين في ذلك حديْ هذا المدى وكذلك الوسط الحسابي والانحراف المعياري للمتغير. ونظراً لاختلاف حديْ المدى من حالة لأخرى وكذلك اختلاف الوسط الحسابي والانحراف المعياري من توزيع لآخر، فـإن تكوين جداول لحساب المساحات المختلفة تحت كل منحنى يصبح أمراً بالغ التعقيد. وحيث أنه من الممكن تحويل أي توزيع معتاد - أياً كانت قيمة وسطه الحسابي أو انحرافه المعياري - إلى توزيع معتاد معياري واحد وسطه الحسابي يساوي الصفر وانحرافه المعياري يساوي الواحد الصحيح، فقد استُخدم هذا التوزيع في تكوين جدول المساحات تحت المنحنى المعتاد والموضح في الملحق رقم (١) في نهاية هذا الكتاب. ولأن المنحنى المعتاد المعياري هو منحنى متماثل حول الصفر فقد تم حساب المساحات (الاحتمالات) المختلفة في الجدول وذلك للقيم الموجبة فقط للمتغير ي (القيم المعيارية للمتغير س). وباستخدام خاصية التماثل هذه يمكننا حساب المساحة المحصورة (الاحتمال) بين أي قيمتيْن للمتغير (أياً كانت الإشارة سالبة أو موجبة) وكذلك حساب احتمال أن يقل المتغير عن أو يزيد على قيمة محددة.

وفيما يلي بعض الملاحظات عن التوزيع المعتاد المعياري والتي تساعد على حساب الاحتمالات المختلفة باستخدام جدول المساحات.

أ- إذا كان لدينا متغير عشوائي له توزيع معتاد وسطه الحسابي يساوي الصفر وانحرافه المعياري يساوي الوحدة فإنه يسمى متغيراً معتاداً معيارياً (قياسيّاً) ويُرمز له عادة بالرمز **ي** ولدالة كثافة احتماله بالرمز ف(**ي**). أي أن:

$$ف(ي) = \frac{1}{\sqrt{2ط}} هـ^{-\frac{ي^2}{2}}$$

وإذا كتبنا ϕ(أ) للدلالة على احتمال أن **ي** ≥ أ، أي ع(**ي** ≥ أ) فإن:

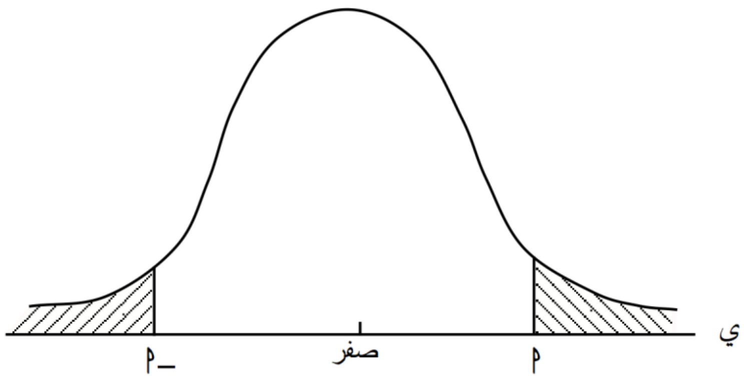
$$\phi(أ) = ع(ي \le أ) = \int_{-\infty}^{أ} ف(ي)\,ءي$$

$$= \int_{-\infty}^{أ} \frac{1}{\sqrt{2ط}} هـ^{-\frac{ي^2}{2}}\,ءي$$

وبالتالي فإن ϕ(**ي**) تعبر عن دالة التوزيع المتجمع للمتغير المعتاد المعياري (**ي**).

ب- نظراً لتماثل منحنى دالة كثافة الاحتمال للتوزيع المعتاد المعياري حول الصفر نجد أن المساحة التي تقع على يسار القيمة (-أ) تساوي المساحة التي تقع على يمين القيمة (+أ). ويتضح ذلك من الشكل (١ – ٥). فعلى سبيل المثال لو أن أ = ١,٩٦ فإننا نجد أن المساحة التي تقع على يسار القيمة -١,٩٦ تساوي المساحة التي تقع على يمين القيمة + ١,٩٦ حيث أن كلاً من المساحتيْن يساوي ٠,٠٢٥ وبصفة عامة نجد أن:

ϕ(-أ) = ١ - ϕ(أ) **(١ – ١٥)**



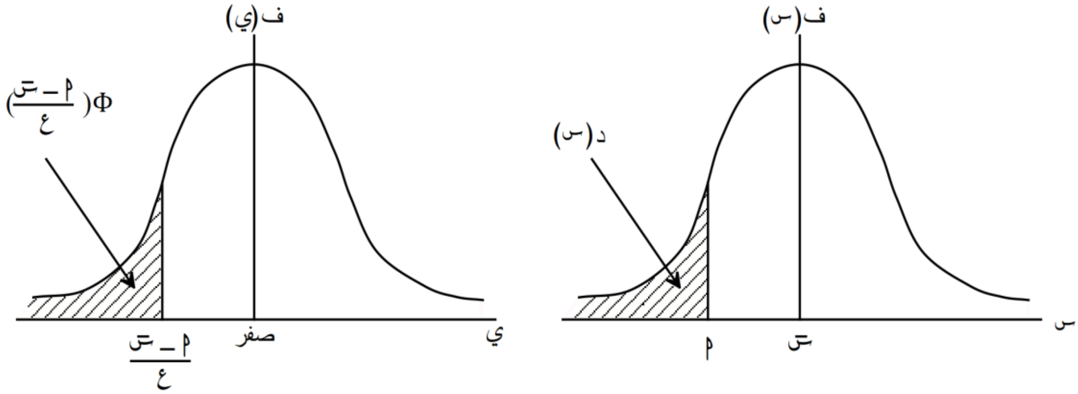
الشكل (١ – ٥)

جـ- تُستخدم القيمة الموجبة فقط للمتغير **ي** وذلك نظراً لتماثل المنحنى المعتدل المعياري حول الصفر. وأما في حالة القيم السالبة فإنه يمكن استخدام العلاقة (١ - ١٥).

د- يمكن تحويل أي متغير عشوائي س له توزيع معتاد وسطه الحسابي $\bar{\text{س}}$ وانحرافه المعياري ع إلى متغير معتاد معياري (**ي**) وذلك بوضع:

$$\text{ي} = \frac{\text{س} - \bar{\text{س}}}{\text{ع}}$$

وبذلك تتحول دالة كثافة الاحتمال ف (س) إلى ف (**ي**) كما تتحـــول د (أ) إلى $\phi\left(\frac{\text{أ} - \bar{\text{س}}}{\text{ع}}\right)$. وهنا يجب أن نلاحظ أن توزيعي س، **ي** لا يختلفان من حيث الشكل وإنما يختلفان فقط في نقطة الأصل وفي وحدات القياس المستخدمة (راجع الشكل ١ – ٦).



الشكل (١ – ٦)

فإذا كانت أ قيمة معينة للمتغير س فإنه يمكننا استنتاج ما يلي:

$$\text{ح}(\text{س} \le \text{أ}) = \text{ح}(\text{س} - \bar{\text{س}} \le \text{أ} - \bar{\text{س}})$$

$$= \text{ح}\left(\frac{\text{س} - \bar{\text{س}}}{\text{ع}} \le \frac{\text{أ} - \bar{\text{س}}}{\text{ع}}\right)$$

$$= \text{ح}\left(\text{ي} \le \frac{\text{أ} - \bar{\text{س}}}{\text{ع}}\right) = \phi\left(\frac{\text{أ} - \bar{\text{س}}}{\text{ع}}\right)$$

أي أن: $\text{د}(\text{أ}) = \phi\left(\frac{\text{أ} - \bar{\text{س}}}{\text{ع}}\right)$

هـ- جدول المساحات تحت المنحنى المعتاد المعياري والموضح في الملحق رقـم (١) في نهاية هذا الكتاب يوضح قيم **ي** المختلفة وقيم ϕ(**ي**) المناظرة لها مع ملاحظة أن المتغير **ي** يبدأ من القيمة صفر وينتهي بالقيمة ٣,٨٩ تقريباً، وفي كل مرة تزداد قيمة **ي** بالمقدار (٠,٠١). وعلى سبيل المثال نجد أن:

$$\phi(٠,٨) = ٠,٧٨٨١ \quad ، \quad \phi(١,٢٣) = ٠,٨٩٠٧$$

والآن لعله من الأيسر فهماً للقارئ أن نوضح كيفية استخدام هذا الجدول في حساب الاحتمالات المختلفة المتعلقة بالتوزيع المعتاد مستعينين في ذلك بالأمثلة التالية:

مثال (٥٠):

إذا كانت **ي** متغير معتاد معياري، احسب - باستخدام الجداول - الاحتمالات التالية:

أ- ع(**ي** > ١,٦٥)　　　　ب- ع(**ي** ≥ ٢,١٥)

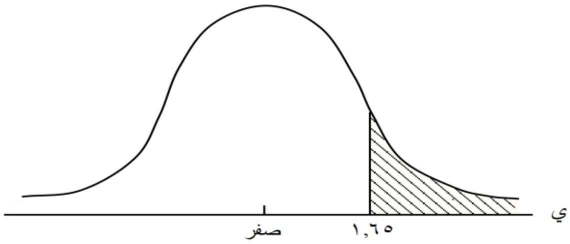
ج- ع(-١,٩٦ ≥ **ي** ≥ ١,٩٦)　　　　د- ع(**ي** ≥ - ٠,٨)

ه- ع(**ي** > - ١,٢)　　　　و- ع(١,٥ ≥**ي** ≥ ٢,١)

ز- ع(صفر ≥ **ي** ≥ ٢,١)　　　　ح- ع(-٠,٩ ≥ **ي** ≥ -٠,٣)

الحــل:

أ-

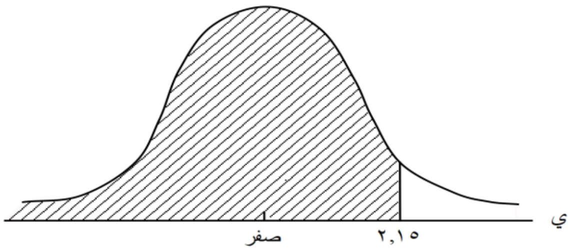


ع(**ي** > ١,٦٥) = ١- ع(**ي** ≥ ١,٦٥) = ١- ϕ(١,٦٥)

ومن الجدول نجد أن القيمة المقابلة لـ ١,٦٥ تساوي ٠,٩٥٠٥

∴ ع(**ي** > ١,٦٥) = ١ - ٠,٩٥٠٥ = ٠,٠٤٩٥

ب-



ع(**ي** ≥ ٢,١٥) = ϕ(٢,١٥)

ومن الجدول التوزيع المعتاد المعياري يمكننا ملاحظة أن:

∴ ع(**ي** ≥ ٢,١٥) = ϕ(٢,١٥) = ٠,٩٨٤٢

جـ-

ع(-١,٩٦ ≥ ي ≥ ١,٩٦) = ع(ي ≥ ١,٩٦) – ع(ي ≥ -١,٩٦)

= ϕ(١,٩٦) – ϕ(-١,٩٦)

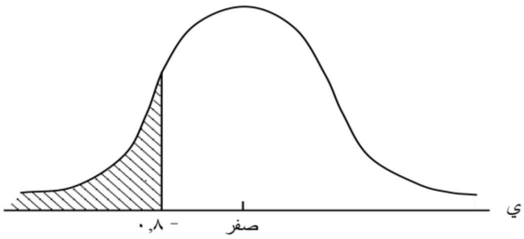
وحيث أن: ϕ(-أ) = ١ – ϕ(أ)

∴ ع(-١,٩٦ ≥ ي ≥ ١,٩٦) = ϕ(١,٩٦) – [١ - ϕ(-١,٩٦)]

= ٢ϕ(١,٩٦) – ١ = ٢ × ٠,٩٧٥ – ١ = ٠,٩٥

وهي النسبة المعروفة والمشار إليها سابقاً في أكثر من موضع من هذا الكتاب.

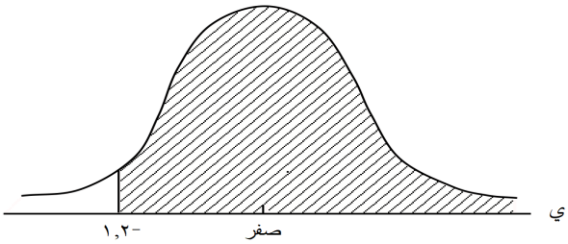
د-



ع(ي ≥ -٠,٨) = ϕ(-٠,٨)

= ١ – ϕ(٠,٨) = ١ – ٠,٧٨٨١ = ٠,٢١١٩

هـ-



ع(ي > -١,٢) = ١ – ع(ي ≥ -١,٢)

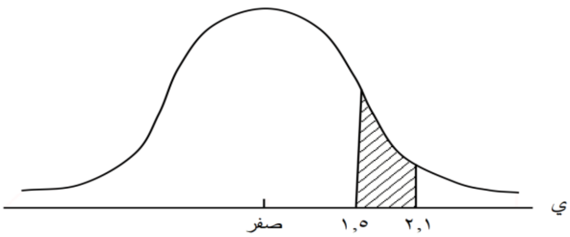
= ١ – ϕ(-١,٢)

وحيث أن: ϕ(-أ) = ١ – ϕ(أ)

∴ ع(ي > -١,٢) = ١ – [١ – ϕ(١,٢)]

= ϕ(١,٢) = ٠,٨٨٤٩

و-

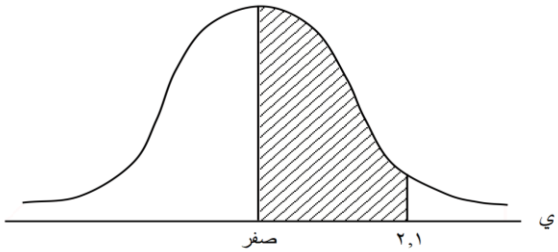


ع(١,٥ ≤ ي ≤ ٢,١) = ع(ي ≤ ٢,١) – ع(ي ≤ ١,٥)

= ϕ(٢,١) – ϕ(١,٥)

= ٠,٩٨٢١ – ٠,٩٣٣٢ = ٠,٠٤٨٩

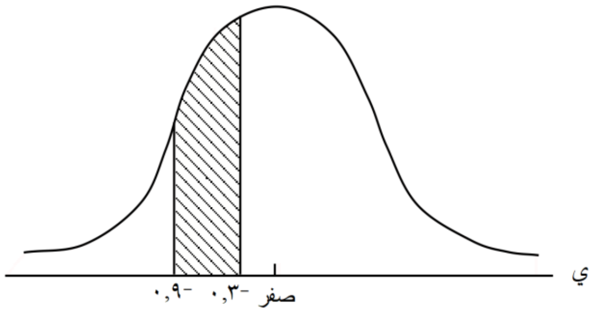
ز-



ع(صفر ≤ ي ≤ ٢,١) = ع(ي ≤ ٢,١) – ع(ي ≤ صفر)

= ϕ(٢,١) – ϕ(صفر) = ٠,٩٨٢١ – ٠,٥ = ٠,٤٨٢١

ح-



ع(-٠,٩ ≤ ي ≤ -٠,٣) = ع(ي ≤ -٠,٣) – ع(ي ≤ -٠,٩)

= ϕ(-٠,٣) – ϕ(-٠,٩)

= ١ – ϕ(٠,٣) – [١ – ϕ(٠,٩)]

= ϕ(٠,٩) – ϕ(٠,٣)

= ٠,٨١٥٩ – ٠,٦١٧٩ = ٠,١٩٨

وتجدر الإشارة إلى أننا حاولنا أن يشمل هذا المثال العديد من الحالات المتنوعة والتي يمكننا أن نلخصها بشكل عام على النحو التالي:

- ع(ي ≤ أ) = ϕ(أ)
- ع(ي > أ) = ١ – ϕ(أ)
- ع(أ ≤ ي ≤ ب) = ϕ(ب) – ϕ(أ)
- ع(-ب ≤ ي ≤ -أ) = ϕ(ب) – ϕ(أ)

حيث: أ < ب

مثال (٥١):

إذا كانت الدرجات التي حصل عليها طلاب الفرقة الثانية في مادة الإحصاء في إحدى كليات التجارة تتبع التوزيع المعتاد بمتوسط مقداره ١٤ درجة وانحراف معياري مقداره ٣ درجات. فإذا تم اختيار واحد من الطلاب بطريقة عشوائية، احسب ما يلي:

أ- احتمال أن تقل درجة الطالب عن ١٦ درجة.

ب- احتمال أن تزيد درجة الطالب عن ١٥ درجة.

ج- احتمال أن تتراوح درجة الطالب ما بين ١١، ١٨ درجة.

د- احتمال أن تتراوح درجة الطالب ما بين ١٧، ٢٠ درجة.

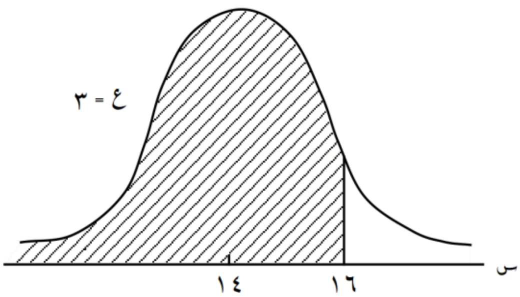
ه- احسب الدرجة التي حصل عليها - كحد أقصى - ٢٥٪ من إجمالي عدد الطلاب.

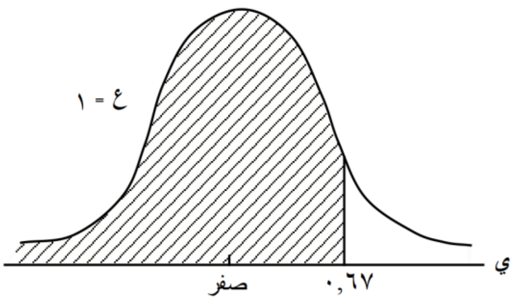
و- احسب الدرجة التي يحصل على أكبر منها ٣٥٪ من إجمالي عدد الطلاب.

الحــل:

بافتراض أن س هي المتغير العشوائي المعتاد الذي يمثل درجات الإحصاء للطلاب، فإنه يمكن حساب الاحتمالات المطلوبة على النحو التالي:

أ-





حيث أن القيمة المعيارية هي:

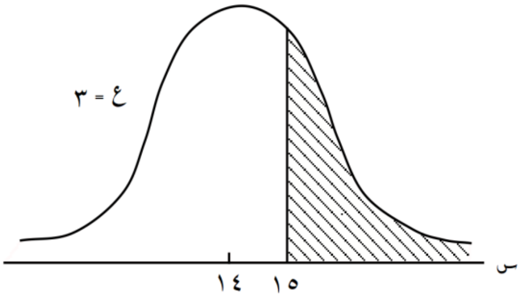
$$\text{ي} = \frac{\text{س} - \bar{\text{س}}}{\text{ع}}$$

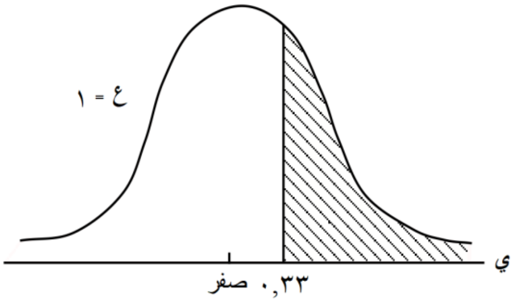
∴ القيمة المعيارية للدرجة ١٦ = $\frac{١٦ - ١٤}{٣}$ = ٠,٦٧

$$\text{ع}(\text{س} \leq ١٦) = \text{ع}\left(\frac{\text{س} - ١٤}{٣} \leq \frac{١٦ - ١٤}{٣}\right)$$

$$= \text{ع}(\text{ي} \leq ٠,٦٧) = \phi(٠,٦٧) = ٠,٧٤٨٦$$

ب-





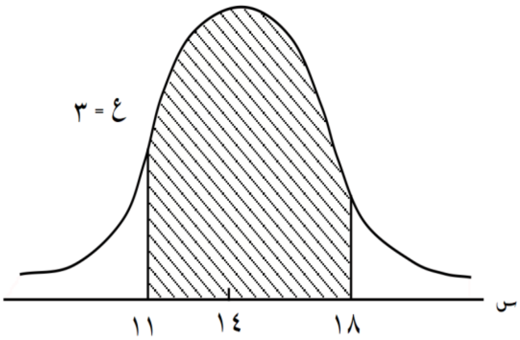
ع(س > ١٥) = ١ – ع(س ≤ ١٥)

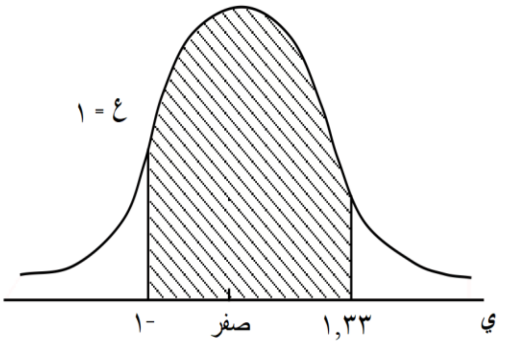
ع(س ≤ ١٥) = ع$\left(\frac{\text{س} - ١٤}{٣} \leq \frac{١٥ - ١٤}{٣}\right)$

= ع(ي ≤ ٠,٣٣) = ϕ(٠,٣٣)

∴ ع(س > ١٥) = ١ – ϕ(٠,٣٣) = ١ – ٠,٦٢٩٣ = ٠,٣٧٠٧

ج-





ع(١١ ≤ س ≤ ١٨) = ع(س ≤ ١٨) – ع(س ≤ ١١)

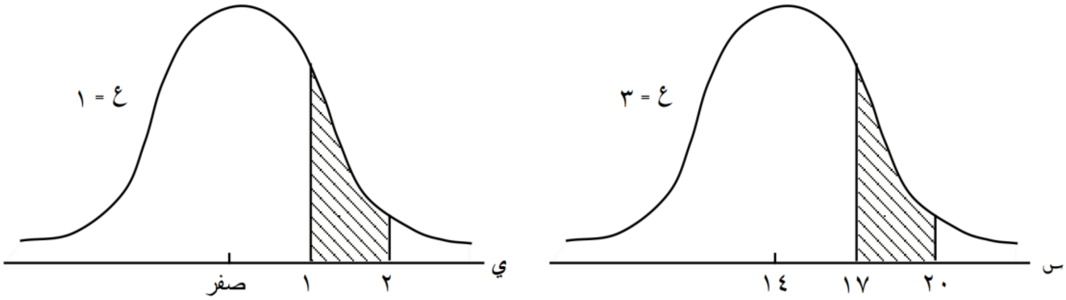
= ع$\left(\frac{\text{س} - ١٤}{٣} \leq \frac{١٨ - ١٤}{٣}\right)$ – ع$\left(\frac{\text{س} - ١٤}{٣} \leq \frac{١١ - ١٤}{٣}\right)$

= ع(ي ≤ ١,٣٣) – ع(ي ≤ -١)

= ϕ(١,٣٣) – [١ – ϕ(١)] = ϕ(١,٣٣) + ϕ(١) – ١

= ٠,٩٠٨٢ + ٠,٨٤١٣ – ١ = ١,٧٤٩٥ – ١ = ٠,٧٤٩٥

د-



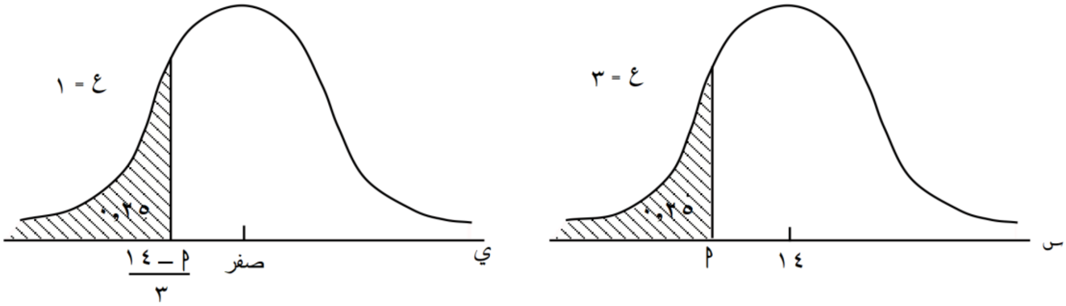
ع(١٧ ≥ س ≥ ٢٠) = ع(س ≥ ٢٠) – ع(س ≥ ١٧)

$$= ع\left(\frac{س - ١٤}{٣} \geq \frac{٢٠ - ١٤}{٣}\right) - ع\left(\frac{س - ١٤}{٣} \geq \frac{١٧ - ١٤}{٣}\right)$$

= ع(ي ≥ ٢) – ع(ي ≥ ١) = ϕ(٢) – ϕ(١)

= ٠,٩٧٧٢ + ٠,٨٤١٣ = ٠,١٣٥٩

هـ -



حيث أن الدرجة المراد تحديدها هي التي يحصل عليها - كحد اقصى - ٢٥٪ من إجمالي عدد الطلاب، فإن قيمة الدرجة تتحدد بتلك النقطة التي تبلغ المساحة الواقعة على يسار العمود المقام عندها ٠,٢٥ وذلك فيما يتعلق بالمنحنى المعتاد المعياري. وعليه يمكن تحديد قيمة الدرجة المطلوبة على النحو الموضح فيما يلي:

ع(س ≥ أ) = ٠,٢٥

وهذا يعنى ان:

$$ع\left(\frac{س - ١٤}{٣} \geq \frac{أ - ١٤}{٣}\right) = ٠,٢٥$$

أي أن: $ع\left(ي \geq \frac{أ - ١٤}{٣}\right) = ٠,٢٥$

$$\therefore \phi\left(\frac{أ - ١٤}{٣}\right) = ٠,٢٥$$

وحيث أن النقطة $(\frac{أ - ١٤}{٣})$ تقع على يسار النقطة (صفر) والتي تمثل الوسط الحسابي للتوزيع المعتاد المعياري، فإنه يمكن استنتاج أن النقطة $(\frac{أ - ١٤}{٣})$ هي مقدار سالب. وحيث أن جدول المساحات يشمل فقط القيم المعيارية الموجبة فإننا نحصل على:

$$\phi(\frac{أ - ١٤}{٣}) = ١ - \phi(\frac{١٤ - أ}{٣}) = ٠,٢٥$$

$$\therefore \phi(\frac{١٤ - أ}{٣}) = ٠,٧٥$$

وبالبحث في جدول المساحة تحت المنحنى المعتاد المعياري يمكننا ملاحظة أن قيمة **ي** المناظرة لـ ϕ(**ي**) = ٠,٧٥ هي ٠,٦٧٥ تقريباً. أي أن:

$$ي = \frac{١٤ - أ}{٣} = ٠,٦٧٥$$

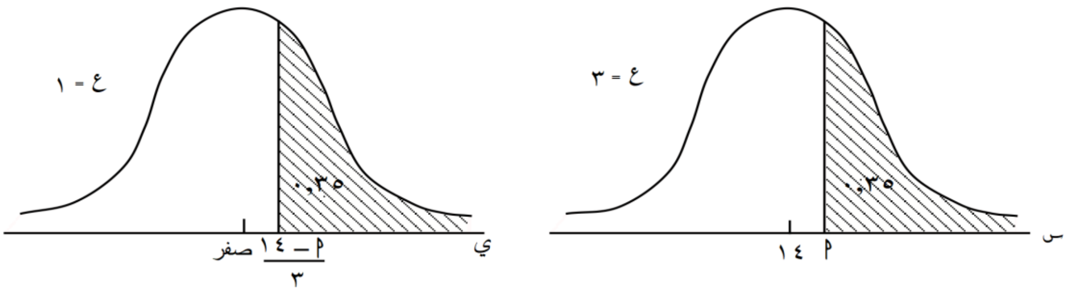
ومنها نحصل على:

أ = ١٤ – ٣ × ٠,٦٧٥ = ١٢ درجة تقريباً.

أي أن ٢٥٪ من إجمالي عدد الطلاب تبلغ درجاتهم ١٢ درجة على الأكثر.

وبلغة الاحتمالات يمكننا القول بأن: ع(س ≤ ١٢) = ٠,٢٥.

و-



حيث أن الدرجة المراد تحديدها هي التي يحصل على أكبر منها ٣٥٪ من إجمالي عدد الطلاب، فإن قيمة هذه الدرجة تتحدد بتلك النقطة التي تبلغ المساحة الواقعة على يمين العمود المقام عندها ٠,٣٥ وذلك فيما يتعلق بالمنحنى المعتاد المعياري.

لذلك يمكن تحديد قيمة الدرجة المطلوبة على النحو التالي:

ع(س > أ) = ٠,٣٥ $\therefore$ ع(س ≤ أ) = ٠,٦٥

وهذا يعنى ان:

$$ع(\frac{س - ١٤}{٣} \leq \frac{أ - ١٤}{٣}) = ٠,٦٥$$

أي أن: $\text{ع}(\text{ي} \ge \frac{\text{أ} - ١٤}{٣}) = ٠,٦٥$

$\therefore \phi(\frac{\text{أ} - ١٤}{٣}) = ٠,٦٥$

وحيث أن النقطة $(\frac{\text{أ} - ١٤}{٣})$ تقع على يمين النقطة (صفر) والتي تمثل الوسط الحسابي للتوزيع المعتاد المعياري، فإنه يمكن استنتاج أن النقطة $(\frac{\text{أ} - ١٤}{٣})$ هي مقدار موجب. وبالتالي يمكننا البحث مباشرةً في جدول المساحة تحت المنحنى المعتاد عن قيمة ي المناظرة لـ $\phi(\text{ي}) = ٠,٦٥$ والتي نجد أنها تساوي تقريباً القيمة ٠,٣٨٥، أي أن: وحيث أن جدول المساحات يشمل فقط القيم المعيارية الموجبة فإننا نحصل على:

$$\text{ي} = \frac{\text{أ} - ١٤}{٣} = ٠,٣٨٥$$

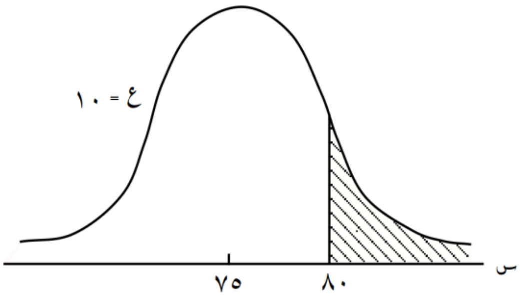
$\therefore$ أ = ١٤ + ٣ × ٠,٣٨٥ = ١٥ درجة تقريباً.

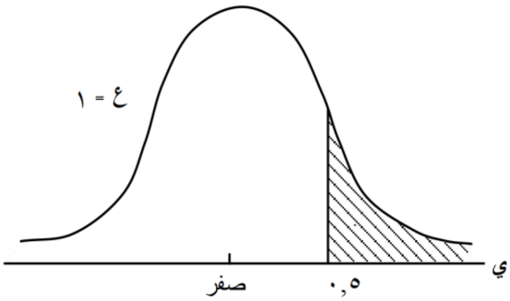
أي أن الدرجة التي يحصل على أكبر منها ٣٥٪ من إجمالي عدد الطلاب هي تقريباً ١٥ درجة.

مثال (٥٢):

إذا بلغ عدد الطلاب المتقدمين للإسكان الجامعي في جامعة ما ٢٠٠٠ طالب، وكانت النسبة المئوية لنجاح هؤلاء الطلاب تتبع توزيعاً معتاداً بمتوسط ٧٥ وانحراف معياري ١٠. فإذا كان الإسكان الجامعي لا يقبل سوى الطلاب الذين تزيد نسبة نجاحهم عن ٨٠٪، احسب عدد الطلاب المتوقع قبولهم في الاسكان الجامعي.

الحــل:





لمعرفة عدد الطلاب المتوقع قبولهم في الإسكان الجامعي فإنه يلزم بداية حساب احتمال أن تزيد نسبة نجاح الطالب (س) عن ٨٠٪. بعد ذلك يمكن حساب العدد المتوقَّع قبوله بضرب الاحتمال المحسوب في إجمالي عدد الطلاب المتقدمين للإسكان الجامعي وهو ٢٠٠٠ طالب.

$\text{ع}(\text{س} > ٨٠) = ١ - \text{ع}(\text{س} \le ٨٠)$

وحيث أن:

$$\text{ع}(\text{س} \le \text{٨٠}) = \text{ع}\left(\frac{\text{س} - \text{٧٥}}{\text{١٠}} \le \frac{\text{٨٠} - \text{٧٥}}{\text{١٠}}\right)$$

$$= \text{ع}(\text{ي} \le \text{٠,٥}) = \phi(\text{٠,٥}) = \text{٠,٦٩١٥}$$

∴ الاحتمال المطلوب = ع(س > ٨٠)

= ١ ـ ٠,٦٩١٥ = ٠,٣٠٨٥

∴ عدد الطلاب المتوقع قبولهم في الإسكان الجامعي = ٢٠٠٠ × ٠,٣٠٨٥

= ٤٦١٧ طالباً

مثال (٥٣):

إذا كان عدد الجمهور الذي يحضر مباراة في كرة القدم ما بين فريقي الأهلي والزمالك يتبع توزيعًا معتاداً بمتوسط مقداره ٣٥٠٠٠ متفرج وانحراف معياري مقداره ٢٥٠٠ متفرج.

احسب ما يلي:

أ- احتمال أن يحضر إحدى المباريات المقامة بين الفريقين ما بين ٣٠٠٠٠ - ٣٧٥٠٠ متفرج.

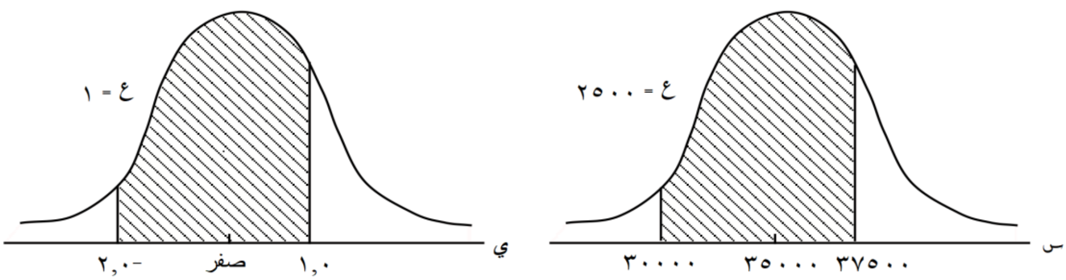
ب- احتمال ألا يزيد عدد جمهور الحاضرين عن ٣٨٠٠٠ متفرج.

ج- الحد الأعلى لعدد الحاضرين في ٩٠٪ من المباريات المقامة بين الفريقيْن.

الحـل:

بافتراض أن المتغير س هو المتغير المعتاد الذي يمثل عدد الحاضرين لمباريات كرة القدم بين فريقي الأهلي والزمالك، فإنه يمكن الإجابة على المطلوب كما يلي:

أ-



$$\text{ع}(\text{٣٠٠٠٠} \le \text{س} \le \text{٣٧٥٠٠}) = \text{ع}(\text{س} \le \text{٣٧٥٠٠}) - \text{ع}(\text{س} \le \text{٣٠٠٠٠})$$

وحيث أن:

$$\text{ع}(\text{س} \le \text{٣٧٥٠٠}) = \text{ع}\left(\frac{\text{س} - \text{٣٥٠٠٠}}{\text{٢٥٠٠}} \le \frac{\text{٣٧٥٠٠} - \text{٣٥٠٠٠}}{\text{٢٥٠٠}}\right)$$

$$= \text{ع}(\text{ي} \le \text{١,٠}) = \phi(\text{١,٠}) = \text{٠,٨٤١٣}$$

وكذلك:

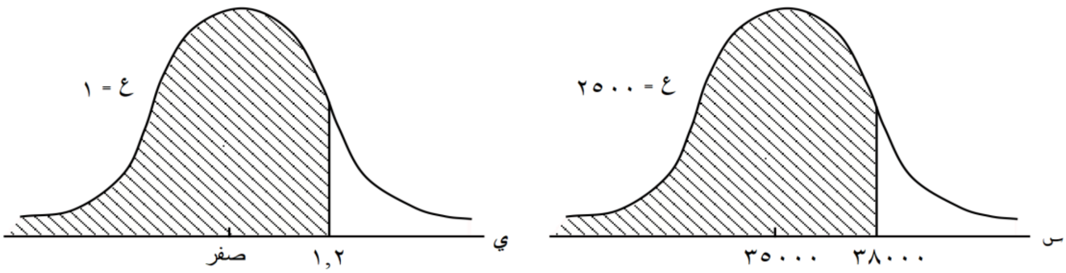
$$\text{ع}(\text{س} \le ٣٠٠٠٠) = \text{ع}\left(\frac{\text{س} - ٣٥٠٠٠}{٢٥٠٠} \le \frac{٣٠٠٠٠ - ٣٥٠٠٠}{٢٥٠٠}\right)$$

$$= \text{ع}(\text{ي} \le -٢) = ١ - \phi(٢)$$

$$= ١ - ٠,٩٧٧٢ = ٠,٠٢٢٨$$

∴ الاحتمال المطلوب = ٠,٨٤١٣ − ٠,٠٢٢٨ = ٠,٨١٨٥

ب-

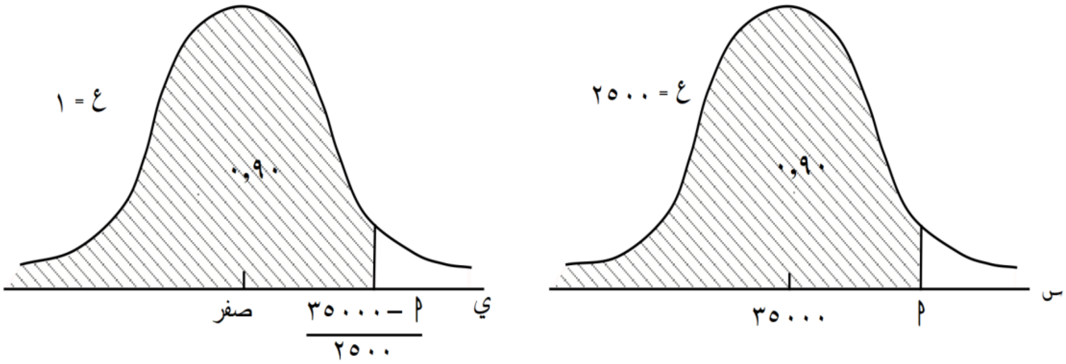


الاحتمال المطلوب = $\text{ع}(\text{س} \le ٣٨٠٠٠)$

$$= \text{ع}\left(\frac{\text{س} - ٣٥٠٠٠}{٢٥٠٠} \le \frac{٣٨٠٠٠ - ٣٥٠٠٠}{٢٥٠٠}\right)$$

$$= \text{ع}(\text{ي} \le ١,٢) = \phi(١,٢) = ٠,٨٨٤٩$$

ج-



يمكن التعبير عن المطلوب في الصورة الاحتمالية التالية:

$$\text{ع}(\text{س} \le \text{أ}) = ٠,٩$$

وقيمة أ هنا هي التي تحدد الحد الأعلى لعدد الحاضرين في ٩٠٪ من المباريات المقامة بين الفريقين. أي أن:

$$\text{ع}\left(\frac{\text{س} - ٣٥٠٠٠}{٢٥٠٠} \le \frac{\text{أ} - ٣٥٠٠٠}{٢٥٠٠}\right) = ٠,٩$$

$\therefore$ ع(ي $\geq \frac{أ - ٣٥٠٠٠}{٢٥٠٠}$) = ٠,٩

أي أن: $\phi\left(\frac{أ - ٣٥٠٠٠}{٢٥٠٠}\right)$ = ٠,٩

وحيث أن النقطة $\left(\frac{أ - ٣٥٠٠٠}{٢٥٠٠}\right)$ تقع على الجانب الأيمن من النقطة (صفر) في المنحنى المعتاد فهي قيمة موجبة. وعليه يمكن البحث في جدول المساحة تحت المنحنى المعتاد المعياري عن قيمة ي المناظــــــرة لـ ϕ(ي) = ٠,٩ حيث نجد أنها تساوي تقريباً القيمة ١,٢٨ وبذلك نجد أن:

$$ي = \frac{أ - ٣٥٠٠٠}{٢٥٠٠} = ١,٢٨$$

أي أن: أ = ٣٥٠٠٠ + ٢٥٠٠ × ١,٢٨ = ٣٨٢٠٠

أي أنه في ٩٠٪ من المباريات المقامة بين الفريقيْن يتوقع ألا يزيد عدد المتفرجين في أية مباراة عن ٣٨٢٠٠ متفرجاً.

التوزيع المعتاد كتقريب لبعض التوزيعات المنفصلة:

أ- التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن:

يُستخدم التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن وذلك عندما تكون ن كبيرة وتقترب قيمتا ل، ك من بعضهما أي أن: ل = ك ≈ $\frac{١}{٢}$. ويرى بعض الإحصائيين أن التوزيع المعتاد يمكن استخدامه كتقريب جيد للتوزيع ثنائي الحديْن إذا كانت ن ≤ ٣٠ وكل من ن ل، ن ك تساوي ٥ على الأقل. كما يرى البعض الآخر أنه لكي يكون التقريب معقولاً ومناسباً فإن قيمة الاحتمال يجب أن تنحصر ما بين ٠,١، ٠,٩ أي أن: ٠,١ ≥ ل ≥ ٠,٩ وأيضا كل من ن ل، ن ك تزيد عن ٥ (أي أن: ن ل ≤ ٥ و ن ك ≤ ٥). وإذا ما تحققت هذه الشروط فإن التواء التوزيع ثنائي الحديْن يكاد ينعدم كما أن معامل تفلطحه يقترب في قيمته من ٣. وهذا يعنى اقتراب التوزيـع بدرجة كافية من التوزيع المعتاد. وبصفة عامة إذا كانت الدقة في التقدير مطلوبة فإن زيادة ن ل، ن ك عن الحدود المشار إليها يصبح أمراً مرغوب فيه. وهنا تجدر الإشارة إلى أن التوزيع ثنائي الحديْن هو توزيع متماثل في الحالة ل = ك = $\frac{١}{٢}$ وذلك بصرف النظر عن قيمة ن. كذلك فإنه كلما اقتربت ل من القيمة $\frac{١}{٢}$ كلما قلت الحاجة إلى زيادة ن.

هذا وفي حالة استخدام التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن يمكن التعبير حينئذ عن التوزيع ثنائي الحديْن بمدرج تكراري ومن ثم منحنى تكراري معتاد يمكن الاستفادة من خصائصه في إيجاد قيم الاحتمالات المختلفة. ولرسم المدرج التكراري لتوزيع منفصل فإننا نفترض أن النقطة المطلوب عندها الاحتمال في حالة التوزيع المنفصل تمثل مركز الفئة في حالة اعتباره متغيراً متصلاً. كما نعتبر أن طول الفئة يساوي الواحد الصحيح. وهذا يعنى أن المدرج التكراري يتكون في هذه الحالة من مستطيلات طول قاعدة كل منها يساوي الواحد الصحيح وارتفاعه يساوي الاحتمال المناظر لمركز الفئة. فعلى سبيل المثال لو كان المطلوب في حالة المتغير المنفصل هو إيجاد الاحتمال عند النقطة ٤، فإنه يمكن رسم مستطيل يقام أحد أضلاعه عند النقطة ٣,٥ بينما يقام الآخر عند النقطة ٤,٥ . وبعد إنشاء المدرج التكراري والمنحنى التكراري بهذه الطريقة يكون المنحنى معتدلاً بمتوسط وانحراف معياري يساويان على الترتيب الوسط الحسابي والانحراف المعياري للتوزيع ثنائي الحديْن.

وحيث أن الوسط الحسابي والانحراف المعياري يمكن تحديدهما على الصورة:

الوسط الحسابي = ن ل

الانحراف المعياري = $\sqrt{\text{ن ل ك}}$

فإنه يمكن إيجاد القيمة المعيارية (**ي**) باستخدام العلاقة:

$$\text{ي} = \frac{\text{س} - \text{ن ل}}{\sqrt{\text{ن ل ك}}}$$

حيث **ي** متغير عشوائي يتبع التوزيع المعتاد المعياري.

وفي هذا الشأن وبناء على ما تقدم فإنه عند استخدام التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن تجدر الإشارة إلى ما يلي:

- الاحتمال عند النقطة ٤ على سبيل المثال في حالة التوزيع ثنائي الحديْن يساوي في قيمته مساحة الجزء المحصور ما بين ٣,٥، ٤,٥ تحت المنحنى المعتدل.

ويمكن ايجاد هذا الاحتمال بتحويل كل من هاتيْن القيمتيْن إلى قيمة معيارية، ثم يحسب الاحتمال بالطريقة المشار اليها سابقاً. أي أن:

$$\text{ع}(\text{س} = ٤) = \text{ع}(\text{س} \le ٤,٥) - \text{ع}(\text{س} \le ٣,٥)$$

بعد ذلك يمكن حساب كل من الاحتماليْن في الطرف الأيسر وذلك بتحويل س إلى **ي** بإتباع الأسلوب المشار إليه في الأمثلة السابقة.

- عند استخدام التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن فإنه يجب معاملة المتغير المنفصل كمتغير متصل. حينئذ يجب إجراء التصحيح الملائم على الاحتمال المطلوب في حالة التوزيع ثنائي الحديْن. وهذا التعديل تتوقف طبيعته على شكل الاحتمال وذلك على النحو الموضح فيما يلي:

الاحتمال في حالة التوزيع ثنائي الحديْن	الاحتمال المناظر في حالة التوزيع المعتاد
ع(ر ≤ أ)	ع(س > أ - ٠,٥)
ع(ر > أ)	ع(س > أ - ٠,٥)
ع(ر ≥ أ)	ع(س ≥ أ + ٠,٥)
ع(ر < أ)	ع(س < أ + ٠,٥)

أي أنه يتم طرح ٠,٥ في الحالتيْن (أكبر من أو يساوي، أقل من) كما يتم إضافة نصف في الحالتيْن (أقل من أو يساوي، أكبر من). وهنا تجدر الإشارة إلى أنه إذا كانت ں كبيرة فإن إجراء هذا التصحيح لا يؤدى إلى تغيير يُذكر في النتائج. ومع أنه لا توجد قاعدة عامة لتحديد جدوى إجراء التصحيح من عدمه، إلا أنه يمكن القول إن التصحيح قد يصبح أمراً غير ذي جدوى إذا لم يؤد إلى تغيير في الرقم العشري الثاني في قيمة **ي**.

مثال (٥٤):

لو افترضنا أن دراسة إحصائية أظهرت أن ٢٠٪ من الطلاب الحاصلين على نسبة نجاح ٧٥٪ فأكثر في الثانوية العامة يحصلون على تقدير عام جيد على الأقل في دراستهم لمرحلة البكالوريوس. احسب احتمال أن يحصل خمسة طلاب على تقدير عام جيد على الأقل في دراستهم لمرحلة البكالوريوس وذلك من بين ٣٠ طالباً حصلوا على نسبة نجاح ٧٥٪ فأكثر في الثانوية العامة، مستخدماً في ذلك:

أ- التوزيع ثنائي الحديْن.

ب- التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن.

علِّق على ما تحصل عليه من نتائج.

الحــل:

أ- باستخدام التوزيع ثنائي الحديْن:

حيث أن:

ں = ٣٠ ، ل = ٠,٢ ، ك = ٠,٨ ، ر = ٥

∴ الاحتمال المطلوب = $^{٣٠}ق_{٥}(٠,٢)^{٥}(٠,٨)^{٢٥}$

$$= \frac{٣٠!}{٥!٢٥!}(٠,٢)^{٥}(٠,٨)^{٢٥}$$

$$= \frac{٣٠ \times ٢٩ \times ٢٨ \times ٢٧ \times ٢٦}{٥ \times ٤ \times ٣ \times ٢}(٠,٠٠٠٣٢)(٠,٠٠٣٧٨) = ٠,١٧$$

ب- باستخدام التوزيع المعتاد:

حيث أن:

ن = ٣٠ ، ن ل= ٣٠(٠,٢) = ٦ وهي أكبر من ٥ ، ن ك = ٣٠(٠,٨) = ٢٤

وهي أيضاً أكبر من ٥.

فإنه يمكن استخدام التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن، حيث يمكن حساب الاحتمال المطلوب على النحو التالي:

حيث أن عدد الطلاب يمثله متغير منفصل فإنه يمكن حساب الاحتمال باستخدام التوزيع المعتاد وذلك باستخدام الصورة التالية:

ع(ر = ٥) = ع(س ≤ ٥,٥) – ع(س ≤ ٤,٥)

وباستخدام التوزيع المعتاد يمكن حساب كل من الاحتماليْن ع(س ≤ ٥,٥) و ع(س ≤ ٤,٥) ولكن بعد إيجاد الوسط الحسابي والانحراف المعياري وذلك كما يلي:

الوسط الحسابي = ن ل = ٣٠(٠,٢) = ٦،

الانحراف المعياري = $\sqrt{\text{ن ل ك}}$

$= \sqrt{٣٠(٠,٢)(٠,٨)} = \sqrt{٤,٨} = ٢,١٩$

∴ ع(س ≤ ٥,٥) = ع$\left(\frac{\text{س} - ٦}{٢,١٩} \le \frac{٥,٥ - ٦}{٢,١٩}\right)$

= ع(ي ≤ -٠,٢٣) = ϕ(-٠,٢٣)

= ١ – ϕ(٠,٢٣) = ١ – ٠,٥٩١ = ٠,٤٠٩

وكذلك:

ع(س ≤ ٤,٥) = ع$\left(\frac{\text{س} - ٦}{٢,١٩} \le \frac{٤,٥ - ٦}{٢,١٩}\right)$

= ع(ي ≤ -٠,٦٨٥) = ϕ(-٠,٦٨٥)

= ١ – ϕ(٠,٦٨٥) = ١ – ٠,٧٥٣٣ = ٠,٢٤٦٧

∴ الاحتمال المطلوب = ع(س ≤ ٥,٥) – ع(س ≤ ٤,٥)

= ٠,٤٠٩ – ٠,٢٤٦٧ = ٠,١٦٢٣

أي أن الاحتمال المحسوب باستخدام التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن يبلغ في قيمته ٩٤,٢٪ من قيمة الاحتمال الحقيقي. وهذا يعنى بدوره أن التوزيع المعتاد يعتبر في هذه الحالة تقريباً جيداً للتوزيع ثنائي الحديْن. ولعل ذلك يبرر استخدام التوزيع المعتاد بدلاً من التوزيع ثنائي الحديْن في مثل تلك الحالات خاصة إذا ما أخذنا في الاعتبار صعوبة العمليات الحسابية عند استخدامنا للتوزيع الأخير.

مثال (٥٥):

بالرجوع إلى بيانات المثال السابق، احسب احتمال أن يحصل أربعة طلاب على الأكثر على تقديم عام جيد على الأقل وذلك من بين الثلاثين طالباً المشار إليهم مستخدماً في ذلك:

أ- التوزيع ثنائي الحدّين.

ب- التوزيع المعتاد كتقريب للتوزيع ثنائي الحدّين.

الحــل:

أ- باستخدام التوزيع ثنائي الحدّين:

حيث أن: ن = ٣٠ ، ل = ٠,٢ ، ك = ٠,٨ ، ر = صفر، ١، ٢، ٣، ٤

∴ ع(ر = ٠) = ^{٣٠}ق$_{٠}$(٠,٢)^{صفر}(٠,٨)^{٣٠}

= (٠,٨)^{٣٠} = ٠,٠٠١٢٤

ع(ر = ١) = ^{٣٠}ق$_{١}$(٠,٢)(٠,٨)^{٢٩}

= ٣٠(٠,٢)(٠,٠٠١٥٥) = ٠,٠٠٩٣

ع(ر = ٢) = ^{٣٠}ق$_{٢}$(٠,٢)^{٢}(٠,٨)^{٢٨}

= $\frac{!٣٠}{!٢!٢٨}$(٠,٢)^{٢}(٠,٨)^{٢٨}

= $\frac{٣٠ \times ٢٩}{٢}$(٠,٠٤)(٠,٠٠١٩٣) = ٠,٣٣٥٨

ع(ر = ٣) = ^{٣٠}ق$_{٣}$(٠,٢)^{٣}(٠,٨)^{٢٧}

= $\frac{!٣٠}{!٣!٢٧}$(٠,٢)^{٣}(٠,٨)^{٢٧}

= $\frac{٣٠ \times ٢٩ \times ٢٨}{٣ \times ٢}$(٠,٠٠٨)(٠,٠٠٢٤٢) = ٠,٠٧٨٦

ع(ر = ٣) = ^{٣٠}ق$_{٤}$(٠,٢)^{٤}(٠,٨)^{٢٦}

= $\frac{!٣٠}{!٤!٢٦}$(٠,٢)^{٤}(٠,٨)^{٢٦}

= $\frac{٣٠ \times ٢٩ \times ٢٨ \times ٢٧}{٤ \times ٣ \times ٢}$(٠,٠٠١٦)(٠,٠٠٣٠٢) = ٠,١٣٢٤

∴ الاحتمال المطلوب = ع(ر ≤ ٤)

= ع(ر = صفر) = ع(ر = ١) + ع(ر = ٢) + ع(ر = ٣) + ع(ر = ٤)

= ٠,٠٠١٢٤ + ٠,٠٠٩٣ + ٠,٠٣٣٥٨ + ٠,٠٧٨٦ + ٠,١٣٢٤٢ = ٠,٢٥٥١

ب- باستخدام التوزيع المعتاد:

في المثال السابق تم حساب كل من الوسط الحسابي والانحراف المعياري وكانا كما يلي:

الوســط الحسابي = ٦ ، الانحراف المعياري = ٢,١٩

ولكي نتعامل مع عدد الطلاب في هذه الحالة كمتغير متصل نجد أن الاحتمال المطلوب هو:

$$ع(ر \leq ٤) = ع(س \leq ٤,٥)$$

∴ الاحتمال المطلوب = ع(س ≤ ٤,٥)

$$= ع(\frac{س - ٦}{٢,١٩} \leq \frac{٤,٥ - ٦}{٢,١٩})$$

$$= ع(ي \leq -٠,٦٨٥) = \phi(-٠,٦٨٥)$$

$$= ١ - \phi(٠,٦٨٥) = ١ - ٠,٧٥٣٣ = ٠,٢٤٦٧$$

وهنا أيضاً يمكننا استنتاج أن التوزيع المعتاد يُعتبر تقريباً جيداً للتوزيع ثنائي الحدّيْن حيث تبلغ قيمة الاحتمال المحسوب باستخدام التقريب ٩٦,٧٪ من القيمة الحقيقية للاحتمال.

مثال (٥٦):

المطلوب حل المثال السابق باستخدام التوزيع المعتاد مع عدم اعتبار عدد الطلاب متغيراً متصلاً مع التعليق على ما تحصل عليه من إجابة.

الحــل:

أ- في حالة عدم اعتبار عدد الطلاب كمتغير متصل فإننا نجد أن الاحتمال المطلوب عند استخدامنا للتوزيع المعتاد هو: ع(س ≤ ٤)

∴ الاحتمال المطلوب = ع(س ≤ ٤)

$$= ع(\frac{س - ٦}{٢,١٩} \leq \frac{٤ - ٦}{٢,١٩})$$

$$= ع(ي \leq -٠,٩١) = \phi(-٠,٩١)$$

$$= ١ - \phi(٠,٩١) = ١ - ٠,٨١٨٦ = ٠,١٨١٤$$

وحيث أن القيمة الحقيقة للاحتمال (أي المحسوبة باستخدام التوزيع ثنائي الحدّيْن) يساوي ٠,٢٥٥١ فإنه يمكن استنتاج أن قيمة الاحتمال في حالة عدم التعامل مع عدد الطلاب كمتغير متصل بلغت ٧١,١٪ فقط من القيمة الحقيقية للاحتمال. وهذا يعنى أنه ما لم نعامل المتغير المنفصل - في حالة استخدام التوزيع المعتاد - كمتغير متصل فإن التقريب الناتج لا يكون على نفس الدرجة من الدقة.

ب- استخدام التوزيع المعتاد كتقريب لتوزيع بواسون:

يمكن استخدام التوزيع المعتاد كتقريب لتوزيع بواسون وذلك في الحالات التي تكون فيها λ كبيرة نسبياً. ويرى بعض الإحصائيين أن الحد الملائم لقيمة λ حتى يمكننا استخدام هذا التقريب هو $\lambda \geq$ ١٠.

هذا مع ملاحظة أن الوسط الحسابي والانحراف المعياري يعطيان على الصورة التالية:

الوســط الحسابي = λ ، الانحراف المعياري = $\sqrt{\lambda}$

وبالتالي نجد أن القيمة المعيارية (ي) تتحدد كما يلي:

$$ي = \frac{س - \lambda}{\sqrt{\lambda}}$$

ويُستنتج من ذلك أنه عند استخدام التوزيع المعتاد كتقريب لتوزيع بواسون فإنه يجب التعامل مع المتغير المنفصل كمتغير متصل. ويترتب على ذلك إجراء التصحيح الملائم على ضوء ما ورد ذكره في حالة التوزيع ثنائي الحديْن.

مثال (٥٧):

إذا كانت نسبة الوحدات المعيبة في إنتاج إحدى الآلات هي ١٥٪، وأن عدد الوحدات المعيبة يتبع توزيع بواسون. احسب الاحتمالات التالية:

أ- الحصول على خمس وحدات معيبة من بين ٨٠ وحدة مختارة عشوائياً من إنتاج هذه الآلة.

ب- الحصول على ثلاث وحدات معيبة على الأكثر من بين ٥٠ وحدة تم اختيارها عشوائياً من إنتاج هذه الآلة.

مستخدماً في ذلك:

- توزيع بواسون - التوزيع المعتاد كتقريب لتوزيع بواسون

قارن بين نتائج التوزيعيْن في كل من أ، ب مع إبداء التعليق المناسب.

الحـــل:

أ- باستخدام توزيع بواسون:

حيث أن: $ن = ٨٠$ ، $ل = ٠,١٥$ ، $ر = ٥$.

$\lambda = نل = ٨٠(٠,١٥) = ١٢$

$$\therefore \text{ الاحتمال المطلوب } = \frac{هـ^{-١٢}(١٢)^{٥}}{٥!}$$

$$= \frac{١,٥٢٩}{١٢٠} = ٠,٠١٢٧$$

باستخدام التوزيع المعتاد:

حيث أن عدد الوحدات المعيبة هو متغير منفصل فإنه يمكن إيجاد الاحتمال المطلوب على النحو التالي:

$$ع(ر = ٥) = ع(س \leq ٥,٥) - ع(س \leq ٤,٥)$$

وحيث أن:

الوسط الحسابي = λ = ١٢

الانحراف المعياري = $\sqrt{\lambda}$ = $\sqrt{١٢}$ = ٣,٤٦٤

$$\therefore ع(س \leq ٥,٥) = ع(\frac{س - ١٢}{٣,٤٦٤} \leq \frac{٥,٥ - ١٢}{٣,٤٦٤})$$

$$= ع(ي \leq -١,٨٨) = \phi(-١,٨٨)$$

$$= ١ - \phi(١,٨٨) = ١ - ٠,٩٧٠٦ = ٠,٠٢٩٤$$

وكذلك: $ع(س \leq ٤,٥) = ع(\frac{س - ١٢}{٣,٤٦٤} \leq \frac{٤,٥ - ١٢}{٣,٤٦٤})$

$$= ع(ي \leq -٢,١٦٥) = \phi(-٢,١٦٥)$$

$= ١ - \phi(٢,١٦٥) = ١ - ٠,٩٨٤٨ = ٠,٠١٥٢$ تقريباً

$$\therefore ع(ر = ٥) = ع(س \leq ٥,٥) - ع(س \leq ٤,٥)$$

$$= ٠,٠٣٠١ - ٠,٠١٥٢ = ٠,٠١٤٩$$

وهذا يعنى أن الاحتمال المحسوب باستخدام التقريب زاد بمقدار بنسبة ١٩,٧٪ عن قيمة الاحتمال الحقيقي وهي ٠,٠١٢٧.

ب- باستخدام توزيع بواسون:

حيث أن: ن = ٥٠ ، ل = ٠,١٥ ، ر = ٠، ١، ٢، ٣

$$\lambda = ن ل = ٥٠(٠,١٥) = ٧,٥$$

$$\therefore \text{الاحتمال المطلوب} = ع(ر \leq ٣)$$

$$= ع(ر = صفر) + ع(ر = ١) + ع(ر = ٢) + ع(ر = ٣)$$

حيث يمكن إيجاد كل من هذه الاحتمالات على النحو التالي:

$$ع(ر = صفر) = \frac{هـ^{-٧,٥}(٧,٥)^{٠}}{٠!} = هـ^{-٧,٥} = ٠,٠٠٠٥٥$$

$$ع(ر = ١) = \frac{هـ^{-٧,٥}(٧,٥)}{١!} = ٠,٠٠٠٥٥(٧,٥) = ٠,٠٠٤١٣$$

$$ع(ر = ٢) = \frac{هـ^{-٧,٥}(٧,٥)^{٢}}{٢!}$$

$$= \frac{٠,٠٠٠٥٥(٥٦,٢٥)}{٢} = ٠,٠١٥٤٧$$

$$ع(ر = ٣) = \frac{هـ^{-٧,٥}(٧,٥)^{٣}}{٣!}$$

$$= \frac{٠,٠٠٠٥٥(٤٢١,٨٧٥)}{٦} = ٠,٠٣٨٦٧$$

$$\therefore\ ع(ر \le ٣) = ٠,٠٠٠٥٥ + ٠,٠٠٤١٣ + ٠,٠١٥٤٧ + ٠,٠٣٨٦٧$$
$$= ٠,٠٥٨٨$$

باستخدام التوزيع المعتاد:

حيث أن عدد الوحدات المعيبة هو متغير منفصل فإن إجراء التصحيح الملائم يضع الاحتمال المطلوب في الصورة التالية:

$$ع(ر \le ٣) = ع(س \le ٣,٥)$$

وحيث أن:

الوسط الحسابي = λ = ٧,٥

الانحراف المعياري = $\sqrt{\lambda} = \sqrt{٧,٥} = ٢,٧٤$

$\therefore$ الاحتمال المطلوب = $ع(س \le ٣,٥)$

$$= ع(\frac{س - ٧,٥}{٢,٧٤} \le \frac{٣,٥ - ٧,٥}{٢,٧٤})$$

$$= ع(ي \le -١,٤٦) = \phi(-١,٤٦)$$

$$= ١ - \phi(١,٤٦) = ١ - ٠,٩٢٧٩ = ٠,٠٧٢١$$

وهذا يعنى أن الاحتمال المحسوب باستخدام التقريب زاد بنسبة ٢٢,٦٪ عن القيمة الحقيقة للاحتمال.

وبمقارنة النتائج في كل من أ، ب نستطيع استنتاج أن التقريب في الحالة (أ) كان أكثر دقة وقرباً للواقع منه في الحالة (ب). ويرجع ذلك إلى أن قيمة λ في الحالة الأولى كانت أكبر منها في الحالة الثانية (١٢ > ٧,٥).

التوزيع المعتاد والاستدلال الإحصائي (العينات الصغيرة):

ذكرنا فيما سبق أن التحليل الإحصائي يقوم أساساً على أسلوب المعاينة، أي يقوم على تعميم نتائج العينات على المجتمعات التي سحبت منها هـذه العينات. وفي الواقع فإنه في ظل الأحجام الكبيرة للعينات ($ن \geq ٣٠$) يستطيع الباحث استخدام خصائص التوزيع المعتاد في عمليات الاستدلال الإحصائي (التقديرات واختبار الفرضيات) حتى لو لم يكن الانحراف المعياري للمجتمع معلوما لديه، أو حتى إذا كانت المجتمعات الأصلية التي سحبت منها العينات لا تتبع التوزيع المعتاد. ذلك أنه في حالة العينات الكبيرة فإن توزيعات الظواهر المختلفة تقترب - من الناحية النظرية - من التوزيع المعتاد. كما أن الانحراف المعياري للمجتمع (المجهول عادة) يمكن تقديره باستخدام بيانات العينة وذلك بدرجة ثقة كبيرة. ولكن المشكلة تنشأ أساساً في الحالات التي يكون فيها حجم العينة صغيراً ($ن < ٣٠$) والانحراف المعياري للمجتمع غير معروف. ففي مثل تلك الحالات يصبح الاعتماد على خصائص التوزيع المعتاد في عمليات التقدير واختبار الفرضيات أمراً مشكوكاً في صحته.

ولعل أهم المشكلات التي تواجه الباحث عند دراسته لأحد المجتمعات الإحصائية المعتادة أو القريبة من الاعتدال هي أن تباين المجتمع يكون مجهولاً وفي نفس الوقت لا يمكن تقديره بالدقة المطلوبة نظراً لصغر حجـم العينة المسحوبة من المجتمع. هذا وتوزيع **ت** والذي سنعرّفه فيما يلي يقدم لنا حلّاً لتلك المشكلة.

٢- توزيع ت: T-Distribution

في عام ١٩٠٨ قام وليام جوست William S. Gosset تحت اسم مستعار (ستودنت Student) بنشر بحث يعالج فيه المشكلة الإحصائية المشار اليها حيث قدم توزيعاً يقترب إلى درجة كبيرة من التوزيع المعتاد المعياري، وهذا التوزيع هو ما يعرف بتوزيع **ت** والذي عُرِف بعد ذلك باسم Student. وفيما يلي نورد بإيجاز خصائص هذا التوزيع:

أ- دالة كثافة الاحتمال للمتغير س الذي يتبع توزيع ت بدرجات حرية ن تأخذ الصورة التالية:

$$ف(س) = أ\left(١ + \frac{س^٢}{ن}\right)^{-\frac{(ن+١)}{٢}}$$

حيث: $- \infty \geq س \geq \infty$ ،

ن درجات الحرية ،

أ مقدار ثابت يمكن تحديد قيمته بالتكامل:

$$\int_{-\infty}^{\infty} ف(س)\, ءس = ١$$

وهنا يمكننا وببساطة تعريف درجات الحرية بأنها تشير إلى عدد القيم التي يكون لنا حرية تحديدها. فعلى سبيل المثال لو افترضنا أننا نتعامل مع عينة حجمها ٣ مفردات (ں = ٣) ونعلم أن متوسط هذه العينة هو ٨، فإن لنا الحرية في تحديد مفردتيْن فقط. فإذا كانت قيمة المفردة الأولى تساوي ٩ وقيمة المفردة الثانية تساوي ٥ فإن قيمة المفردة الثالثة يجب أن تساوي ١٠ وذلك لكي يكون المتوسط ٨. عندئذ تقول ان لدينا درجات حرية تساوي (ں - ١) أي (٣ - ١ = ٢). وبالمثل، إذا كانت ں = ١٨ فإن ذلك يعنى أن لدينا حرية تحديد ١٧ قيمة فقط منها.

ب- الوسط الحسابي (القيمة المتوقَّعة) والتباين لتوزيع ت يُعرَّفان كما يلي:

$$\textbf{توقع (س) = صفر،} \qquad (١ - ١٦)$$

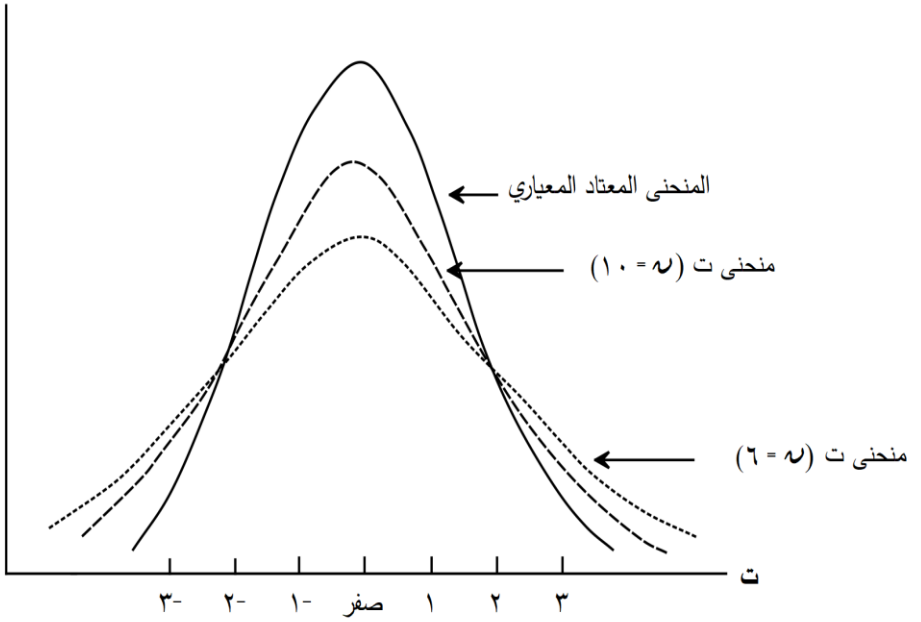
$$\textbf{تباين (س)} = \frac{ں}{ں - ٢} \qquad (١ - ١٧)$$

جـ- توزيع ت - مثله مثل التوزيع المعتاد المعياري - هو توزيع جرسي الشكل ومتماثل حول الصفر، ويقترب فيه المنحنى من المحور الأفقي تدريجياً ولكنه لا يلامسه أبداً.

د- يختلف توزيع ت عن التوزيع المعتاد المعياري في أن الأول يكون أكثر تشتتاً من الثاني، أي أن طرفي توزيع ت أكثر امتداداً على الجانبيْن ويزداد هذا التشتت كلما صغر حجم العينة (ں). حينئذ نجد أن منحنى توزيع ي (المنحنى المعتاد المعياري) يعلو منحنى توزيع **ت** كما أن طرفي المنحنى الثاني يعلوان طرفي المنحنى الأول عند الذيليْن {الشكل (١ - ٧)} ويزداد اقتراب توزيع ت من توزيع ي (التوزيع المعتاد المعياري) كلما زاد حجم العينة حتى يصل هذا الحجم إلى ٣٠ على الأقل (ں ≥ ٣٠). وعندما ں ← ∞ يتساوى التوزيعان من حيث التشتت (يلاحظ أن تباين التوزيع ت والمعرف بالصورة (١ - ١٦) يؤول إلى الواحد الصحيح عندما تؤول ں إلى ما لا نهاية). أي أن:

$$\underset{ں \leftarrow \infty}{نهـا}\left(\frac{ں}{ں - ٢}\right) = ١$$

هناك اختلاف آخر بين توزيع ت والتوزيع المعتاد المعياري وهو أن التوزيعات المعتادة مهما اختلفت أوساطها الحسابية أو انحرافاتها المعيارية يمكن تحويلها إلى توزيع معتاد معياري واحد وسطه الحسابي الصفر وتباينه الوحدة. وأما توزيع ت فهو عبارة عن مجموعة من المنحنيات يتوقف شكل كل منها على قيمة ں، وبزيادة قيمة ں يقترب منحنى توزيع ت من المنحنى المعتاد المعياري كما يظهر في الشكل (١ – ٧).



الشكل (١ – ٧)
مقارنة بين منحنى توزيع ت والمنحنى المعتاد المعياري

هـ- كما أشرنا من قبل فإنه في الوقت الذي نجد فيه أن هناك توزيعاً معتاداً معيارياً واحداً وبالتالي منحنى معتاداً معيارياً واحداً وذلك مهما اختلفت الأوساط الحسابية والانحرافات المعيارية للتوزيعات المعتادة، فإننا نجد أن توزيع ت ما هو إلا عبارة عن مجموعة من المنحنيات التي يتوقف شكل كل منها على حجم العينة (وبالتالي على درجات الحرية). لذلك فإنه للحصول على قيم (ت) المناظرة لدرجات حرية معينة واحتمالات محددة فإنه يمكننا استخدام جدول توزيع ت الموضح في الملحق (٢) في نهاية هذا الكتاب. حيث نجد أن لكل درجة حرية قيمة مناظرة للمتغير ت وذلك في إطار احتمالات معينة.

هذا علما بأن جدول (ت) يعطى الاحتمال ع(س ≤ $ت_{ن-١}$) حيث ن - ١ هي درجات الحرية. وعلى سبيل المثال، وعند استخدام جدول ت نجد أنه:

إذا كانت س تتبع توزيع ت بدرجات حرية ١٠ فإن:

ع(س ≤ ٢,٢٢٨) = ٠,٩٧٥ ، ع(س ≤ ٣,١٦٩) = ٠,٩٩٥

وهذا في حد ذاته يعنى ما يلي:

ع(س > ٢,٢٢٨) = ٠,٠٢٥ ، ع(س > ٣,١٦٩) = ٠,٠٠٥

و- يمكننا التحقق من أنه كلما زاد حجم العينة ن (وبالتالي درجات الحرية: ن - ١) كلما اقترب توزيع ت من التوزيع المعتاد المعياري حيث نلاحظ أنه بزيادة درجات الحرية إلى ∞ (والتي يمثلها الصف الأخير من جدول ت) نجد ما يلي على سبيل المثال:

ع(س ≤ ١,٩٦) = ٠,٩٧٥ ، ع(س ≤ ٢,٥٧٦) = ٠,٩٩٥

وذلك في كل من جدولي التوزيع المعتاد المعياري وتوزيع ت عند درجات حرية ∞.

وسوف نتوقف عند هذا الحد في تعريفنا لتوزيع ت وذلك بالقدر الذي يخدم حدود استخداماته في عمليات الاستدلال الإحصائي في ظل الأحجام الصغيرة للعينات على نحو ما سنرى في الفصل التالي.

تماريـــن
الفصل الأول

١- صندوق يحتوي على ٤ كرات بيضاء و٨ كرات سوداء، فإذا سحبت كرتان بطريقة عشوائية من الصندوق (بفرض عدم الإعادة)، احسب الاحتمالات التالية:

أ- أن تكون كلٌ من الكرتيْن المسحوبتيْن بيضاء.

ب- أن تكون كلٌ من الكرتيْن سوداء.

ج- أن تكون إحدى الكرتيْن بيضاء والأخرى سوداء.

٢- إذا أُلقيت زهرتا نرد على سطح أملس، احسب الاحتمالات التالية:

أ- الحصول على رقم ٧ على الزهرة الأولى.

ب- الحصول على رقميْن مجموعهما ٩.

٣- إذا كان احتمال نجاح الطالب في مادة الإحصاء هو ٠,٨٥ واحتمال نجاحه في مادة الإدارة المالية هو ٠,٧ فإذا تم اختيار أحد الطلاب بطريقة عشوائية، احسب الاحتمالات التالية:

أ- رسوبه في المادتيْن.

ب- نجاحه في المادتيْن.

ج- نجاحه في الإدارة المالية ورسوبه في الإحصاء.

د- نجاحه في الإحصاء أو الإدارة المالية.

ه- رسوبه في مادة واحدة على الأقل.

٤- صندوق يحتوي على ٤ كرات بيضاء و٦ كرات سوداء، وصندوق أخر به ٣ كرات بيضاء و٥ سوداء، تم اختيار أحد الصندوقيْن عشوائياً ثم سحبت منه كرة، المطلوب:

أ- احتمال أن تكون الكرة المسحوبة بيضاء.

ب- إذا كانت الكرة المسحوبة سوداء. ما هو احتمال أن تكون قد سحبت من الصندوق الثاني.

٥- سحبت ورقة بطريقة عشوائية من مجموعة أوراق اللعب.

المطلوب:

أ- احتمال أن تكون الورقة المسحوبة صورة.

ب- إذا كانت الورقة المسحوبة من اللون الأسود، ما هو احتمال أن تكون ثمانية؟

ج- احتمال أن تكون الورقة المسحوبة سبعة أو ورقة من نوع السباتى.

د- إذا كانت الورقة المسحوبة صورة، ما هو احتمال أن تكون بنتاً؟

٦- في التمرين (٢)، احسب احتمال الحصول على رقميْن مجموعهما ٦ على أن يكون الرقم على الزهرة الثانية ٤.

٧- أظهرت التجارب السابقة أنه من بين كل ١٠٠٠ وحدة منتَجة بواسطة الآلة (أ) في أحد المصانع هناك ٥٠ وحدة معيبة، وأنه من بين كل ٢٠٠٠ وحدة منتَجة بواسطة الآلة (ب) في نفس المصنع هناك ٢٠٠ وحدة معيبة. فإذا كان الانتاج اليومي للمصنع يتمثل في ٩٠٠ وحدة من إنتاج الآلة (أ) و٦٠٠ وحدة من إنتاج الآلة (ب). وإذا سحبت وحدة من الإنتاج اليومي للمصنع ومقداره ١٥٠٠ وحدة، احسب الاحتمالات التالية:

أولاً: أن تكون الوحدة المسحوبة من إنتاج الآلة (أ).

ثانياً: أن تكون معيبة ومن إنتاج الآلة (ب).

ثالثاً: أن تكون معيبة.

رابعاً: إذا كانت الوحدة المسحوبة سليمة، ما هو احتمال أن تكون من إنتاج الآلة (ب)؟

٨- أوجد التوزيع الاحتمالي والتوزيع التجميعي للمتغيرات العشوائية المعبر عنها كما يلي:

أ- عدد مرات ظهور الوجه الذي عليه الكتابة عند رمي قطعتيْ عملة.

ب- مجموع الرقميْن عند رمي زهرتيْ نرد.

٩- إذا كانت س متغير عشوائي يمكن أن يأخذ أي قيمة داخل المدى من ٥ إلى ١٠ وذلك باحتمالات متساوية، المطلوب:

أ- إيجاد دالة كثافة الاحتمال.

ب- إيجاد دالة التوزيع الاحتمالي التجميعي.

ج- حساب الاحتمالات: ع(س ≤ ٧)، ع(٦ ≤ س ≤ ٩).

١٠- احسب القيمة المتوقَّعة والتباين لعدد مرات ظهور الوجه الذي عليه الشعار (الصورة) وذلك عند رمي قطعتي عملة.

١١- فيما يلي التوزيع الاحتمالي للمتغير العشوائي (س):

س	صفر	١	٢	٣
ع(س)	$\frac{١}{٨}$	$\frac{١}{٤}$	$\frac{١}{٢}$	$\frac{١}{٨}$

المطلوب:

أ- احسب التوقُّع والتباين للمتغير س.

ب- احسب: ت(س - ١) ، ت(س + ٤) ، ت($س^{٢}$)

ج- الانحراف المتوسط للمتغير س.

١٢- إذا كانت دالة كثافة الاحتمال للمتغير العشوائي (س) هي:

$$ف(س) = \begin{cases} \frac{٣}{٤}(٢س - س^{٢}) & \text{صفر} \le س \le ٢ \\ \text{صفر} & \text{خلاف ذلك} \end{cases}$$

المطلــوب:

أ- أوجد القيمة المتوسطة والوسيط والمنوال لهذا المتغير، ثم علق على ما تحصل عليه من نتائج.

ب- احسب التباين للمتغير العشوائي س.

ج- احسب الاحتمالات التالية:

$ع(س \le ١)$ ، $ع(س > ١,٥)$ ، $ع(س \le ٠,٨)$

١٣- إذا كانت دالة إحدى الظواهر تتبع التوزيع التالي:

$$ف(س) = \begin{cases} \frac{١}{٢} هـ^{-\frac{س}{٢}} & ٠ \le س \le \infty \\ \text{صفر} & \text{خلاف ذلك} \end{cases}$$

المطلوب:

أ- تحقق من أن هذه دالة تمثل دالة كثافة احتمال.

ب- أوجد دالة التوزيع الاحتمالي.

ج- احسب الوسط الحسابي والتباين.

د- احسب الاحتمالات التالية:

$ع(س \le ٣)$، $ع(س > ١٠)$

١٤- أُلقيت زهرة نرد خمس مرات، احسب الاحتمالات التالية:

أ- الحصول على الرقم ٤ ثلاث مرات.

ب- الحصول على رقم زوجي مرة واحدة على الأقل.

ج- عدم الحصول على رقم ٥.

١٥- أوجد دالة كثافة الاحتمال للمتغير العشوائي (س) والذي يمثل عدد الوحدات المعيبة في عينة مكونة من ١٠ وحدات أُخذت من إنتاج آلة نسبة المعيب في إنتاجها تبلغ ٥٪.

١٦- ألقيت قطعة عملة أربع مرات، المطلوب:

أ- أوجد دالة كثافة الاحتمال للمتغير العشوائي س والذي يمثل ظهور الوجه الذي عليه الشعار.

ب- احسب احتمال ظهور الوجه الذي عليه الشعار مرة واحدة على الأقل.

ج- أوجد القيمة المتوقَّعة والتباين للمتغير س.

١٧- إذا كان ١٠٪ من الطلاب الذين يلتحقون بجامعة ما يتركونها قبل إتمام الدراسة بها. أوجد احتمال أنه من بين عشرة طـــلاب تـــم اختيارهم عشوائياً من بين طلاب هذه الجامعة سوف يكون هناك طالبان فقط يتركون الجامعة قبل إتمام دراستهم بها.

١٨- إذا علمت أن نسبة الوحدات المطابقة للمواصفات من إنتاج أحد المصانع هي ٩٠٪. فإذا كان قسم مراقبة الإنتاج في المصنع يقوم بفحص ٨ وحدات يومياً، احسب احتمال أن يكون من بينها:

أ- خمس وحدات مقبولة فنياً.

ب- وحدتان على الأكثر غير مطابقة للمواصفات.

ج- ما هو العدد المتوقَّع لعدد الوحدات المقبولة فنياً.

د- احسب الانحراف المعياري لعدد الوحدات المقبولة فنياً.

ه- هل التوزيع الاحتمالي للوحدات المقبولة فنياً متماثل؟ لماذا؟

١٩- إذا كان احتمال كشف جهاز رادار لطائرة معادية هو ٠,٨ فإذا كانت إحدى القواعد تمتلك ٥ أجهزة من هذا النوع، احسب الاحتمالات التالية:

أ- ظهور الطائرة المعادية على أربعة أجهزة.

ب- ظهور الطائرة المعادية على جهازين على الأقل.

ج- كشف الطائرة المعادية.

٢٠- إذا كانت نسبة المعيب في إنتاج إحدى الآلات هي ١٠٪ ، وأُخذت عينة مكونة من خمس وحدات من إنتاج هذه الآلة، احسب الاحتمالات التالية:

أ- ألا يكون بالعينة أية وحدات معيبة.

ب- أن يكون بالعينة وحدتين معيبتيْن.

ج- ألا يزيد عدد الوحدات السليمة عن ٤ وحدات.

د- احسب العدد المتوقَّع للوحدات المعيبة.

٢١- تشير سجلات أحد أقسام البوليس بأن هناك ثلاث حوادث في المتوسط عند تقاطع ما شهرياً، احسب الاحتمالات التالية:

أ- وقوع خمس حوادث خلال شهريْن عشوائياً.

ب- عدم وقوع حوادث خلال شهر كامل.

ج- القيمة المتوقَّعة والانحراف المعياري لعدد الحوادث خلال شهريْن.

٢٢- إذا أظهرت الخبرة الماضية أن ١٪ من الأطفال يصابون بمرض معين خلال عام. فإذا اخترنا ١٠٠ طفل عشوائياً، المطلوب:

أ- احتمال أن هناك ٥ أطفال سوف يصابون بهذا المرض خلال العام.

ب- ما هي القيمة المتوقَّعة لعدد الأطفال الذين يصابون بالمرض خلال العام.

٢٣- تنتج إحدى الشركات نوعاً معيناً من الآلات الحاسبة الصغيرة. فإذا أوضحت دراسات الشركة أن ٢٪ من الآلات المُباعة تُرد وتُستبدل خلال فترة الضمان. وأن إحدى المؤسسات قامت بشراء ٢٠٠ آلة حاسبة من إنتاج هذه الشركة. احسب احتمال أنه خلال فترة الضمان:

أ- أن آلة واحدة على الأقل سوف يتم استبدالها.

ب- ألا يتم استبدال أية آلة.

ج- أن يتم استبدال خمس آلات.

د- احسب توقُّع وتباين عدد الآلات الحاسبة التي يتم استبدالها.

٢٤- من سجلات إحدى محطات تزويد السيارات بالوقود وُجد أنه تتوقف عشر سيارات في المتوسط خلال نصف ساعة وذلك للتزود بالوقود.

أ- احتمال توقف ٣ سيارات للتزويد بالوقود خلال نصف ساعة.

ب- توقف سيارة واحدة على الأقل للتزود بالوقود خلال ساعة.

ج- العدد المتوقَّع للسيارات التي تتوقف للتزود بالوقود خلال ثلاث ساعات.

٢٥- إذا كانت سجلات قسم الصيانة في أحد المصانع تشير إلى أنه خلال الخمسين أسبوعاً الماضية حدثت خمسة أعطال لنوع معين من الآلات. فإذا كان التشغيل الناجح للمصنع يتطلب وجود ثلاث آلات على الأقل غير معطلة، كما أن المصنع يمتلك خمس آلات من هذا النوع. احسب ما يلي:

أ- احتمال التشغيل الناجح ليوم معين.

ب- احتمال ألا يحدث أكثر من عطل واحد خلال شهر قادم (الشهر = ٣٠ يوماً)

٢٦- توضح الخبرة الماضية أن ١٪ من مصابيح الكهرباء المنتَجة في أحد المصانع هي مصابيح معيبة. فإذا تم سحب عينة مكونة من ٣٠ مصباحاً، احسب الاحتمالات التالية:

أ- وجود ٣ مصابيح معيبة.

ب- وجود أكثر من مصباح معيب.

وذلك باستخدام:

أولاً: التوزيع ثنائي الحديْن.

ثانياً: توزيع بواسون كتقريب لتوزيع ثنائي الحديْن.

ثم قارن بين نتائجك في الحالتيْن مبدياً ما تراه مناسباً من تعليق.

٢٧- المطلوب حل التمرين السابق بافتراض أن حجم العينة ١٠٠مصباح. قارن بين نتائجك في هذه الحالة وبين النتائج التي حصلت عليها من قبل مبدياً ما تراه مناسباً من تعليق.

٢٨- وضح أنه عندما تكون ن = ١٠٠، ل = ٠,٠٢ فإن توزيع بواسون يعطى تقريباً جيداً للتوزيع ثنائي الحديْن.

٢٩- إذا كانت نسبة المعيب في إنتاج شركة ما هي ٠,٥٪ وكانت منتَجات هذه الشركة معبأة في صناديق بكل منها ١٠٠ وحدة. فإذا كانت الشركة تضمن إرجاع الصندوق الذي يحتوي على أكثر من وحدتيْن معيبتيْن، وأن إرجاع مثل هذا الصندوق يكلف الشركة جنيهان. فإذا فكرت الشركة في خطة فحص لمنتجاتها مع استبعاد المعيب منها، وكانت عملية الفحص تكلف الشركة خمسة عشر قرشاً لكل صندوق.

المطلوب:

استخدام توزيع بواسون كتقريب للتوزيع ثنائي الحديْن في توضيح ما إذا كانت خطة الفحص تستحق الاهتمام أم لا.

٣٠- صندوق به ٧ كرات بيضاء و٣ كرات خضراء. سُحبت منه ثلاث كرات بطريقة عشوائية، احسب الاحتمالات التالية:

أولاً: الحصول على كرة واحدة بيضاء من بين الكرات المسحوبة.

ثانياً: الحصول على كرة خضراء و٣ كرات بيضاء.

وذلك بافتراض:

أ- السحب تم مع الإعادة ب- السحب تم مع عدم الإعادة

٣١- مجتمع مكون من عشرة طلاب منهم ٦ ذكور، ٤ إناث، فإذا اختيرت مجموعة مكونة من خمسة طلاب بطريقة عشوائية من مجتمع الطلاب وبدون إعادة، احسب:

أ- احتمال أن تكون المجموعة المختارة مكونة من:

- ٣ ذكور و٢ إناث - ذكر واحد وأربع إناث.

ب- عدد الإناث المتوقَّع في المجموعة المختارة.

٣٢- إذا كانت ي متغير معتاد معياري، احسب - باستخدام الجداول - الاحتمالات التالية:

أ - ع(ي > ١,٥) **ب-** ع(ي ≥ ٢,٤) **جـ-** ع(٠,٨ ≥ ي ≥ ٢,٧)

د- ع(ي > -١,٦) **هـ -** ع(-٣ ≥ ي ≥ ١,٨)

٣٣- إذا كان من المعروف أن توزيع أعمار المصابيح الكهربائية المنتجة بواسطة أحد المصانع يتبع معتادا بمتوسط ١٠٠ ساعة وانحراف معياري ٨ ساعات. فإذا تم اختيار أحد المصابيح الكهربائية عشوائياً، احسب الاحتمالات التالية:

أ- أن يقل عمر المصباح عن ٩٠ ساعة.

ب- أن يزيد عمر المصباح عن ١٢٠ ساعة.

ج- أن يقل عمر المصباح عن ١١٠ ساعة.

د- أن يتراوح عمر المصباح ما بين ١٠٨ و١٠٥ ساعة.

هـ- أن يتراوح عمر المصباح ما بين ١٠٨ و ١١٢ ساعة.
و- أن يتراوح عمر المصباح ما بين ٨٠ و ٩٠ ساعة.

٣٤- إذا كانت أوزان مجموعة كبيرة من الطلاب تتبع توزيعاً معتاداً متوسطه ٧٥ كجم وانحرافه المعياري ٥ كجم. فإذا تم اختيار أحد الطلاب بطريقة عشوائية من بين مجموعة الطلاب، احسب ما يلي:
أ- احتمال أن يقل وزن الطالب عن ٨٠ كجم.
ب- احتمال أن يتراوح وزن الطالب ما بين ٦٥ و ٨٥ كجم.
ج- ما هو الوزن الذي يزيد عليه ٦٥٪ من الطلاب.
د- ما هو الوزن الذي يقل عنه أو يساويه ٨٥٪ من الطلاب.

٣٥- إذا كانت الدرجات التي حصل عليها طلاب الفرقة الثانية في إحدى كليات التجارة في مادة ما تتبع توزيعاً معتاداً وسطه الحسابي ٧٠ درجة وانحرافه المعياري ١٠ درجات.
احسب الدرجة التي حصل على أكبر منها ٢٥٪ من إجمالي عدد الطلاب.

٣٦- إذا كان العمر الافتراضي لنوع معين من إطارات السيارات هو ٢٥٠٠٠ كم والانحراف المعياري له هو ١٠٠٠ كم. فإذا افترضنا أن العمر الافتراضي للإطارات يتبع التوزيع المعتاد، احسب ما يلي:
أ- النسبة المئوية المتوقَّعة للإطارات التي يتراوح عمرها الافتراضي ما بين ٢٤٠٠٠، ٢٧٠٠٠ كم.
ب- العدد المتوقَّع للإطارات التي يزيد عمرها الافتراضي عن ٢٦٠٠٠كم.
ج- العمر الافتراضي الذي تبلغه - كحد أقصى - ٧٥٪ من الإطارات.

٣٧- إذا كان عدد الطلاب الذين يرغبون في الالتحاق بكلية الشرطة هو ١٠ آلاف طالب. وكانت أطوالهم تتبع توزيعاً معتاداً متوسطه ١٧٠ سم وانحرافه المعياري ١٥ سم. فإذا علمت أن كلية الشرطة تقبل الأشخاص الذين تتراوح أطوالهم ما بين ١٤٥، ١٧٥ سم (بافتراض اجتيازهم للاختبارات المختلفة)، ما هو عدد الطلاب المتوقَّع قبولهم؟

٣٨- لو افتراضنا بأن الخبرة الماضية قد أشارت بأن ٦٥٪ فقط من الطلاب الملتحقين بالفرقة الأولى بكلية التجارة بإحدى الجامعات يحصلون على درجة البكالوريوس. فإذا اختير وبطريقة عشوائية ٣٥ طالباً من الملتحقين حديثا بكلية التجارة، احسب احتمال أن يحصل منهم ٢٨ طالباً على الأقل على درجة البكالوريوس وذلك باستخدام:
أ- التوزيع ثنائي الحديْن.
ب- التوزيع المعتاد كتقريب للتوزيع ثنائي الحديْن.
قارن بين نتائجك: أ و ب مبدياً ما تراه مناسباً من تعليق

٣٩- إذا كانت نسبة الوحدات المعيبة في إنتاج أحد المصانع هي ٢٠٪، وأن عدد الوحدات المعيبة يتبع توزيع بواسون. احسب احتمال الحصول على وحدتيْن معيبتيْن من بين ٥٠ وحدة مختارة عشوائياً من إنتاج هذا المصنع، مستخدماً في ذلك.

أ- توزيع بواسون.

ب- التوزيع المعتاد كتقريب لتوزيع بواسون.

قارن بين نتائجك في أ و ب، مع إبداء التعليق المناسب.

٤٠- إذا كان المتغير س يتبع توزيع ت بدرجات حرية ٨ درجات. احسب الاحتمالات التالية: مستخدماً في ذلك جدول المساحة تحت المنحنى الاحتمالي لتوزيع ت:

أ- ع(س ≤ ٢,٨٩٦)　　　**ب-** ع(س ≤ ٢,٣٠٦)

جـ- ع(س > ١,٣٩٧)

الفصل الثاني

التقدير الإحصائي

مقدمـــــة:

سبق أن أشرنا في الفصل الأول من هذا الكتاب إلى أن الأساليب الإحصائية تعتمد أساساً في تطبيقاتها على أسلوب المعاينة. إذ أن أسلوب الحصر الشامل تعترضه العديد من الصعوبات فنيةً كانت أم ماليةً وحيث لا يمكن استخدامه في العديد من الحالات. ونظراً لما يتميز به أسلوب المعاينة من دقة في البيانات، بالإضافة إلى كونه الأسرع والأقل تكلفة، فقد أصبح هذا الأسلوب على جانب كبير من الأهمية في عمليات التحليل الإحصائي. وعلى سبيل المثال يمكن الاستدلال على طبيعة الجوانب المختلفة للمجتمع وذلك باستخدام البيانات المتاحة عن عينة مسحوبة من هذا المجتمع. فباتباع أساليب الاستدلال الإحصائي يمكننا الحصول على تقديرات لمعلمات المجتمع (الوسط الحسابي، النسبة، التباين،... الخ) وذلك عن طريق المعلومات المتاحة عن هذا المجتمع من واقع بيانات العينة المأخوذة منه.

وحيث أن هناك عدداً كبيراً من العينات التي يمكن اختيارها من المجتمع فإن تقدير معلمة من معلمات المجتمع باستخدام بيانات العينة يختلف من حيث الدقة من عينة إلى أخرى. ولعل المثال التالي يوضح ذلك:

مثال (١):

لنفترض أن لدينا مجتمعاً مكوناً من ستة طلاب، وسوف نرمز إليهم على الترتيب بالرمز أ، ب، جـ، ء، هـ، و. وكانت درجاتهم في مـادة الاقتصاد هي على الترتيب: ١٨، ١٤، ١٠، ١٥، ١٧، ١٦. فإذا اردنـا تقدير متوسط درجات الطلاب في المجتمع وذلك باستخدام بيانات عينة عشوائية بسيطة مكونة من أربعة طلاب ومسحوبة من هذا المجتمع، ما هي النتائج التي يمكن الحصول عليها؟

الحـــل:

في هذا المثال نجد أن عدد العينات المختلفة ذات الحجم ٤ والتي يمكن سحبها من المجتمع هو ($^{٦}ق_{٤}$ = ١٥ عينة). كما أن المتوسط الحقيقي للمجتمع يمكن حسابه كما يلي:

$$\text{المتوسط الحقيقي للمجتمع } (\text{م}) = \frac{١٨ + ١٤ + ١٠ + ١٥ + ١٧ + ١٦}{٦}$$

= ١٥ درجة

وباستعراض جميع العينات وحساب الوسط الحسابي لكل منها وذلك كما يوضحه الجدول (٢ - ١) يمكننا أن نستنتج أن هناك عينة واحدة فقط (العينة رقم ٧) يتساوى متوسطها مع المتوسط الحقيقي للمجتمع ($\bar{س}$ = م). وأما بقية العينات فإن موسطاتها تختلف زيادة أو نقصاناً عن المتوسط الحسابي الحقيقي للمجتمع، مع ملاحظة أن مجموع انحرافات المتوسطات المختلفة للعينات عن المتوسط الحقيقي يساوي الصفر. وهذا يعنى أن احتمال سحب عينة حجمها ٤ يتساوى متوسطها مع المتوسط الحقيقي للمجتمع هو ($\frac{١}{١٥}$ = ٠,٠٧). ويمكننا أن نتصور إذا ما كان حجم المجتمع كبيراً فإن عدد العينات ذات الحجم الواحد والتي يمكن سحبها من المجتمع سوف يكون كبيراً جداً، وبالتالي يقل احتمال سحب عينة عشوائية تعكس الواقع الحقيقي للمجتمع تماماً. ويكون ذلك بصف خاصة إذا كان حجم العينة صغيراً بدرجة كافية بالنسبة إلى حجم المجتمع المسحوبة منه. فمن المعروف إحصائياً أنه كلما زاد حجم العينة كلما زادت درجة دقة التقديرات.

وهنا فإنه من المنطقي أن يُثار التساؤل التالي: إذا كان المقياس الإحصائي المحسوب يختلف من عينة إلى أخرى وذلك نتيجة لاختلاف المفردات الممثَّلة في العينة. وإذا كنا لا نعرف القيمة الحقيقية للمقياس الإحصائي في المجتمع. فما هو المعيار الذي يمكن من خلاله الحكم على درجة الدقة في تقديرات عينة بذاتها؟

الجدول (٢ - ١)
بيانات جميع العينات ذات الحجم ٤ والممكن سحبها من مجتمع حجمه ٦ مفردات

رقم العينة	مفردات العينة	بيانات العينة	الوسط الحسابي للعينة ($\bar{س}$)	انحراف الوسط الحسابي للعينة عن الوسط الحسابي للمجتمع ($\bar{س}$ - م)
١	أ ب ج د	١٨ ١٤ ١٠ ١٥	١٤,٢٥	-٠,٧٥
٢	أ ب ج هـ	١٨ ١٤ ١٠ ١٧	١٤,٧٥	-٠,٢٥
٣	أ ب ج و	١٨ ١٤ ١٠ ١٦	١٤,٥٠	-٠,٥٠
٤	أ ب د هـ	١٨ ١٤ ١٥ ١٧	١٦,٠٠	١,٠٠
٥	أ ب د و	١٨ ١٤ ١٥ ١٦	١٥,٧٥	٠,٧٥
٦	أ ب هـ و	١٨ ١٤ ١٧ ١٦	١٦,٢٥	١,٢٥
٧	أ ج د هـ	١٨ ١٠ ١٥ ١٧	١٥,٠٠	صفر
٨	أ ج د و	١٨ ١٠ ١٥ ١٦	١٤,٧٥	-٠,٢٥
٩	أ ج هـ و	١٨ ١٠ ١٧ ١٦	١٥,٢٥	٠,٢٥
١٠	أ د هـ و	١٨ ١٥ ١٧ ١٦	١٦,٥٠	١,٥٠
١١	ب ج د هـ	١٤ ١٠ ١٥ ١٧	١٤,٠٠	-١,٠٠
١٢	ب ج د و	١٤ ١٠ ١٥ ١٦	١٣,٧٥	-١,٢٥
١٣	ب ج هـ و	١٤ ١٠ ١٧ ١٦	١٤,٢٥	-٠,٧٥
١٤	ب د هـ و	١٤ ١٥ ١٧ ١٦	١٥,٥٠	٠,٥٠
١٥	ب د هـ و	١٠ ١٥ ١٧ ١٦	١٤,٥٠	-٠,٥٠

وفي واقع الأمر فإن الاستدلال الإحصائي يقدم لنا الإجابة على هذا التساؤل وذلك من خلال استخدامه لنظرية الاحتمالات والنظرية الإحصائية الحديثة. وينقسم الاستدلال الإحصائي إلى نوعيْن هما:

١- التقدير الإحصائي: وهو موضوع هذا الفصل.

٢- اختبـــار الفــــرضيات: والذي يتم تناوله في الفصل الثالث من هـذا الكتاب.

ولعله من المناسب هنا ان نتناول بعض التعريفات والمفاهيم المستخدمة في التقدير الإحصائي عموماً وذلك كما يلي:

بعض التعريفات والمفاهيم في التقدير الإحصائي:

١- التقدير بنقطة: Point Estimation

وهذا التقدير يُعبَّر عنه بعدد واحد، مثال ذلك أن يقدر متوسط دخـل الأسرة في مجتمع ما بـ ٦٥٠ جنيهاً مثلا وذلك من خلال سحب عينة من هذا المجتمع. وبصفة عامة يمكننا القول إن متوسط العينة ($\bar{س}$) هو تقدير لمتوسط المجتمع (م)، أو أن نسبة وجود ظاهرة ما في العينة ($\hat{ل}$) هي تقدير لنسبة الظاهرة في المجتمع (ل). وقس على ذلك بالنسبة لبقية المعلمات المراد تقديرها.

وفي مثـال (١) يمكننا ملاحظ أن المتوسط الحقيقي للمجتمـــع (م) ما هو إلا عبارة عن الوسط الحسابي لمتوسطات العينات المختلفة ($\bar{س}_{١}$، $\bar{س}_{٢}$ ....، $\bar{س}_{١٥}$). وهنا يمكننا القول إن (م) تساوي توقع ($\bar{س}$)، أي أن الوسط الحسابي للعينة ($\bar{س}$) هو تقدير غير متحيز لمتوسط المجتمع (م).

ملحوظة:

يُقال لمقدّر ما أنه غير متحيز في تقديره لمعلمة ما من معلمات المجتمع إذا كانت القيمة المتوقَّعة لهذا المقدّر (أي متوسطه) تساوي قيمة المعلمة الحقيقية في المجتمع.

وحيث أن تقدير الوسط الحسابي للمجتمع - على سبيل المثال - بقيمة واحدة محددة يختلف من حيث الدقة من عينة إلى أخرى (يلاحَظ أن متوسطات العينات في المثال المشار إليه تتراوح ما بين ١٣,٧٥، ١٦,٢٥ درجة في حين أن الوسط الحقيقي للمجتمع يبلغ ١٥ درجة)، فإنه يمكننا القول بأن الاعتماد على قيمة واحدة لتقدير لمعلمة المجتمع قد يحمل في طياته قدراً كبيراً من الريبة والشك حول صحة النتائج المُتوصَّل إليها. ولعل الأمر يكون أكثر دقة لو أننا عرفنا حجم الخطأ الذي يحتمل الوقوع فيه نتيجة لاتباعنا أسلوب المعاينة. وهذا الأمر تتناوله الصورة الثانية من صورتي التقدير وهي التقدير بفترة.

٢- التقدير بفترة: Interval Estimation

يشير التقدير بفترة إلى المدى أو الفترة التي تقع معلمة المجتمـــع غـير المعلومة داخلها وذلك بدرجة ثـقة معينة.

مثال:

يمكننا أن نقول إن متوسط دخل الأسرة في المجتمع المسحوب منه العينة يتراوح ما بين ٧٥٠، ٨٥٠ جنيهاً وذلك بدرجة ثـقة ٩٥٪.

وفي هذا الشأن لعله من المفيد أن نعرض بعض المفاهيم والتعريفات مستعينين في ذلك بالمثال المذكور:

أ- فترة الثـقة: Confidence Interval

وهي المدى الذي يُتوقع أن تقع فيه القيمة الحقيقية لمعلمة المجتمع. وتعكس فترة الثـقة درجة الدقة في التقدير وكذلك حجم الخطأ الذي يمكننا الوقوع فيه. ويتوقف اختيار فترة الثـقة على مدى إلمام الإحصائي بطبيعة المشكلة موضوع البحث.

وفي مثالنا هذا تتحدد فترة الثـقة بالعدديْن (٧٥٠، ٨٥٠)، أي أن المتوسط الحقيقي لدخل الأسرة في المجتمع يمكن أن يأخذ آية قيمة ابتداءً من ٧٥٠ جنيهاً وحتى ٨٥٠ جنيهاً.

ب- حدا الثـقة: Confidence Limits

وهما الحد الأدنى والحد الأعلى للثـقة، حيث يمثل الحد الأدنى أقل قيمة تشملها فترة الثـقة بينما يمثل الحد الأعلى أكبر قيمة تشملها هذه الفترة.

وفي مثالنا الحالي نجد أن:

الحد الأدنـى للثـقة = ٧٥٠ جنيهاً ، الحد الأعلى للثـقة = ٨٥٠ جنيهاً.

جـ- معامل الثـقة: Confidence Coefficient

يشير معامل الثـقة إلى احتمال وقوع القيمة الحقيقية لمعلمة المجتمع داخل فترة الثـقة. أي أنه يعبر عن درجة الثـقة في صحة التقدير الذي حصلنا عليه. وبتعبير آخر: لو أخذنا عدداً كبيراً من العينات من المجتمع المراد دراسته وحددنا فترة ثـقة في كل مرة نختار فيها عينة، فإنه من المتوقع أن نجد انه في ٩٥٪ من الحالات (هذا بافتراض أن معامل الثـقة = ٩٥٪) سوف تقع القيمة الحقيقية لمعلمة المجتمع داخل فترة الثـقة المحسوبة. ومعامل الثـقة هذا يحدده الباحث قبل أن يقوم بحساب فترة الثـقة، وعادة ما تتراوح قيمة ما بين ٩٠٪، ١٠٠٪. ويُرمز إلى معامل الثـقة بالرمز: $(١ - \alpha)$.

وتبلغ قيمة الثـقة في مثالنا ٠,٩٥ وعليه نستطيع القول إن المتوسط الحقيقي لدخل الأسرة في المجتمع (م) يتراوح ما بين ٧٥٠ جنيهاً، ٨٥٠ جنيهاً وذلك بمعامل ثـقة ٠,٩٥ وهذا بدوره يعنى أن:

ع(م > ٧٥٠) = ع(م < ٨٥٠) = ٠,٠٢٥.

د- درجة الثـقة: Confidence Level

وهي عبارة عن معامل الثـقة معبراً عنه في شكل نسبة مئوية، أي أن:

درجة الثـقة = (١ – α) × ١٠٠

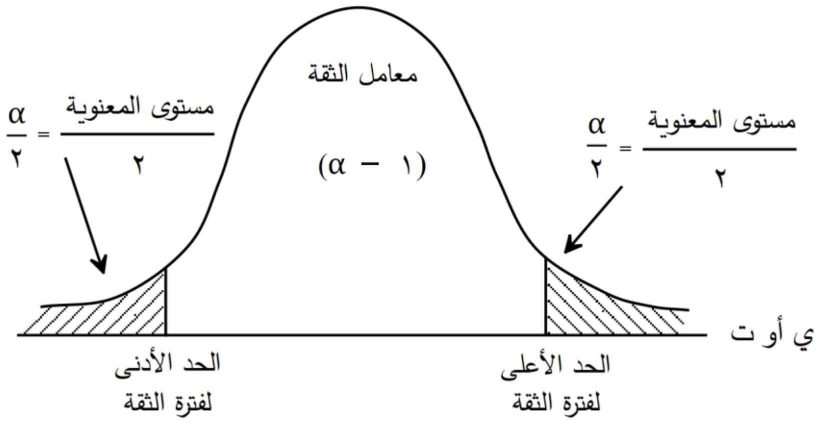
وفي مثالنا الحالي نجد أن:

درجة الثـقة = ٠,٩٥ × ١٠٠ = ٩٥٪.

هـ- مستوى المعنوية: Level of Significance

أشرنا من قبل إلى أن التقدير بفترة يأخذ في الاعتبار احتمالات الوقوع في خطأ عند التقدير. إذ أن القيمة الحقيقية لمعلمة المجتمع هي أساساً قيمة مجهولة وبالتالي فهناك احتمال أو فرصة لأن يخطئ الباحث في تقديرها.

واحتمال الخطأ هذا يُعبَّر عنه بمستوى المعنوية، أي أن مستوى المعنوية هو احتمال عدم وقوع القيمة الحقيقية لمعلمة المجتمع داخل فترة الثـقة. وبمعنى آخر فإن مستوى المعنوية يشير إلى احتمال أن تكون القيمة الحقيقية لمعلمة المجتمع أقل من الحد الأدنى أو أكبر من الحد الأعلى لفترة الثـقة. وبتعبير آخر: لو اخترنا عدداً كبيراً من العينات وحددنا فترة ثـقة في كل مرة نختار فيها عينة، فإنه من المتوقَّع أن نجد أنه في ٥٪ من الحالات (بافتراض أن مستوى المعنوية = ٥٪) سوف تقع القيمة الحقيقية لمعلمة المجتمع خارج فترة الثـقة ويرمز إلى مستوى المعنوية بالرمز (α).



الشكل (٢ - ١)

وفي مثالنا الحالي حيث أن مستوى المعنوية يساوي ٠,٠٥ فإننا نستطيع القول إن احتمال أن يقل المتوسط الحقيقي لدخل الأسرة في المجتمع عن ٧٥٠ جنيهاً أو يزيد على ٨٥٠ جنيهاً يساوي ٠,٠٢٥. أي أن:

$$ع(م < ٧٥٠) = ع(م > ٨٥٠) = \frac{\alpha}{٢} = ٠,٠٢٥.$$

هذا ونستطيع التعبير - بيانياً - عن المفاهيم والتعريفات التي أشرنا إليها وذلك كما هو موضح في الشكل (٢ - ١).

وفي الأجزاء التالية من هذا الفصل سوف نقوم بدراسة بعض جوانب موضوع التقدير بفترة ثقة وذلك لبعض معلمات المجتمع موضع الاهتمام. وتيسيراً على القارئ سوف ندعم عرضنا للموضوع بالأمثلة الرقمية وذلك على النحو الموضح فيما يلي:

التقدير بفترة ثقة لمتوسط المجتمع

في البداية نرى أنه من الملائم أن نتعرض بالدراسة لبعض المفاهيم والتعريفات الإحصائية المتعلقة بتطبيق نظرية التقدير على متوسط المجتمع.

أولاً: توزيع المعاينة للمتوسط

إذا أخذنا عينات عشوائية متكرِّرة من مجتمع ما (أو جميع العينات الممكنة ذات حجم معين) وحسبنا الوسط الحسابي في كل مرة نختار فيها عينة من المجتمع فإننا سوف نجد أن هذا الوسط يختلف من عينة لأخرى. أي أن متوسط العينة يعتبر متغيراً عشوائياً يمكن إيجاد التوزيع الاحتمالي له والذي يعرف عادة باسم "توزيع المعاينة للمتوسط". وهذا يعنى أن توزيع المعاينة للمتوسط هو التوزيع الاحتمالي لمتوسطات العينات ذات الحجم الواحد والمأخوذة من نفس المجتمع. وهذا التوزيع له وسط حسابي وانحراف معياري (أو خطأ معياري) سوف نرمز لهما على الترتيب بالرمزين $م_{\bar{س}}$ ، $\sigma_{\bar{س}}$. وفي هذا الشأن، هناك نظريتان هامتان تربطان ما بين توزيع المعاينة للمتوسط وبين معلمات المجتمع الأصلي سوف نتناول كلاً منهما على النحو التالي:

نظرية (١):

إذا كان لدينا مجتمع حجمه $ن$ من المفردات وله توزيع معتاد وسطه الحسابي $م$ وتباينه $\sigma^{٢}$، وسحبنا منه عينات عشوائية متكرر حجم كل منها $ن$ (أو جميع العينات الممكنة ذات الحجم $ن$). فإن الوسط الحسابي للعينة ($\bar{س}$) تختلف قيمته من عينة لأخرى، ويمثل متغيراً عشوائياً له توزيع معتاد وسطه الحسابي وانحرافه المعياري يُعرَّفان كما يلي:

$$م_{\bar{س}} = م \qquad (٢ - ١)$$

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} \quad \text{في حالة المجتمع غير المحدود أو } ن < ٠,٠٥ ن \qquad (٢ - ٢)$$

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} \sqrt{\frac{ن - ن}{ن - ١}} \quad \text{في حالة المجتمع غير المحدود أو } ن > ٠,٠٥ ن \qquad (٢ - ٣)$$

حيث:

$\text{م}_{\bar{\text{س}}}$ الوسط الحسابي لتوزيع المعاينة (أي الوســــط الحسابي لمتوسطات العينات)،

م الوسط الحسابي للمجتمع،

$\sigma_{\bar{\text{س}}}$ الانحراف المعياري لتوزيع المعاينة (أي الانحراف المعياري لمتوسطات العينات)، ويُعرف بالخطأ المعياري،

σ الانحراف المعياري للمجتمع،

ن حجم كل عينة من العينات المحسوبة،

ن حجم المجتمع.

وهذه النظرية صحيحة دائماً أياً كان حجم العينة ن صغيراً كان أم كبيراً. وفيما يلي سوف نتحقق - حسابياً - من صحة هذه النظرية مستعينين في ذلك ببيانات المثال (١) وذلك على النحو التالي:

الجدول (٢ - ١) يوضح الوسط الحسابي لكل عينة ذات حجم ٤ مفردات يمكن سحبها من مجتمع حجمه ٦ مفردات. وباستخدام متوسطات العينات المختلف يمكننا حساب المتوسط الحسابي لهذه المتوسطات ($\text{م}_{\bar{\text{س}}}$) وكذلك الانحراف المعياري للمتوسطات أي الخطأ المعياري ($\sigma_{\bar{\text{س}}}$) وذلك كما يلي:

$$\text{م}_{\bar{\text{س}}} = \frac{\sum_{\text{ر}=١}^{١٥} \bar{\text{س}}_{\text{ر}}}{١٥}$$

$$= \frac{١٤,٢٥ + ١٤,٧٥ + ١٤,٥ + ... + ١٥,٥ + ١٤,٥}{١٥}$$

$$= \frac{٢٢٥}{١٥} = ١٥ \text{ درجة}$$

وحيث أن:

$$\text{م} = \frac{١٨ + ١٤ + ١٠ + ١٥ + ١٧ + ١٦}{٦} = \frac{٩٠}{٦} = ١٥ \text{ درجة}$$

فإنه يمكننا استنتاج أن:

$$\text{م}_{\bar{\text{س}}} = \text{م}$$

أي أن الوسط الحسابي لمتوسطات جميع العينات الممكنة ذات الحجم الواحد يساوي في قيمته الوسط الحسابي للمجتمع الذي أخذت منه هذه العينات.

وهذه العلاقة صحيحة سواء كان المجتمع محدوداً أو غير محدود، مع ملاحظة أنه في حالة المجتمع المحدود يجب أخذ جميع العينات الممكنة ذات الحجم الواحد (ن). وأما إذا كان

المجتمع غير محدود فإنه يجب الاستمرار في أخذ عينات متكررة ذات الحجم (ر) وذلك إلى ما لا نهاية. ويمكن أن نعبر عن ذلك بصورة أخرى كما يلي:

$$\text{ت}(\bar{\text{س}}) = \text{م}$$

أي أن $\bar{\text{س}}$ هو تقدير غير متحيز لمتوسط المجتمع (م)، وهو ما سبق الإشارة إليه.
كما يمكننا حساب الانحراف المعياري للمتوسطات (الخطأ المعياري) وذلك على الصورة:

$$\sigma_{\bar{\text{س}}} = \sqrt{\frac{\sum(\bar{\text{س}} - \text{م}_{\bar{\text{س}}})^{\text{٢}}}{\text{ك}}}$$ نلاحظ أن ك هنا ترمز إلى عدد العينات

$$= \sqrt{\frac{\sum(\bar{\text{س}} - \text{م})^{\text{٢}}}{\text{ك}}}$$

$$= \sqrt{\frac{(\text{-٠,٧٥})^{\text{٢}} + (\text{-٠,٢٥})^{\text{٢}} + \ldots\ldots + (\text{-٠,٥})^{\text{٢}}}{\text{١٥}}} = \sqrt{\frac{\text{١٠}}{\text{١٥}}} = \text{٠,٨١٦}$$ درجة

وهنا تجب الإشارة إلى أن عدد العينات يساوي $\text{ق}^{\text{ن}}_{\text{ر}}$ حيث ن حجم المجتمع و ر حجم العينة.
وحيث أن الانحراف المعياري للمتوسطات يقيس درجة انحراف متوسط العينة عن متوسط المجتمع (يلاحَظ أن: $\text{م}_{\bar{\text{س}}} = \text{م}$)، فإن الانحراف المعياري للمتوسطات إنما يقيس درج الخطأ في قيمة الوسط الحسابي للعينـة والذي (أي الخطأ) ينشأ في المتوسط نتيجة لاتباع المعاينة العشوائية. وهذا هو سبب تسميته بالخطأ المعياري للمتوسط Standard Error ($\sigma_{\bar{\text{س}}}$).
وعند تناولنا لهذا المقياس سوف ندعوه من الآن فصاعداً - وذلك من باب التيسير – باسم الخطأ المعياري.
وحيث أن:

$$\sigma = \sqrt{\frac{\sum(\text{س} - \text{م})^{\text{٢}}}{\text{ن}}}$$

$$= \sqrt{\frac{(\text{١٨} - \text{١٥})^{\text{٢}} + (\text{١٤} - \text{١٥})^{\text{٢}} + \ldots\ldots + (\text{١٦} - \text{١٥})^{\text{٢}}}{\text{٦}}} = \text{٢,٥٨٢}$$ درجة

وفي هذا المثال نجد أن: ن = ٦ ، ر = ٤

أي أن المجتمع محدود، كما أن حجم العينة كبير نسبياً بالنسبة لحجم المجتمع. ولذلك فإنه يمكن تطبيق العلاقة (٢ - ٣) لحساب الخطأ المعياري، أي أن:

$$\sigma_{\bar{\text{س}}} = \frac{\sigma}{\sqrt{\text{ر}}}\sqrt{\frac{\text{ن} - \text{ر}}{\text{ن} - \text{١}}} = \frac{\text{٢,٥٨٢}}{\sqrt{\text{٤}}}\sqrt{\frac{\text{٦} - \text{٤}}{\text{٦} - \text{١}}} = \frac{\text{٢,٥٨٢}}{\text{٢}}\sqrt{\text{,٦٣٢}} = \text{٠,٨١٦}$$

وهي نفس النتيجة التي حصلنا عليها باستخدامنا للطريقة المباشرة في حساب الخطأ المعياري. وهنا تجدر الإشارة إلى أن $\sigma_{\bar{س}}$ تكون دائما أصغر من σ وذلك لأن الأوساط الحسابية للعينات تكون أقل اختلافاً أو انتشاراً من قيم المجتمع.

وفي حالة ما إذا كان المجتمع غير محدود (أي أن: $ن \leftarrow \infty$)، أو كانت $ى$ صغيرة نسبياً بالنسبة إلى $ن$ فإنه يلاحظ أن المقدار $\frac{ن - ى}{ن - ١}$ يؤول إلى الواحد الصحيح. حينئذ نجد أن العلاقة (٢ - ٣) تؤول إلى العلاقـــة (٢ - ٢). ويسمى المقدار $\sqrt{\frac{ن - ى}{ن - ١}}$ بمعامل التصحيح للمجتمعات المحـدودة Finite Correction.

وبصفة عامة يمكن القول إنه يمكن إهمال معامل التصحيح وذلك في الحالات التي يكون فيها حجم المجتمع كبيراً أو أن حجم العينة يكون صغيراً جداً بالنسبة لحجم المجتمع. وفي كلتا الحالتيْن تقترب قيمة معامل التصحيح من الواحد الصحيح. وقد جرت العادة على إهمال هذا المعامل في الحالات التي تكون فيها ($ى < ٠,٠٥ \; ن$). وفي جميع الحالات فإن الخطأ المعياري يتناسب طردياً مع الانحراف المعياري للمجتمع وعكسياً مع حجم العينة. وعلى سبيل المثال لو افترضنا أننا ضاعفنا حجم العينة أربع مرات فإن ذلك سيؤدى إلى زيادة دقة $\bar{س}$ كتقدير للوسط الحسابي للمجتمع (م) وذلك بإنقاص قيمة الخطأ المعياري ($\sigma_{\bar{س}}$) إلى النصف (راجع العلاقة ٢ - ٢).

ويوضح الجدول (٢ - ٢) توزيع المعاينة للمتوسط وذلك من واقع بيانات المثال (١).

هذا ولو حاولنا حساب الوسط الحسابي والخطأ المعياري لمتوسطات العينات مستخدمين في ذلك توزيع المعاينة للمتوسط (والمكوَّن من الصفيْن الأول والثالث من الجدول (٢ - ٢)) - وذلك بدلاً من القيم الأصلية - لحصلنا على نفس النتائج السابقة. أي أن:

$$م_{\bar{س}} = ١٥ \text{ درجة} \quad ، \quad \sigma_{\bar{س}} = ٠,٨١٦ \text{ درجة}$$

مع ملاحظة أن الوسط الحسابي والخطأ المعياري يمكن حسابهما باستخدام توزيع المعاينة وذلك على النحو التالي:

$$م_{\bar{س}} = \sum \bar{س} \times ع(\bar{س})$$

$$= ١٣,٧٥ \times \frac{١}{١٥} + ١٤ \times \frac{١}{١٥} + ١٤,٢٥ \times \frac{٢}{١٥} + ...$$

$$+ ١٦,٥ \times \frac{١}{١٥} = \frac{٢٢٥}{١٥} = ١٥ \text{ درجة}$$

الجدول (٢ - ٢)
توزيع المعاينة لمتوسط عينة حجمها ٤ مفردات مسحوبة من مجتمع حجمه ٦ مفردات

قيمة المتوسط ($\bar{س}$)	١٣,٧٥	١٤	١٤,٢٥	١٤,٥	١٤,٧٥	١٥	١٥,٢٥	١٥,٥	١٥,٧٥	١٦	١٦,٢٥	١٦,٥
عدد مرات التكرار	١	١	٢	٢	٢	١	١	١	١	١	١	١
ع($\bar{س}$)	$\frac{١}{١٥}$	$\frac{١}{١٥}$	$\frac{٢}{١٥}$	$\frac{٢}{١٥}$	$\frac{٢}{١٥}$	$\frac{١}{١٥}$	$\frac{١}{١٥}$	$\frac{١}{١٥}$	$\frac{١}{١٥}$	$\frac{١}{١٥}$	$\frac{١}{١٥}$	$\frac{١}{١٥}$

$$\sigma_{\bar{س}} = \sum (\bar{س} - م)^{٢} \times ع(\bar{س})$$

$$= (١٣,٧٥ - ١٥)^{٢} \times \frac{١}{١٥} + (١٤ - ١٥)^{٢} \times \frac{١}{١٥}$$

$$+ (١٤,٢٥ - ١٥)^{٢} \times \frac{٢}{١٥} + ... + (١٦,٥٠ - ١٥)^{٢} \times \frac{١}{١٥}$$

$$= \frac{١٠}{١٥} = ٠,٦٦٦٧$$

$\therefore \ \sigma_{\bar{س}} = \sqrt{٠,٦٦٦٧} = ٠,٨١٦$ درجة

وهي نفس النتائج التي حصلنا عليها من قبل.

وفي هذا الشأن يلزم التأكيد مرة أخرة على أن الوسط الحسابي للعينة ($\bar{س}$) هو مقدِّر غير متحيز Unbiased Estimator للوسط الحسابي للمجتمع (م) والذي يكون مجهولاً عادة، مع ملاحظة أن المقدِّر يكون غير متحيز إذا كانت قيمته المتوقَّعة مساوية لمعلمة المجتمع موضع التقدير. وهو ما ينطبق على $\bar{س}$.

وأما فيما يتعلق بالانحراف المعياري للمجتمع (غير المعروف عادة)، فإنه يمكننا استخدام الانحراف المعياري المحسوب من العينة (ع) وذلك كتقدير للانحراف المعياري للمجتمع (σ)، حيث:

$$ع = \sqrt{\frac{\sum (س - \bar{س})^{٢}}{ن - ١}}$$

$$= \sqrt{\frac{١}{ن - ١}\left[\sum س^{٢} - \frac{\left(\sum س\right)^{٢}}{ن}\right]} \qquad (٢ - ٤)$$

حينئذ يعطي الخطأ المعياري بالصورة:

$$ع_{\bar{س}} = \frac{ع}{\sqrt{ن}} \qquad (٢ - ٥)$$

في حالة المجتمع غير المحدود أو $ن < ٠,٠٥ \ ن$،

$$ع_{\bar{س}} = \frac{ع}{\sqrt{ن}} \sqrt{\frac{ن - ن}{ن - ١}} \qquad (٢ - ٦)$$

في حالة المجتمع المحدود أو $ن > ٠,٠٥ \ ن$

وأما الانحراف المعياري المعرَّف بالصورة:

$$ع = \sqrt{\frac{\sum(س - \bar{س})^2}{ن}}$$

$$ع = \sqrt{\frac{\sum س^2}{ن} - \left(\frac{\sum س}{ن}\right)^2} \qquad (٢ - ٧)$$

فإنه مقدِّر غير ملائم للانحراف المعياري للمجتمع (σ)، وسوف نتعرض لإيضاح هذه النقطة في موضع لاحق في هذا الكتاب.

لذلك فإنه عند دراستنا لموضوع فترات الثقة سوف نستخدم الصورة (٢ - ٤) وذلك لحساب الانحراف المعياري (ع) من بيانات العينة.

مثال (٢):

إذا كان عدد طلاب الفرقة الثالثة في إحدى كليات التجارة هو ١٢٠٠ طالب، وكان الوسط الحسابي والانحراف المعياري لدرجاتهم في مادة المحاسبة ١٥ درجة و٥ درجات على الترتيب. فإذا قمنا بسحب عينة من طلاب الفرقة الثالثة بهذه الكلية، احسب الوسط الحسابي والخطأ المعياري توزيع المعاينة للمتوسط وذلك في حالة:

أ- حجم العينة يساوي ٤٠ طالباً.

ب- حجم العينة يساوي ١٠٠ طالب.

الحــل:

أ- حيث أن: $م_{\bar{س}} = م$

$\therefore$ $م_{\bar{س}} = ١٥$ درجة

وحيث أن حجم العينة هو ٤٠ طالباً، وهو أقل من ٥٪ من حجم المجتمع فإنه يمكن إهمال معامل التصحيح في حسابنا للخطأ المعياري أي أن:

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} = \frac{٥}{\sqrt{٤٠}} = \frac{٥}{٦,٣٢} = ٠,٧٩ \text{ درجة}$$

وهذا يعنى أنه لو قمنا بسحب عدد كبير من العينات حجم كل منها ٤٠ طالباً وذلك من مجتمع طلاب الفرقة الثالثة فإنه من المتوقَّع أن تكون مجموعة المتوسطات للعينات المسحوبة ذات وسط حسابي ١٥ درجة وانحراف معياري ٠,٧٩ درجة.

ب- هنا أيضاً:

$م_{\bar{س}} = م = ١٥$ درجة

وحيث أن حجم العينة هو ١٠٠ طالب، وهو أكبر من ٥٪ من حجم المجتمع فإنه يجب أخذ معامل التصحيح في الاعتبار عند حسابنا للخطأ المعياري. أي أن:

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} \sqrt{\frac{ن - ن}{ن - ١}}$$

$$= \frac{٥}{\sqrt{١٠٠}} \sqrt{\frac{١٢٠٠ - ١٠٠}{١٢٠٠ - ١}}$$

$$= \frac{٥}{١٠} \times \sqrt{\frac{١١٠٠}{١١٩٩}} = ٠{,}٤٨$$

أي أننا لو قمنا بسحب عدد كبير من العينات حجم كل منها ١٠٠ طالب فإنه من المتوقَّع أن يبلغ الوسط الحسابي لمتوسطات العينات ١٥ درجة كما يبلغ الانحراف المعياري لهذه المتوسطات ٠,٤٨ درجة.

مثال (٣):

مجتمع مكون من خمسة طلاب، فإذا كانت درجاتهم في مادة الإحصاء هي ١٦، ١٨، ١٥، ١٩، ١٢ درجة. المطلوب:

أ- حساب م، σ (الوسط الحسابي والانحراف المعياري للمجتمع).

ب- إيجاد توزيع المعاينة لمتوسط عينة حجمها طالبان.

ج- حساب الوسط الحسابي والخطأ المعياري لتوزيع المعاينة (أي $م_{\bar{س}}$ و $\sigma_{\bar{س}}$)

الحــل:

أ- الوسط الحسابي للمجتمع (م) $= \frac{\sum س}{ن}$

$$= \frac{١٦ + ١٨ + ١٥ + ١٩ + ١٢}{٥} = \frac{٨٠}{٥} = ١٦ \text{ درجة}$$

الانحراف المعياري للمجتمع (σ) $= \sqrt{\frac{\sum (س - م)^٢}{ن}}$

$$= \sqrt{\frac{(١٦ - ١٦)^٢ + (١٨ - ١٦)^٢ + ... + (١٢ - ١٦)^٢}{٥}}$$

$$= \sqrt{\frac{٣٠}{٥}} = \sqrt{٦} = ٢{,}٤٥ \text{ درجة}$$

ب- توزيع المعاينة لمتوسط عينة حجمها طالبان هو عبارة عن متوسطات كل العينات ذات الحجم ٢ والتي يمكن سحبها من المجتمع واحتمالات الحصول على هذه المتوسطات.

وحيث أن حجم المجتمع = ٥ فإنه يمكن سحب عدداً من العينات يساوي ($^{5}ق_{2}$ = ١٠ عينات) وبيانها هو:
(١٦، ١٨) ، (١٦، ١٥) ، (١٦، ١٩) ، (١٦، ١٢) ، (١٨، ١٥) ، (١٨، ١٩) ، (١٨، ١٢) ، (١٥، ١٩) ، (١٥، ١٢) ، (١٩، ١٢)

والأوساط الحسابية لهذه العينات هي على الترتيب:
١٧ ، ١٥,٥ ، ١٧,٥ ، ١٤ ، ١٦,٥ ، ١٨,٥ ، ١٥ ، ١٧ ، ١٣,٥ ، ١٥,٥

هذا ويوضح الجدول (٣ - ٣) توزيع المعاينة المطلوب.

الجدول (٣ - ٣)
توزيع المعاينة لمتوسط عينة
حجمها ٢ مسحوب من مجتمع حجمه ٥

قيمة المتوسط ($\bar{س}$)	عدد مرات التكرار	الاحتمال ع($\bar{س}$)
١٣,٥	١	٠,١
١٤,٠	١	٠,١
١٥,٠	١	٠,١
١٥,٥	٢	٠,٢
١٦,٥	١	٠,١
١٧,٠	٢	٠,٢
١٧,٥	١	٠,١
١٨,٥	١	٠,١
المجموع		١,٠

ج- بتطبيق نظرية (١) المشار إليها سابقاً نجد أن:

$$م_{\bar{س}} = م = ١٦ \text{ درجة}$$

وحيث أن حجم العينة = ٢ وهو أكبر من ٥٪ من حجم المجتمع (أي أن: ن > ٠,٠٥ ن)
فإنه يمكن حساب الخطأ المعياري كما يلي:

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} \sqrt{\frac{ن - ن}{ن - ١}}$$

$$= \frac{٢,٤٥}{\sqrt{٢}} \sqrt{\frac{٥ - ٢}{٥ - ١}} = \frac{٢,٤٥}{١,٤١} \times \sqrt{\frac{٣}{٤}} = ١,٥ \text{ درجة}$$

ملحوظة:

يمكن للقارئ التحقق من أنه لو حسبنا الانحراف المعياري للأوساط الحسابية للعينات المختلفة فإننا نحصل على نفس النتيجة.

مثال (٤):

إذا كانت حسابات التوفير لدى مكتب بريد إحدى القرى تتبع توزيعاً معتاداً له متوسط ١٥٠ جنيهاً وانحراف معياري ٥٠ جنيهاً. فإذا سحبنا عينة عشوائية مكونة من ٢٥ حساباً، احسب احتمال أن يزيد متوسطها عن ١٤٥ جنيهاً.

الحــل:

حيث أن المجتمع الأصلي موزع توزيعاً معتاداً فإن توزيع المعاينة للمتوسط يتبع أيضاً التوزيع المعتاد وذلك بغض النظر عن حجم العينة (نظرية ١). وأما الوسط الحسابي والخطأ المعياري لتوزيع المعاينة فيمكن حسابهما كما يلي:

$$م_{\bar{س}} = م = ١٥٠ \text{ جنيهاً}$$

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} = \frac{٥٠}{\sqrt{٢٥}} = ١٠ \text{ جنيهات}$$

وحيث أن: $ي = \frac{\bar{س} - م}{\sigma_{\bar{س}}}$ تتبع التوزيع المعتاد المعياري، فإنه يمكن حساب الاحتمال المطلوب على النحو التالي:

$$ع(\bar{س} > ١٤٥) = ١ - ع(\bar{س} \leq ١٤٥)$$

$$= ١ - ع(\frac{\bar{س} - ١٥٠}{١٠} \leq \frac{١٤٥ - ١٥٠}{١٠})$$

$$= ١ - ع(ي \leq -٠,٥) = ١ - \phi(-٠,٥)$$

$$= ١ - [١ - \phi(٠,٥)] = \phi(٠,٥) = ٠,٦٩١٥$$

نظرية (٢): نظرية الحد المركزية

إذا كان لدينا مجتمع ما (ليس بالضرورة أن يكون معتاداً) وسطه الحسابي م وتباينه $\sigma^٢$، وسحبنا منه عينة حجمها ن. فإنه عندما تكون ن كبيرة بدرجة كافية فإن توزيع الأوساط الحسابية لجميع العينات الممكنة (ذات الحجم ن) يقترب من التوزيع المعتاد بوسط حسابي وانحراف معياري يعرَّفان كما يلي:

$$م_{\bar{س}} = م$$

$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}}$ في حالة المجتمع غير المحدود أو: $ن < ٠,٠٥ \, N$

أو $\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} \sqrt{\frac{N - ن}{N - ١}}$ في حالة المجتمع المحدود أو: $ن > ٠,٠٥ \, N$

ويكون التقريب جيداً إذا بلغ حجم العينة ٣٠ مفردة على الأقل، أي أن توزيع المعاينة للمتوسط يقترب من التوزيع المعتاد أياً كان شكل المجتمع المسحوب منه العينة وذلك إذا كان حجم العينة ٣٠ مفردة على الأقل. ويزداد توزيع المعاينة للمتوسط اقتراباً من التوزيع المعتاد كلما زاد حجم العينة.

هذا ونظرية الحد المركزية هي أهم نظرية في الاستدلال الإحصائي، حيث ترجع أهميتها إلى أنها تسمح لنا باستخدام التوزيع المعتاد في تقدير معلمات المجتمع وذلك دون معرفة أي شيء عن طبيعة المجتمع الأصلي والتي عادة ما تكون مجهولة.

وكنتيجة للنظرية الحد المركزية فإن توزيع المتغير (ي) حيث:

$$ي = \frac{\bar{س} - م}{\frac{\sigma}{\sqrt{ن}}}$$

يؤول إلى التوزيع المعتاد المعياري وذلك بغض النظر عن شكل التوزيع الخاص بالمتغير س في المجتمع الأصلي وبشرط أن تكون ن كبيرة كبراً كافياً (ن ≥ ٣٠).

مثال (٥):

يتكون الجهاز الإداري لإحدى الشركات الكبرى من ٢٠٠٠ موظف. وكان متوسط الدخل السنوي الموظفين في هذه الشركة هو ١٢٠ والانحراف المعياري للدخل هو ٣٥. فإذا أخذت الشركة عينة عشوائية مكونة من ٣٠٠ موظف، احسب احتمال أن متوسط الدخل السنوي الموظفين في العينة سوف يبلغ ١٢٥ على الأكثر (الدخل مقاس بالألف جنيه).

الحـــل:

قبل حساب الاحتمال المطلوب يلزم إيجاد توزيع المعاينة للمتوسط وذلك على النحو التالي: $م_{\bar{س}} = م = ١٢٠$

وحيث أن ن = ٣٠٠ > ٠,٠٥ ن، فإن يمكن حساب الخطأ المعياري آخذين في الاعتبار معامل التصحيح. أي أن:

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} \sqrt{\frac{ن - ن}{ن - ١}}$$

$$= \frac{٣٥}{\sqrt{٣٠٠}} \sqrt{\frac{٢٠٠٠ - ٣٠٠}{٢٠٠٠ - ١}}$$

$$= \frac{٣٥}{١٧,٣٢} \times \sqrt{٠,٨٥} = \frac{٣٥}{١٧,٣٢} \times ٠,٩٢ = ١,٨٦$$

وحيث أن: $ن > ٣٠$ فإن توزيع المعاينة للمتوسط يتبع التوزيع المعتاد وذلك بغض النظر عن توزيع المرتبات في مجتمع الشركة. أي أن: $ي = \frac{\bar{س} - م}{\sigma_{\bar{س}}}$ تتبع التوزيع المعتاد المعياري:

$$\therefore\ ع(\bar{س} \le ١٢٥) = ع(\frac{\bar{س} - ١٢٠}{١,٨٦} \le \frac{١٢٥ - ١٢٠}{١,٨٦})$$

$$= ع(ي \le ٢,٦٩) = \phi(٢,٦٩) = ٠,٩٩٦٤$$

ثانياً: تقدير فترة الثقة لمتوسط المجتمع

أشرنا فيما سبق إلى أنه إذا كان لدينا مجتمع (ليس بالضرورة أن يكون معتاداً) متوسطه $م$ وتباينه $\sigma^٢$، وسحبنا جميع العينات الممكنة والتي حجم كل منها $ن$، فإنه إذا كانت $ن$ كبيرة كبراً كافياً ($ن \ge ٣٠$) فإن توزيع المعاينة لمتوسطات هذه العينات يقترب من التوزيع المعتاد بمتوسط $م$ وتباين $\frac{\sigma^٢}{ن}$ على الترتيب. وأما إذا كان التوزيع الأصلي للمجتمع معتاداً بمتوسط $م$ وتباين $\sigma^٢$ فإن توزيع المعاينة لمتوسطات العينات هو توزيع معتاد بمتوسط $م$ وتباين $\frac{\sigma^٢}{ن}$ وذلك بغض النظر عن حجم العينة صغيراً كان أم كبيراً.

وباستخدام هذه الخواص وبافتراض أن المجتمع يتبع التوزيع المعتاد نجد أن المتغير العشوائي:

$$ي = \frac{\bar{س} - م}{\frac{\sigma}{\sqrt{ن}}}$$

يتبع التوزيع المعتاد المعياري.

وهنا لو أردنا تكوين فترة ثقة لمتوسط المجتمع وذلك بمعامل ثقة يساوي ٠,٩٥. على سبيل المثال، فإنه باستخدام جدول المساحات تحت المنحنى المعتاد يمكننا استنتاج أن:

$$ع(-١,٩٦ \le ي \le ١,٩٦) = ٠,٩٥$$

$$\therefore\ ع(-١,٩٦ \le \frac{\bar{س} - م}{\frac{\sigma}{\sqrt{ن}}} \le ١,٩٦) = ٠,٩٥$$

وهذا يعني أن:

$$ع(-١,٩٦(\frac{\sigma}{\sqrt{ن}}) \le \bar{س} - م \le ١,٩٦(\frac{\sigma}{\sqrt{ن}}) = ٠,٩٥$$

$$ع(\bar{س} - ١,٩٦(\frac{\sigma}{\sqrt{ن}}) \le م \le \bar{س} + ١,٩٦(\frac{\sigma}{\sqrt{ن}})) = ٠,٩٥$$

ومن ذلك نستنتج أن الوسط الحسابي للمجتمع (م) يقع في الفترة:

$$\left(\bar{س} - ١,٩٦\left(\frac{\sigma}{\sqrt{ن}}\right) ، \bar{س} + ١,٩٦\left(\frac{\sigma}{\sqrt{ن}}\right)\right)$$

وذلك باحتمال يساوي ٠,٩٥، حيث:

$\bar{س} - ١,٩٦\left(\frac{\sigma}{\sqrt{ن}}\right)$ هو الحد الأدنى لفترة ثقة ٩٥٪.

$\bar{س} + ١,٩٦ \frac{\sigma}{\sqrt{ن}}$ هو الحد الأعلى لفترة ثقة ٩٥٪.

وأما إذا كان معامل الثقة المستخدم هو ٠,٩٩ فإن الحد الأدنى والحد الأعلى للثقة يتحددان كما يلي:

$$\bar{س} - ٢,٥٨\left(\frac{\sigma}{\sqrt{ن}}\right) \quad و \quad \bar{س} + ٢,٥٨\left(\frac{\sigma}{\sqrt{ن}}\right)$$

أي أن:

$$ع\left(\bar{س} - ٢,٥٨\left(\frac{\sigma}{\sqrt{ن}}\right) \leq م \leq \bar{س} + ٢,٥٨\left(\frac{\sigma}{\sqrt{ن}}\right)\right) = ٠,٩٩$$

وبشكل عام، فإن الحد الأدنى والحد الأعلى للثقة يتحددان كما يلي:

$$ع\left(\bar{س} - ي_{\frac{\alpha}{٢}} \frac{\sigma}{\sqrt{ن}} \leq م \leq \bar{س} + ي_{\frac{\alpha}{٢}} \frac{\sigma}{\sqrt{ن}}\right) = ١ - \alpha$$

حيث:

$\bar{س}$: الوسط الحسابي للعينة ، م: الوسط الحسابي للمجتمع،

σ: الانحراف المعياري للمجتمع ، $١ - \alpha$ معامل الثقة،

$ي_{\frac{\alpha}{٢}}$: قيمة الدرجة المعيارية (ي) وتستخرج من جدول المساحات تحت المنحنى المعتاد المعياري بحيث: $\phi(ي) = ١ - \frac{\alpha}{٢}$

وحيث:

$ي_{\frac{\alpha}{٢}} = ١,٦٤٥$ في حالة: $\alpha = ٠,١٠$ ، أي $١ - \alpha = ٠,٩٠$

$= ١,٩٦$ في حالة: $\alpha = ٠,٠٥$ ، أي $١ - \alpha = ٠,٩٥$

$= ٢,٥٨$ في حالة: $\alpha = ٠,٠١$ ، أي $١ - \alpha = ٠,٩٩$

$= ٣,٠٠$ في حالة: $\alpha = ٠,٠٠٣$ ، أي $١ - \alpha = ٠,٩٩٧$

كما يمكن إيجاد قيم $ي_{\frac{\alpha}{٢}}$ المناظرة للقيم المختلفة الأخرى لـ α.

هذا وإذا كان المطلوب هو تقدير إحدى معلمات مجتمع ما. فإنه باستخدام نتائج عينة عشوائية مسحوبة من هذا المجتمع سوف تختلف قيمة التقدير من عينة لأخرى ويؤخذ تأثير هذا الاختلاف في الاعتبار باستخدام فترة ثـقة لتقدير المعلمة المطلوبة. وبشكل عام، يمكن تحديد هذه الفترة على النحو التالي:

الحد الأدنى لفترة الثـقة = المقدِّر بنقطة للمعلمة – الدرجة المعيارية × الانحراف المعياري للمقدِّر بنقطة للمعلمة

الحد الأعلى لفترة الثـقة = المقدِّر بنقطة للمعلمة + الدرجة المعيارية × الانحراف المعياري المقدِّر بنقطة للمعلمة

أي أن فترة الثـقة تتحدد كما يلي:

المقدِّر بنقطة للمعلمة ± الدرجة المعيارية × الانحراف المعياري للمقدِّر بنقطة للمعلمة

هذا مع ملاحظة أن الانحراف المعياري لمقدِّر المعلمة في حالتنا هذه ($\bar{س}$) هو ما يعرف بالخطأ المعياري ($\sigma_{\bar{س}}$). أي أنه يمكن التعبير عن فترة الثـقة لمتوسط المجتمع ($م$) على النحو التالي:

الحد الأدنى لفترة الثـقة = $\bar{س}$ – الدرجة المعيارية × الخطأ المعياري ($\sigma_{\bar{س}}$)

الحد الأعلى لفترة الثـقة = $\bar{س}$ + الدرجة المعيارية × الخطأ المعياري ($\sigma_{\bar{س}}$)

وفي هذا الشأن، تجدر الإشارة إلى عدة نقاط:

١- أن مركز فترة الثـقة (أي نقطة منتصفها) هو الوسط الحسابي للعينة ($\bar{س}$) والذي يمثِّل تقديراً غير متحيز لمتوسط المجتمع ($م$).

٢- أن طول فترة الثـقة يتوقف على ثلاثة عوامل هي:

درجة الثـقة المطلوبة ($١ - \alpha$) × ١٠٠:

والتي تؤثر بدورها (طردياً) على قيمة الدرجة المعيارية. حيث نجد أنه كلما زادت درجة الثـقة المطلوبة كلما أدى ذلك إلى زيادة طول فترة الثـقة، والعكس صحيح. هذا بافتراض ثبات العوامل الأخرى. وربما يؤدي ذلك لا يؤدي إلى الفهم الخاطئ بأن زيادة طول فترة الثـقة إلى حد كبير والذي من شأنه أن يجعلنا نثق في تقديرنا ثـقة تامة (أي بدرجة ١٠٠٪) هو أمر مقبول. إذ أن ذلك يعني في بعض الحالات عدم جدوى عملية التقدير في حد ذاتها. ولتوضيح هذه النقطة دعنا نفترض المثال التالي:

إذا قلنا إننا نثق ثـقة تامة (أي بدرجة ١٠٠٪) بأن متوسط أعمار الطلاب في إحدى كليات التجارة يتراوح ما بين عشر سنوات ومائة سنة. هنا، وفي هذا المثال هل تضمن درجة الثـقة – وهي تامة – جدوى ما توصلنا إليه من نتيجة؟ الإجابة هي قطعاً بالنفي، إذ أن النتائج في مثل تلك الحالات لا تخرج عن مجرد كونها نوعاً من البديهيات التي لا تحتاج أساساً إلى

عمليات التقدير. ومن هنا، فإن الأمر يحتاج في الحكم على جدواه إلى عوامل أخرى مثل دقة التقدير وهو ما نوضحه فيما يلي:

حجم العينة ($ن$):

كلما زاد حجم العينة كلما أدى ذلك إلى نقص طول فترة الثقة وبالتالي إلى زيادة دقة التقدير والعكس صحيح، وذلك بافتراض ثبات العوامل الأخرى. وهذا أمر منطقي، إذ إنه كلما زاد حجم العينة كلما كانت أكثر قرباً لواقع المجتمع التي أخذت منه، مما يعني ارتفاع دقة التقدير وكذلك زيادة درجة الثقة في هذا التقدير. والحالة المتطرفة في هذا الشأن هي أن نفترض أن حجم العينة زاد إلى الحد الذي وصل فيه إلى حجم المجتمع نفسه (تحولت الحالة هنا إلى حصر شامل بدلاً من معاينة). حينئذ فإن دقة التقدير ودرجة الثقة فيه تبلغان حد الـ ١٠٠٪.

تباين المجتمع ($\sigma^{٢}$):

حيث أنه بزيادة تباين المجتمع $\sigma^{٢}$ وبالتالي زيادة σ فإن طول فترة الثقة يزداد والعكس الصحيح. ولعل تلك العلاقة تتضح فيما لو افترضنا حالتيْن متطرفتيْن، أولهما: أن تباين المجتمع يساوي الصفر، وهذا يعني أن المجتمع متجانس تجانساً تاماً، وهذا يؤدي إلى انه وفقاً لتعريف الحديْن الأدنى والأعلى لفترة الثقة فإن الوسط الحسابي للمجتمع يساوي الوسط الحسابي للعينة لأن قيمة المقدار المضاف إلى أو المطروح من $\bar{س}$ يساوي الصفر (لأن σ = صفر)، وهذا صحيح في حالة التجانس التام للمجتمع. ويلاحظ أن طول فترة الثقة في هذه الحالة هو الصفر. أي أن عملية التقدير تحولت إلى تقدير بنقطة بدلاً من تقدير بفترة ثقة.

وأما الحالة الثانية فهي إذا ما افترضنا أن تباين المجتمع كبير جداً، بمعنى أن المجتمع يتكون من طبقات غير متجانسة إلى حد بعيد. وفي هذه الحالة فإنه ما لم تكن $ن$ كبيرة كبراً كافياً حتى يمثِّل المجتمع بالعينة تمثيلاً صحيحاً فإن كبر $\sigma^{٢}$ سوف يؤدي بالضرورة إلى زيادة طول فترة الثقة وبالتالي إلى عدم دقة التقدير (يُلاحَظ أن زيادة $\sigma^{٢}$ يؤدي إلى زيادة قيمة المقدار المضاف إلى أو المطروح من $\bar{س}$)

هذا ونظراً لأن الانحراف المعياري للمجتمع لا يكون معروفاً دائماً. كما أن حجم العينة أحياناً ما يكون صغيراً ($ن < ٣٠$). أضف إلى ذلك أن توزيع المجتمع الأصلي قد يكون معتاداً وقد يكون غير ذلك. فإن التقدير بفترة لمتوسط المجتمع يختلف من حالة لأخرى. حيث يتوقف ذلك على:

- الانحراف المعياري للمجتمع: هل هو معلوم أم مجهول؟
- حجم العينة: هل هو كبير ($ن \geq ٣٠$) أم صغير ($ن < ٣٠$)؟
- توزيع المجتمع الأصلي: هل هو معتاد أم غير معتاد؟

لذلك فإنه من الملائم أن نستعرض – في هذا الشأن – الحالات المختلفة وعددها (٢^{٣} = ٨ حالات) مستعينين في ذلك ببعض الأمثلة. وفي جميع الأمثلة التي سنوردها اعتبرنا – من قَبيل التيسير في الحسابات – أن المجتمع غير محدود. ومن ثم فإن الخطأ المعياري ($\sigma_{\bar{س}}$) والمستخدم في تلك الأمثلة يأخذ دائماً الصورة (٢ - ٢). وأما من الناحية العملية فإنه يمكن استخدام الخطأ المعياري ($\sigma_{\bar{س}}$) كما هو مُعرف في (٢ - ٣) وذلك في الحالات التي تتطلب ذلك.

١- في حالة المجتمع المعتاد:

هناك أربع حالات مختلفة لتقدير الوسط الحسابي للمجتمع وذلك إذا كان المجتمع الذي سحبت منه العينة يتبع التوزيع المعتاد. وهنا تلزم الإشارة إلى أنه عندما نقول بأن المجتمع يتبع التوزيع المعتاد فإننا نقصد بذلك أن قياسات الظاهرة أو المتغير موضع الدراسة في المجتمع لها توزيع معتاد. وهذه الحالات هي:

الحالة الأولى:

تقدير متوسط المجتمع عندما يكون التباين (σ^{2}) معلوماً وحجم العينة كبيراً (ن ≥ ٣٠):

وفي هذه الحالة يمكننا استخدام التوزيع المعتاد في عملية التقدير وذلك تطبيقاً لنظرية (١). أي أن فترة الثقة تتحدد كما يلي:

الحد الأدنى لفترة الثقة = $\bar{س} - ي_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{ن}}\right)$

الحد الأعلى لفترة الثقة = $\bar{س} + ي_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{ن}}\right)$

وذلك بمعامل ثقة يساوي (١ – α)، أي بدرجة ثقة (١ – α) × ١٠٠.

مثال (٦):

سُحبت عينة حجمها ٦٤ مفردة وذلك من مجتمع معتاد متوسطه مجهول وتباينه (σ^{2} = ٩). فإذا علم أن الوسط الحسابي للعينة هو ١٥، أوجد بفترة ثقة ٩٥٪ الوسط الحسابي للمجتمع.

الحــل:

حيث أن المجتمع يتبع توزيعاً معتاداً معلومٌ تباينه، فإنه – وبغض النظر عن حجم العينة كبيراً كان أم صغيراً – يمكن استخدام التوزيع المعتاد في عملية تقدير الوسط الحسابي للمجتمع (μ).

الخطأ المعياري = $\frac{\sigma}{\sqrt{ن}} = \frac{٣}{\sqrt{٦٤}} = ٠,٣٧٥$

وحيث أن درجة الثقة المطلوبة هي ٩٥٪، فإنه من جدول المنحنى المعتاد المعياري يمكننا استنتاج أن: $ي_{\frac{\alpha}{٢}} = ١,٩٦$

وحيث أن: $\phi(ي) = ١ - \frac{\alpha}{٢}$

$$= ١ - \frac{٠,٠٥}{٢} = ٠,٩٧٥$$

وبالبحث في جدول المساحة تحت المنحنى المعتاد المعياري نجد أن قيمة ي المناظرة للمساحة ٠,٩٧٥ هي ١,٩٦.

∴ الحد الأدنى لفترة الثقة = $\bar{س} - ي_{\frac{\alpha}{٢}} (\frac{\sigma}{\sqrt{ن}})$

$$= ١٥ - ١,٩٦ \times ٠,٣٧٥ = ١٥ - ٠,٧٣٥ = ١٤,٢٦٥$$

∴ الحد الأعلى لفترة الثقة = $\bar{س} + ي_{\frac{\alpha}{٢}} (\frac{\sigma}{\sqrt{ن}})$

$$= ١٥ + ١,٩٦ \times ٠,٣٧٥ = ١٥,٧٣٥$$

أي أن فترة الثقة ٩٥٪ للوسط الحسابي للمجتمع (م) هي [١٤,٢٦٥، ١٥,٧٣٥]. وهذا يعني أن:

$$ع(١٤,٢٦٥ \leq م \leq ١٥,٧٣٥) = ٠,٩٥$$

أي أننا لو سحبنا من المجتمع عدداً كبيراً من العينات (أو جميع العينات الممكنة ذات الحجم ٦٤ مفردة) وأنشأنا في كل مرة فترة الثقة ٩٥٪، فإنه من المتوقَّع أن تقع القيمة الحقيقية للوسط الحسابي للمجتمع داخل ٩٥٪ من فترات الثقة المكوَّنة.

مثال (٧):

أُختيرت عينة عشوائية حجمها ٣٦ مفردة وذلك من مجتمع معتاد انحرافه المعياري ٩. فإذا كان متوسط هذه العينة يساوي ٣٥، قدر الوسط الحسابي للمجتمع مستخدماً.

أ- فترة ثقة ٩٩٪ (ب)- فترة ثقة ٨٥٪

قارن بين نتائجك في الحالتيْن مع ابداء التعليق المناسب.

الحــل:

هنا أيضاً يمكن استخدام التوزيع المعتاد في عملية التقدير حيث أن المجتمع يتبع توزيعاً معتاداً معروفٌ تباينه (تطبيقاً لنظرية ١).

أ- فترة الثقة ٩٩٪:

$$\text{الخطأ المعياري} = \frac{\sigma}{\sqrt{ن}} = \frac{٩}{\sqrt{٣٦}} = ١,٥$$

وحيث أن: $\phi(ي_{\frac{\alpha}{٢}}) = ١ - \frac{\alpha}{٢}$

$$= ١ - \frac{٠,٠١}{٢} = ٠,٩٩٥$$

وبالبحث في جدول المساحة تحت المنحنى المعتاد المعياري نجد أن قيم **(ي)** المقابلة للمساحة ٠,٩٩٥ هي ٢,٥٨. أي أن: $ي_{\frac{\alpha}{٢}} = ٢,٥٨$

$\therefore$ الحد الأدنى لفترة الثقة $= \bar{س} - ي_{\frac{\alpha}{٢}} (\frac{\sigma}{\sqrt{ن}})$

$$= ٣٥ - ٢,٥٨ \times ١,٥ = ٣١,١٣$$

$\therefore$ الحد الأعلى لفترة الثقة $= \bar{س} + ي_{\frac{\alpha}{٢}} (\frac{\sigma}{\sqrt{ن}})$

$$= ٣٥ + ٢,٥٨ \times ١,٥ = ٣٨,٨٧$$

أي أن فترة الثقة ٩٩٪ للوسط الحسابي (م) هي (٣١,١٣، ٣٨,٨٧).

ب- فترة الثقة ٨٥٪:

تم حساب الخطأ المعياري في **(أ)** ووجدنا أنه يساوي ١,٥. وحيث أن:

$$\phi(ي_{\frac{\alpha}{٢}}) = ١ - \frac{\alpha}{٢} = ١ - \frac{٠,١٥}{٢} = ٠,٩٢٥$$

وبالبحث في جدول المساحة تحت المنحنى المعتاد المعياري عن قيمة **ي** المناظرة للمساحة ٠,٩٢٥ نجد أن: $ي_{٠,٠٧٥} = ١,٤٤$

$\therefore$ الحد الأدنى لفترة الثقة $= \bar{س} - ي_{\frac{\alpha}{٢}} (\frac{\sigma}{\sqrt{ن}})$

$$= ٣٥ - ١,٤٤ \times ١,٥ = ٣٢,٨٤$$

$\therefore$ الحد الأعلى لفترة الثقة $= \bar{س} + ي_{\frac{\alpha}{٢}} (\frac{\sigma}{\sqrt{ن}})$

$$= ٣٥ + ١,٤٤ \times ١,٥ = ٣٧,١٦$$

أي أن فترة الثقة ٨٥٪ للوسط الحسابي للمجتمع (م) هي [٣٢,٨٤، ٣٧,١٦].

وبمقارنة النتائج التي حصلنا عليها في أ، ب نجد أن فترة الثقة المكوَّنة في الحالة الأولى هي أكثر اتساعاً منها في الحالة الثانية. ويرجع ذلك إلى أن درجة الثقة المطلوبة في أ كانت أكبر منها في ب. وهذا يؤكد ما سبق أن أشرنا إليه وهو أن فترة الثقة تزداد بزيادة درجة الثقة المطلوبة، هذا مع ثبات حجم العينة والانحراف المعياري للمجتمع.

الحالة الثانية:

تقدير متوسط المجتمع عندما يكون التباين (σ^2) معلوماً وحجم العينة صغيراً (ن < ٣٠):

في هذه الحالة، طالما أن توزيع المجتمع هو توزيع معتاد تباينه معلوم فإنه يمكن استخدام التوزيع المعتاد في تقديرنا للوسط الحسابي للمجتمع، صغيراً كان حجم العينة أم كبيراً (نظرية ١). وعليه فإنه ليس هناك جديد يقال عن تلك الحالة أكثر مما قيل عن سابقتها. أي أن:

الحد الأدنى لفترة الثقة = $\bar{س} - ي_{\frac{\alpha}{2}} (\frac{\sigma}{\sqrt{ن}})$

الحد الأعلى لفترة الثقة = $\bar{س} + ي_{\frac{\alpha}{2}} (\frac{\sigma}{\sqrt{ن}})$

وذلك بمعامل ثقة يساوي (١ − α)، أي بدرجة ثقة (١ − α) × ١٠٠.

مثال (٨):

عينة عشوائية حجمها ٢٠ مفردة ووسطها الحسابي يساوي ٦٠، اختيرت من مجتمع معتاد انحرافه المعياري ١٥. أوجد التقدير بفترة ثقة ٩٥٪ للوسط الحسابي للمجتمع.

الحــل:

في هذا المثال، حيث أن المجتمع يتبع توزيعاً معتاداً تباينه معلوم فإننا نستخدم التوزيع المعتاد في عملية التقدير وذلك بغض النظر عن حجم العينة.

الخطأ المعياري = $\frac{\sigma}{\sqrt{ن}} = \frac{١٥}{\sqrt{٢٠}} = \frac{١٥}{٤,٤٧} = ٣,٣٦$

وحيث أن درجة الثقة المطلوبة هي ٩٥٪، فإننا نجد أن:

α = ٠,٠٥ ، $ي_{\frac{\alpha}{2}}$ = ١,٩٦

∴ الحد الأدنى لفترة الثقة = $\bar{س} - ي_{\frac{\alpha}{2}} (\frac{\sigma}{\sqrt{ن}})$

= ٦٠ − ١,٩٦ × ٣,٣٦ = ٥٣,٤١

∴ الحد الأعلى لفترة الثقة = $\bar{س} + ي_{\frac{\alpha}{2}} (\frac{\sigma}{\sqrt{ن}})$

= ٦٠ + ١,٩٦ × ٣,٣٦ = ٦٦,٥٩

أي أن فترة الثقة ٩٥٪ للوسط الحسابي للمجتمع (م) هي [٥٣,٤١، ٦٦,٥٩].

مثال (٩):

المطلوب هو حل المثال السابق وذلك بافتراض أن حجم المجتمع المسحوب منه العينة المذكورة هو ١٠٠ مفردة. قارن بين نتائجك في هذه الحالة وبين ما حصلت عليه من نتائج في المثال السابق مع إبداء التعليق المناسب.

الحـل:

لا تختلف طريقة حل هذا المثال عن الطريقة المتبعة في المثال السابق إلا في كيفية حساب الخطأ المعياري. ففي المثال السابق لم يذكر شيء عن حجم المجتمع وبالتالي اعتبرناه مجتمعاً غير محدود، وقمنا بحساب الخطـأ المعياري بناءً على ذلك. وأما في هذا المثال وحيث يبلغ حجم المجتمع ١٠٠ مفردة فإننا نلاحظ أن: ه > ٠,٠٥ ن

وعليه يمكن حساب الخطأ المعياري على الصورة:

$$\sigma_{\bar{\text{س}}} = \frac{\sigma}{\sqrt{\text{ه}}}\sqrt{\frac{\text{ن} - \text{ه}}{\text{ن} - ١}}$$

$$= \frac{١٥}{\sqrt{٢٠}}\sqrt{\frac{١٠٠ - ٢٠}{١٠٠ - ١}} = ٣,٣٥ \times ٠,٩ = ٣,٠٢$$

الحد الأدنى لفترة الثقة = ٦٠ – ١,٩٦ × ٣,٠٢ = ٥٤,١

الحد الأعلى لفترة الثقة = ٦٠ + ١,٩٦ × ٣,٠٢ = ٦٥,٩

وبمقارنة نتائج هذا المثال بالنتائج التي حصلنا عليها في المثال السابق (رقم ٨) لوجدنا أنه نظراً لأن درجة الثقة في المثالين واحدة وهي ٩٥٪ فإنه يمكننا القول بأن نتائج هذا المثال أكثر دقة من سابقتها وذلك لأن طول فترة الثقة في هذا المثال هو أقل من نظيره في مثال (٨). ويرجع هذا إلى أن العينـــة (ه = ٢٠) كانت أكثر قرباً أو تمثيلاً لواقع المجتمع في هذا المثال وذلك لأن المجتمع كـان محـدوداً (ن = ١٠٠) في الوقت الذي اعتُبِر فيه هذا المجتمع غير محدود (ن ← ∞) في مثال (٨).

الحالة الثالثة:

تقدير متوسط المجتمع عندما يكون التباين (σ^2) مجهولاً وحجم العينة كبيراً (ه ≥ ٣٠):

في هذه الحالة نجد أن المجتمع يتبع التوزيع المعتاد كما أن حجم العينة كبير، وهي شروط تتوافق مع استخدام التوزيع المعتاد في تقدير متوسط المجتمع (م). هذا ونظراً لأن الانحراف المعياري للمجتمع غير معروف، كما أن حجم العينة كبير فإننا نستخدم الانحراف المعياري المحسوب من بيانات العينة (ع) كتقدير للانحراف المعياري للمجتمع (σ)، حيث تعرَّف ع في هذه الحالة كما يلي:

$$ع = \sqrt{\frac{1}{ن - 1}\left[\sum س^{2} - \frac{\left(\sum س\right)^{2}}{ن - 1}\right]}$$

وذلك لأن تباين العينة: $ع^{2} = \frac{1}{ن - 1}\left[\sum س^{2} - \frac{\left(\sum س\right)^{2}}{ن - 1}\right]$

هو تقدير غير متحيز لتباين المجتمع σ^{2}

ومن غير الملائم استخدام مربع الصيغة (٢ - ٧) للانحراف المعياري في إيجاد تقدير لتباين المجتمع وذلك لأن المقدِّر في هذه الحالة سوف يكون متحيزاً.

وهنا يمكننا أن نطبق نظرية الحد المركزية في إيجاد فترة الثقة لمتوسط المجتمع، حيث نجد أن الحد الأدنى والحد الأعلى للثقة في هذه الحالة يتحددان كما يلي:

$$\text{الحد الأدنى لفترة الثقة} = \bar{س} - ي_{\frac{\alpha}{2}} \left(\frac{ع}{\sqrt{ن}}\right)$$

$$\text{الحد الأعلى لفترة الثقة} = \bar{س} + ي_{\frac{\alpha}{2}} \left(\frac{ع}{\sqrt{ن}}\right)$$

وذلك بمعامل ثقة يساوي (١ - α) أي بدرجة ثقة (١ - α) × ١٠٠، وحيث ع معرَّفة بالصيغة (٢ - ٤).

مثال (١٠):

سحبت عينة حجمها ٢٢٥ طالباً من مجتمع طلاب الفرقة الثانية بإحدى كليات التجارة. فإذا وجدنا أن الوسط الحسابي والانحراف المعياري لدرجات طلاب العينة في مادة الإحصاء هما على الترتيب ١٦ درجة و٥ درجات. وإذا كان من المعروف أن توزيع درجات طلاب الفرقة الثانية في مادة الإحصاء يتبع توزيعاً معتاداً متوسطه وتباينه مجهولان. أوجد - بفترة ثقة ٩٩٪ - تقديراً للوسط الحسابي لدرجات الطلاب في الفرقة الثانية.

الحــل:

نجد في هذا المثال أن المجتمع يتبع توزيعاً معتاداً غير معروف وسطه الحسابي وانحرافه المعياري. وحيث أن حجم العينة كبيراً ($ن \geq ٣٠$) فإنه تطبيقاً لنظرية الحد المركزية يمكننا استخدام التوزيع المعتاد مع استخدام ع كما هي معرَّفة في (٢ - ٤) كتقدير للانحراف المعياري للمجتمع σ وذلك في تقدير الوسط الحسابي للمجتمع.

$$\text{الخطأ المعياري} = \frac{ع}{\sqrt{ن}} = \frac{٥}{\sqrt{٢٢٥}} = ٠,٣٣$$

وحيث أن درجة الثقة المطلوبة هي ٩٩٪، فإننا نجد أن:

$$\alpha = ٠,٠١ \quad ، \quad ي_{\frac{\alpha}{٢}} = ٢,٥٨$$

$$\therefore \text{الحد الأدنى لفترة الثقة} = \bar{س} - ي_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$$

$$= ١٦ - ٢,٥٨ \times ٠,٣٣ = ١٥,١٥ \text{ درجة}$$

$$\therefore \text{الحد الأعلى لفترة الثقة} = \bar{س} + ي_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$$

$$= ١٦ + ٢,٥٨ \times ٠,٣٣ = ١٦,٨٥ \text{ درجة}$$

أي أن فترة الثقة ٩٩٪ لمتوسط درجات الإحصاء لطلاب الفرقة الثانية هي [١٥,١٥، ١٦,٨٥].

الحالة الرابعة:

تقدير متوسط المجتمع عندما يكون التباين ($\sigma^٢$) مجهولاً وحجم العينة صغيراً ($ن < ٣٠$):

سبق الإشارة في الفصل الأول من هذا الكتاب إلى أن توزيع ت يستخدم في حالة العينات الصغيرة ($ن < ٣٠$). وفي واقع الأمر فإن تطبيق توزيع ت في حالة تقدير متوسط مجتمع يتطلب عدة شروط نوجزها فيما يلي:

- أن يكون توزيع المجتمع طبيعياً (معتاداً).
- تباين المجتمع ($\sigma^٢$) مجهول.
- حجم العينة صغير ($ن < ٣٠$).

وهذه الشروط هي واقع حالتنا هذه، ومن ثم فإن توزيع **ت** هو التوزيع الملائم للاستخدام في تقدير متوسط المجتمع.

$$\text{وحيث أن: } ت = \frac{\bar{س} - م}{\frac{ع}{\sqrt{ن}}}$$

هي متغير يتبع توزيع ت بدرجات حرية $ن - ١$، فإنه وعلى غرار ما تم في حالة التوزيع المعتاد يمكننا استنتاج أن فترة الثقة في هذه الحالة تأخذ الصورة التالية:

$$\text{الحد الأدنى لفترة الثقة} = \bar{س} - ت_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$$

$$\text{الحد الأعلى لفترة الثقة} = \bar{س} + ت_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$$

وذلك بمعامل ثقة يساوي ($١ - \alpha$) أي بدرجة ثقة ($١ - \alpha$) $\times$ ١٠٠، وحيث ع معرَّفة بالصيغة (٢ - ٤).

وفي مثل تلك الحالات فإننا نستخدم جدول المساحة تحت منحنى ت والملحق بنهاية هذا الكتاب (الملحق ٢) وذلك بدلاً من جدول المساحة تحت المنحنى المعتاد المعياري والمستخَدم في الحالات السابقة. مع ملاحظة أن $ت_{\frac{\alpha}{٢}}$ تتشابه في تعريفها مع $ي_{\frac{\alpha}{٢}}$.

وهنا نود أن نذكِّر القارئ بأن هناك اختلافاً جوهرياً بين التوزيع المعتاد وتوزيع ت وذلك فيما يتعلق بالكشف عن المساحات المناظرة للقيم المعيارية $ي_{\frac{\alpha}{٢}}$، $ت_{\frac{\alpha}{٢}}$. ففي التوزيع المعتاد نجد أنه بالنسبة لمستوى معنوية معين (α) توجد قيمة واحدة مناظرة لـ $ي_{\frac{\alpha}{٢}}$. وأما في توزيع ت نجد أن هناك أكثر من قيمة للمساحة (الاحتمال) وجميعها مناظر لـ $ت_{\frac{\alpha}{٢}}$. ويرجع هذا الاختلاف إلى ما سبق أن أشرنا إليه وهو أن جميع التوزيعات المعتادة مهما اختلفت أوساطها الحسابية وانحرافاتها المعيارية يمكن تحويلها جميعاً إلى توزيع معتاد معياري واحد وسطه الحسابي يساوي الصفر وتباينه يساوي الوحدة (أي الواحد الصحيح). هذا في الوقت الذي نجد فيه أن توزيع **ت** ما هـو إلا عبارة عن مجموعة من المنحنيات التي يعتمد شكل كل منها على حجم العينة ں، راجع الشكل (١ - ٧). لذلك نجد أن قيمة $ت_{\frac{\alpha}{٢}}$ لا تعتمد على قيم α وحدها وإنما تعتمد أيضاً على المقدار (ں ــ ١) وهو ما يسمى بـ "درجات الحرية". أي أن:

درجات الحرية = حجم العينة ــ ١

= ں ــ ١

هذا وللكشف عن قيم ت المناظرة لدرجات الحرية المختلفة في ظل مستوى معنوية محدد يمكننا اتباع الأسلوب المُشار إليه من قبل في الفصل الأول من هذا الكتاب.

ومن باب الإيضاح، وعلى سبيل المثال، نجد أن:

$ت_{(١٥، ٠,٠٢٥)} = ٢,١٣١$ ، $ت_{(٢٥، ٠,٠١)} = ٢,٤٨٥$

$ت_{(\infty، ٠,٠٢٥)} = ١,٩٦$

وفيما يلي بعض الأمثلة التي توضح كيفية استخدام توزيع ت في إيجاد تقدير لمتوسط المجتمع:

مثال (١١):

أُخذت عينة عشوائية حجمها ٢٠ مفرد بمتوسط ٦٠ وانحراف معياري ١٥ وذلك من مجتمع يتبع التوزيع المعتاد. المطلوب:

أ- أوجد الوسط الحسابي للمجتمع وذلك بفترة ثقة ٩٥٪.

ب- فسِّر نتائجك في أ.

ج- قارن نتائجك في أ بنتائج مثال (٨) مع إبداء التعليق المناسب.

الحـــل:

أ- حيث أن المجتمع يتبع توزيعاً معتاداً مجهول التباين، كما أن حجم العينة صغير ($ن < ٣٠$) فإن التوزيع الملائم لتقدير الوسط الحسابي للمجتمع هو توزيع ت. أي أن:

$$\text{الخطأ المعياري} = \frac{ع}{\sqrt{ن}} = \frac{١٥}{\sqrt{٢٠}} = ٣{,}٣٦$$

وحيث أن درجة الثقة المطلوبة هي ٩٥٪، فإننا نجد أن:

$$\alpha = ٠{,}٠٥ \quad، \quad ت_{(ن-١)، \frac{\alpha}{٢}} = ت_{١٩، ٠{,}٠٢٥} = ٢{,}٠٩٣$$

$$\therefore \text{الحد الأدنى لفترة الثقة} = \bar{س} - ت_{\left[(ن-١)، \frac{\alpha}{٢}\right]} \left(\frac{ع}{\sqrt{ن}}\right)$$

$$= ٦٠ - ٢{,}٠٩٣ \times ٣{,}٣٦ = ٥٢{,}٩٧$$

$$\therefore \text{الحد الأعلى لفترة الثقة} = \bar{س} + ت_{\left[(ن-١)، \frac{\alpha}{٢}\right]} \left(\frac{ع}{\sqrt{ن}}\right)$$

$$= ٦٠ - ٢{,}٠٩٣ \times ٣{,}٣٦ = ٦٧{,}٠٣$$

أي أن الوسط الحسابي للمجتمع ($م$) يقع بين حدي الثقة ٥٢,٩٧ و ٦٧,٠٣ وذلك بدرجة ثقة ٩٥٪.

ب- يمكن تفسير النتائج في أ بأننا لو أخذنا من المجتمع عينات عشوائية متكرِّرة حجم كل منها ٢٠ مفردة، وأنشأنا فترة الثقة ٩٥٪ لمتوسط المجتمع في كل مرة نسحب فيها عينة. فإنه من المتوقَّع أن ٩٥٪ من فترات الثقة المكوَّنة سوف تضم القيمة الحقيقية لمتوسط المجتمع غير المعروف.

ج- بمقارنة فترة الثقة المكوَّنة في هذا المثال (باستخدام توزيع ت) بتلك الفترة المكوَّنة في مثال (٨) المستخدم فيها التوزيع المعتاد يلاحظ أن كلاً من $\bar{س}$، $ن$ واحدة في المثاليْن، كما أن σ في المثال (٨) تساوي $ع$ في مثال (١١)، فإننا نستنتج - وكما هو متوقَّع - أن فترة الثقة في حالة توزيع ت أكبر منها في حالة التوزيع المعتاد. ويرجع ذلك إلى أن توزيع ت هو أكبر تشتتاً من التوزيع المعتاد عندما يكون حجم العينة صغيراً ($ن < ٣٠$). ومع ذلك فإن الفرق بين طولي فترتي الثقة ليس كبيراً إذ أنه عند $ن = ٢٠$ فإن توزيع ت يقترب إلى حد ما من التوزيع المعتاد. ومن المتوقَّع أن يقل هذا الفارق أكثر فأكثر كلما زاد حجم العينة $ن$، والعكس صحيح.

مثال (١٢):

أخذت عينة عشوائية حجمها ٨ مفردات وذلك من مجتمع معتاد، فإذا كانت مشاهدات العينة هي: ٤ ، ٨ ، ٧ ، ٥ ، ٧، ٦، ٢ ، ٩. أوجد فترة الثقة ٩٩٪ للوسط الحسابي للمجتمع المسحوب منه العينة.

الحـــل:

حيث أن المجتمع يتبع توزيعاً معتادا تباينه غير معلوم، كما أن حجم العينة صغير ($ن < ٣٠$) فإن توزيع ت هو التوزيع الملائم لإيجاد فترة الثقة لمتوسط المجتمع (μ). وحيث أن درجة الثقة المطلوبة هي ٩٩٪ فإننا نجد أن:

$$١ - \alpha = ٠,٩٩ \quad ، \quad \alpha = ٠,٠١$$

$$\therefore ت_{(ن-١)، \frac{\alpha}{٢}} = ت_{٧، ٠,٠٠٥} = ٣,٤٩٩$$

وحيث أن الانحراف المعياري للمجتمع غير معلوم فإننا سوف نستخدم الانحراف المعياري للعينة (ع) وذلك كتقدير للانحراف المعياري للمجتمع σ وحيث ع تعطى بالصورة:

$$ع = \sqrt{\frac{١}{ن-١}\left[\sum س^٢ - \frac{(\sum س)^٢}{ن}\right]}$$

وبحساب مجموع مشاهدات العينة ومجموع مربعاتها فإننا نحصل على:

$$\sum س = ٤٨ \quad ، \quad \sum س^٢ = ٣٢٤$$

أي أن:

$$ع = \sqrt{\frac{١}{٧}\left[٣٢٤ - \frac{(٤٨)^٢}{٨}\right]} = \sqrt{٥,١٤} = ٢,٢٧$$

$$\therefore \text{الخطأ المعياري} = \frac{ع}{\sqrt{ن}} = \frac{٢,٢٧}{\sqrt{٨}} = \frac{٢,٢٧}{٢,٨٣} = ٠,٨$$

$$\text{الوسط الحسابي } (\bar{س}) = \frac{\sum س}{ن} = \frac{٤٨}{٨} = ٦$$

$$\therefore \text{الحد الأدنى لفترة الثقة} = \bar{س} - ت_{\left[(ن-١)، \frac{\alpha}{٢}\right]} \frac{ع}{\sqrt{ن}}$$

$$= ٦ - ٣,٤٩٩ \times ٠,٨ = ٣,٢$$

$$\therefore \text{الحد الأعلى لفترة الثقة} = \bar{س} + ت_{\left[(ن-١)، \frac{\alpha}{٢}\right]} \frac{ع}{\sqrt{ن}}$$

$$= ٦ - ٣,٤٩٩ \times ٠,٨ = ٨,٨$$

أي أن الوسط الحسابي يتراوح ما بين ٣,٢ ، ٨,٨ وذلك بدرجة ثقة ٩٩٪.

٢- في حالة المجتمع غير المعتاد:

هناك أيضاً أربع حالات مختلفة لتقدير الوسط الحسابي للمجتمع الذي يتبع توزيعاً غير معتاد. وفيما يلي نستعرض تلك الحالات مستعينيــن في ذلك ببعض الأمثلة. علماً بأنه ما لم يُذكر صراحة في المثال بأن المجتمع يتبع التوزيع المعتاد فإننا سنعتبر أن المجتمع يتبع توزيعاً غير معتاد.

الحالة الأولى:

تقدير متوسط المجتمع عندما يكون التباين ($\sigma^٢$) معلوماً وحجم العينة كبيراً ($ن \geq ٣٠$):

إذا أردنا تقدير متوسط المجتمع في هذه الحال فإنه يمكننا استخدام التوزيع المعتاد وذلك تطبيقاً لنظرية الحد المركزية (نظرية ٢) والتي تفيد بأنه إذا كان تباين المجتمع معلوماً وحجم العينة كبيراً كبراً كافياً ($ن \geq ٣٠$) فإن توزيع المعاينة للمتوسط يتبع توزيعاً معتاداً وذلك بغض النظر عن شكل المجتمع المسحوب منه العينة. أي أن:

$$\text{الحد الأدنى لفترة الثقة} = \bar{س} - ي_{\frac{\alpha}{٢}} \left(\frac{\sigma}{\sqrt{ن}}\right)$$

$$\text{الحد الأعلى لفترة الثقة} = \bar{س} + ي_{\frac{\alpha}{٢}} \left(\frac{\sigma}{\sqrt{ن}}\right)$$

وذلك بمعامل ثقة ($١ - \alpha$) أي بدرجة ثقة ($١ - \alpha$) × ١٠٠.

مثال (١٣):

سُحبت عينة عشوائية حجمها ١٠٠ مفردة من مجتمع تباينه ١٦. فإذا كان الوسط الحسابي للعينة هو ٢٥، احسب فترة الثقة ٩٥٪ لمتوسط المجتمع (μ).

الحـــل:

حيث أن حجم العينة كبير ($ن = ١٠٠ > ٣٠$) والانحراف المعياري للمجتمع معلوم فإنه يمكن استخدام التوزيع المعتاد في تقدير متوسط المجتمع أياً كان توزيع المجتمع المسحوب منه العينة (تطبيقاً لنظرية الحد المركزية). وحيث أن درجة الثقة المطلوبة هي٩٥٪ فإننا نجد أن:

$$\alpha = ٠,٠٥ \quad ، \quad ي_{\frac{\alpha}{٢}} = ١,٩٦$$

$$\text{الخطأ المعياري} = \frac{\sigma}{\sqrt{ن}} = \frac{٤}{\sqrt{١٠٠}} = ٠,٤$$

$$\therefore \text{الحد الأدنى لفترة الثقة} = \bar{س} - ي_{\frac{\alpha}{٢}} \frac{\sigma}{\sqrt{ن}}$$

$$= ٢٥ - ١,٩٦ \times ٠,٤ = ٢٤,٢٢$$

∴ الحد الأعلى لفترة الثقة = $\bar{س} + ي_{\frac{\alpha}{٢}} \frac{\sigma}{\sqrt{ن}}$

= ٢٥ + ١,٩٦ × ٠,٤ = ٢٥,٧٨

أي أن فترة الثقة ٩٥٪ لمتوسط المجتمع (م) هي [٢٤,٢٢ ، ٢٥,٧٨].

الحالة الثانية:

تقدير متوسط المجتمع عندما يكون التباين (σ^2) مجهولاً وحجم العينة كبيراً (ن ≥ ٣٠):

نظراً لكبر حجم العينة فإنه يمكن تطبيق نظرية الحد المركزية (نظرية ٢) وذلك بعد استخدام الانحراف المعياري للعينة (ع) كتقدير للانحراف المعياري للمجتمع σ. أي أن الحديْن الأدنى والأعلى يعرفان كما يلي:

الحد الأدنى لفترة الثقة = $\bar{س} - ي_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$

الحد الأعلى لفترة الثقة = $\bar{س} + ي_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$

وذلك بمعامل ثقة يساوي (١ - α) أي بدرجة ثقة (١ - α) × ١٠٠، وحيث ع هي الانحراف المعياري للعينة والمعطى بالصورة (٢ - ٤).

مثال (١٤):

أخذت عينة عشوائية حجمها ٢٥٦ مفردة بمتوسط ١٢٠ وانحراف معياري ٤٠ من مجتمع معروف عنه أنه لا يتبع التوزيع المعتاد. أوجد التقدير بفترة ثقة ٩٠٪ للوسط الحسابي للمجتمع (م).

الحـــل:

في هذا المثال نجد أن حجم العينة كبير (ن > ٣٠) حيث يمكن تطبيق نظرية الحد المركزية لاستخدام التوزيع المعتاد في عملية التقدير.

وحيث أن درجة الثقة المطلوبة هي ٩٠٪ فإنه من جداول المساحة تحت المنحنى المعتاد المعياري نجد أن: $ي_{\frac{\alpha}{٢}}$ = ١,٦٤٥

الخطأ المعياري = $\frac{ع}{\sqrt{ن}} = \frac{٤٠}{\sqrt{٢٥٦}}$ = ٢,٥

∴ الحد الأدنى لفترة الثقة = $\bar{س} - ي_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$

= ١٢٠ – ١,٦٤٥ × ٢,٥ = ١١٥,٨٩ ،

∴ الحد الأعلى لفترة الثقة = $\bar{س} + ي_{\frac{\alpha}{2}} (\frac{ع}{\sqrt{ن}})$

$= ١٢٠ + ١,٦٤٥ \times ٢,٥ = ١٢٤,١١$

أي أن متوسط المجتمع (م) يتراوح ما بين ١١٥,٨٩ ، ١٢٤,١١ وذلك بدرجة ثقة ٩٠٪.

الحالة الثالثة:

تقدير متوسط المجتمع عندما يكون التباين ($\sigma^٢$) معلوماً وحجم العينة صغيراً ($ن < ٣٠$):

حيث أن المجتمع الذي أُخذت منه العينة لا يتبع التوزيع المعتاد كما أن حجم العينة صغير، فإنه لا يمكننا - من الناحية النظرية - استخدام التوزيع المعتاد أو توزيع ت في تقدير الوسط الحسابي للمجتمع (م)، وذلك لأن هذا الوضع لا يتوافق مع شروط تطبيق أي من نظرية (١) أو نظرية (٢). ومع ذلك فإن النظرية التي قدمها عالم الرياضيات الروسي تشبتشيف يمكن أن تفيد في هذا الشأن. وفيما يلي تعريف النظرية وكيفية استخدامها في تقدير متوسط المجتمع (م) في مثل تلك الحالات:

نظرية تشبتشيف: Chebyshev's Theorem

تقول النظرية: أنه بصرف النظر عن شكل التوزيع فإن نسبة المشاهدات التي لا تبعد عن الوسط الحسابي بأكثر من ل انحراف معياري هي على الأقل $١ - \frac{١}{ل^٢}$، حيث $ل \geq ١$.

وإذا أردنا تطبيق هذه النظرية على توزيع المعاينة للمتوسط فإننا نستطيع أن نقول إن احتمال أن متوسط العينة ($\bar{س}$) لا يبعد عن الوسط الحسابي للمجتمع (م) بأكثر من ل خطأ معياري هو:

$$ع(|\bar{س} - م| \leq ل \frac{\sigma}{\sqrt{ن}}) \geq ١ - \frac{١}{ل^٢} \qquad (٢ - ٨)$$

هذا وتُعرف الصيغة (٢ - ٨) بمتباينة تشبتشيف، والتي تقوم على افتراض ان الانحراف المعياري للمجتمع (σ) معروف.

متباينة تشبتشيف وتقدير متوسط المجتمع (م) بفترة ثقة:

في حالة استخدامنا لمتباينة تشبتشيف في تقدير فترة ثقة لمتوسط المجتمع (م) فإننا نوجد قيم ل التي تجعل المقدار ($١ - \frac{١}{ل^٢}$) مساوٍ لمعامل الثقة المطلوب. وبعد تحديد قيمة ل فإن الحد الأدنى والحد الأعلى للثقة يمكن إيجادهما على النحو التالي:

حيث أن:

$$\text{ع}\left(|\bar{\text{س}} - \text{م}| \le \text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}}\right) \ge ١ - \frac{١}{\text{ل}^{٢}}$$

فإنه يمكن كتابة الطرف الأيمن من متباينة تشبتشيف على الصورة التالية:

$$\text{ع}\left(|\bar{\text{س}} - \text{م}| \le \text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}}\right) = \text{ع}\left(-\text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}} \le \bar{\text{س}} - \text{م} \le \text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}}\right)$$

$$= \text{ع}\left(\bar{\text{س}} - \text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}} \le \text{م} \le \bar{\text{س}} + \text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}}\right)$$

أي أن:

$$\text{ع}\left(\bar{\text{س}} - \text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}} \le \text{م} \le \bar{\text{س}} + \text{ل}\,\frac{\sigma}{\sqrt{\text{ن}}}\right) \ge ١ - \frac{١}{\text{ل}^{٢}}$$

وهذا يعنى أن الحد الأدنى والحد الأعلى للثقة يتحددان كما يلي:

الحد الأدنى لفترة الثقة $= \bar{\text{س}} - \text{ل}\left(\frac{\sigma}{\sqrt{\text{ن}}}\right)$

الحد الأعلى لفترة الثقة $= \bar{\text{س}} + \text{ل}\left(\frac{\sigma}{\sqrt{\text{ن}}}\right)$

حيث: ل تتحدد قيمتها بحيث: $١ - \frac{١}{\text{ل}^{٢}}$ = معامل الثقة المطلوب.

أي أن: $١ - \frac{١}{\text{ل}^{٢}} = ١ - \alpha$ $\qquad \therefore\ \text{ل} = \frac{١}{\alpha}$

مثال (١٥):

سُحبت عينة حجمها ٢٠ مفردة من مجتمع انحرافه المعياري ١٥. فإذا كان معروفاً لنا أن المجتمع لا يتبع التوزيع المعتاد وأن متوسط العينة يساوي ٦٠. أوجد بفترة ثقة ٩٥٪ الوسط الحسابي للمجتمع. قارن نتائجك بنتائج المثاليْن (٨)، (١١) مع إبداء التعليق المناسب.

الحـــل:

حيــث أن المجتمــع لا يتبـــع التوزيـــع المعتــاد وحجم العينـــة صغــير ($\text{ن} < ٣٠$) فإننا من الناحية النظرية لا نستطيع استخدام التوزيع المعتاد أو توزيع ت في تقدير الوسط الحسابي للمجتمع. ولكن يمكننا استخدام نظرية تشبتشيف لهذا الغرض وذلك على النحو التالي:

حيث أن درجة الثقة المطلوبة هي ٩٥٪.

وبمساواة المقدار $(1 - \frac{1}{ل^2})$ بالقيمة ٠,٩٥ نحصل على:

$$1 - \frac{1}{ل^2} = 0.95 \quad ، \quad \therefore \frac{1}{ل^2} = 0.05$$

أي أن: $ل^2 = 20 \quad \therefore ل = \sqrt{20} = 4.47$

$$\text{الخطأ المعياري} = \frac{\sigma}{\sqrt{ن}} = \frac{15}{\sqrt{20}} = 3.36$$

$$\text{الحد الأدنى لفترة الثقة} = \bar{س} - ل(\frac{\sigma}{\sqrt{ن}})$$

$$= 60 - 4.47 \times 3.36 = 44.98$$

$$\text{الحد الأعلى لفترة الثقة} = \bar{س} + ل(\frac{\sigma}{\sqrt{ن}})$$

$$= 60 + 4.47 \times 3.36 = 75.02$$

أي أن فترة الثقة ٩٥٪ لمتوسط المجتمع (μ) هي [٤٤,٩٨ ، ٧٥,٠٢].

يلاحَظ هنا أيضاً أن كلاً من $\bar{س}$، $ن$، σ (أو ع) واحدة في الأمثلة الثلاثة، وأن الاختلاف هنا هو أن توزيع المجتمع غير معتاد كما أنه لا يتبع توزيع ت.

وبمقارنة فترة الثقة المكونة في هذا المثال بفترة الثقة المكوَّنة في كلٍّ من المثاليْن (٨)، (١١) نجد أن:

فترة الثقة		الأسلوب المستخدم في تكوين فترة الثقة
الحد الأدنى	**الحد الأعلى**	
٥٣,٤١	٦٦,٥٩	التوزيع المعتـــاد
٥٢,٧٩	٦٧,٠٣	توزيـــــــع ت
٤٤,٩٨	٧٥,٠٢	نظرية تشبتشيف

وبذلك يمكننا بسهولة استنتاج أن نظرية تشبتشيف تعطي فترة ثقة أوسع بكثير من فترة الثقة المكوَّنة باستخدام التوزيع المعتاد أو توزيع ت وهو ما يعنى درجة دقة أقل في تقدير متوسط المجتمع (يُلاحَظ أن درجة الثقة واحدة في الحالات الثلاثة). لذلك فإنه من النادر استخدام هذه النظرية في إيجاد فترات ثقة لمتوسط المجتمع غير المعروف. ومع ذلك فإنها تمثل الاختيار الوحيد في مثل هذه الحالة إذا لم نتمكن من زيادة حجم العينة إلى ٣٠ على الأقل حتى نتمكن من استخدام التوزيع المعتاد.

الحالة الرابعة:

تقدير متوسط المجتمع عندما يكون التباين (σ^2) مجهولاً وحجم العينة صغيراً (ن < ٣٠):

نجد أيضاً في هذه الحالة أن المجتمع لا يتبع التوزيع المعتاد وحجم العينة صغير، زد على ذلك أن تباين المجتمع مجهول. الأمر الذي لا يمكن معه استخدام التوزيع المعتاد أو توزيع ت في إيجاد فترة ثقة لمتوسط المجتمع م. وفي ظل وضع كهذا ليس أمامنا إلا استخدام نظرية تشبتشيف بما يشوبها من محاذير وشكوك في حالة كتلك. إذ أنه في هذه الحالة سوف نستخدم الانحراف المعياري للعينة (ع) بدلاً من الانحراف المعياري للمجتمع (σ). ومن المعروف أن استخدام الانحراف المعياري المحسوب من عينة صغيرة هو أمر تنقصه الدقة وقد يعطى نتائج مضللة.

وباستخدام متباينة تشبتشيف فإن الحد الأدنى والحد الأعلى للثقة يتحددان كما يلي:

$$\text{الحد الأدنى لفترة الثقة} = \bar{س} - ل(\frac{ع}{\sqrt{ن}})$$

$$\text{الحد الأعلى لفترة الثقة} = \bar{س} + ل(\frac{ع}{\sqrt{ن}})$$

حيث: ل تتحدد قيمتها بحيث: $١ - \frac{١}{ل^٢}$ = معامل الثقة المطلوب.

أي أن: $ل = \frac{١}{\alpha}$ ، ع كما هي معرَّفة بالصورة (٢ - ٤).

مثال (١٦):

أُخذت عينة عشوائية حجمها ١٦ مفردة بمتوسط ٣٥ وانحراف معياري ٥ وذلك من مجتمع لا يتبع التوزيع المعتاد. أوجد فترة الثقة ٩٩٪ للوسط الحسابي للمجتمع.

الحـــل:

حيث أن المجتمع لا يتبع التوزيع المعتاد وحجم العينة صغير، كما أن الانحراف المعياري للمجتمع مجهول، فإنه - من الناحية النظرية - لا يمكننا استخدام التوزيع المعتاد أو توزيع ت في إيجاد فترة الثقة لمتوسط المجتمع م. والخيار الوحيد أمامنا هو استخدام نظرية تشبتشيف وذلك على النحو التالي:

حيث أن درجة الثقة المطلوبة هي ٩٩٪، وبمساواة المقدار $(١ - \frac{١}{ل^٢})$ بالقيمة ٠,٩٩ نحصل على ما يلي:

$$١ - \frac{١}{ل^٢} = ٠,٩٩ \quad ، \quad \therefore ل = ١٠$$

الخطأ المعياري = $\frac{ع}{\sqrt{ن}}$ = $\frac{٥}{\sqrt{١٦}}$ = ١,٢٥

∴ الحد الأدنى لفترة الثقة = $\bar{س}$ – ل $\frac{ع}{\sqrt{ن}}$

= ٣٥ – ١٠ × ١,٢٥ = ٢٢,٥

∴ الحد الأعلى لفترة الثقة = $\bar{س}$ + ل $\frac{ع}{\sqrt{ن}}$

= ٣٥ + ١٠ × ١,٢٥ = ٤٧,٥

أي أن فترة الثقة ٩٩٪ لمتوسط المجتمع (م) هي [٢٢,٥ ، ٤٧,٥]. ومرةً أخرى يلاحَظ أن نظرية تشبتشيف تعطى فترات ثقة واسعة.

هناك نقطة هامة يجب التأكيد عليها وهي أنه في مثل تلك الحالات لا نستطيع – من الناحية النظرية – أن نستخدم التوزيع المعتاد أو توزيع ت في تقدير فترة الثقة لمتوسط المجتمع المجهول (م). ولا يكون أمامنا إلا أن نستخدم نظرية تشبتشيف أو نزيد حجم العينة العشوائية إلى ٣٠ على الأقل (وهذا هو الافضل) وذلك حتى نتمكن من استخدام التوزيع المعتاد. وفي واقع الأمر فإن استخدام توزيع ت بدلاً من نظرية تشبتشيف في هذه الحالات قد لا يزيد الوضع سوءًا.

وبذلك نكون قد استعرضنا الحالات المختلفة لإيجاد فترة الثقة للوسط الحسابي للمجتمع (م) موضحين التوزيع أو الأسلوب الإحصائي الملائم في كل حالة. ونظراً لتعدد الحالات واختلافها لعله من المفيد أن نلخص تلك الحالات بالشكل الذي ييسر على القارئ سرعة فهمها واستيعابها وذلك على النحو الموضح في الجدول (٢ - ٤).

الجدول (٢ - ٤)
الحالات المختلفة لإيجاد فترات الثقة لمتوسط المجتمع (م)

الحالة	بيان الحالة			التوزيع أو الاسلوب المستخدم	فترة الثقة	مبررات استخدام التوزيع أو الأسلوب الإحصائي
	توزيع المجتمع	حجم العينة (ن)	تباين المجتمع (σ^2)			
١	معتاد	كبير (ن ≥ ٣٠)	معلوم	التوزيع المعتاد	$\bar{س} \pm ي_{\frac{\alpha}{2}} (\frac{\sigma}{\sqrt{ن}})$	تطبيقاً لنظرية (١) *
٢	معتاد	صغير (ن < ٣٠)	معلوم	التوزيع المعتاد	$\bar{س} \pm ي_{\frac{\alpha}{2}} (\frac{\sigma}{\sqrt{ن}})$	تطبيقاً لنظرية (١) *
٣	معتاد	كبير (ن ≥ ٣٠)	مجهول	التوزيع المعتاد	$\bar{س} \pm ي_{\frac{\alpha}{2}} (\frac{ع}{\sqrt{ن}})$	تطبيقاً لنظرية (١) مع استخدام ع كتقدير لـ σ وذلك نظراً لكبر حجم العينة
٤	معتاد	صغير (ن < ٣٠)	مجهول	توزيع ت	$\bar{س} \pm ت_{\frac{\alpha}{2}} (\frac{ع}{\sqrt{ن}})$	بيان الحالة يمثل تماماً شروط استخدام توزيع ت.
٥	غير معتاد	كبير (ن ≥ ٣٠)	معلوم	التوزيع المعتاد	$\bar{س} \pm ي_{\frac{\alpha}{2}} (\frac{\sigma}{\sqrt{ن}})$	تطبيقاً لنظرية الحد** المركزية (نظرية ٢)
٦	غير معتاد	كبير (ن ≥ ٣٠)	مجهول	التوزيع المعتاد	$\bar{س} \pm ي_{\frac{\alpha}{2}} (\frac{ع}{\sqrt{ن}})$	تطبيقاً لنظرية (٢) مع استخدام ع كتقدير لـ σ وذلك نظراً لكبر حجم العينة
٧	غير معتاد	صغير (ن < ٣٠)	معلوم	متباينة تشبتشيف	$\bar{س} \pm ل (\frac{\sigma}{\sqrt{ن}})$ حيث: $ل = \frac{1}{\alpha}$	بيان الحالة يستحيل معه - نظرياً - استخدام التوزيع المعتاد أو توزيع ت.
٨	غير معتاد	صغير (ن < ٣٠)	مجهول	متباينة تشبتشيف	$\bar{س} \pm ل (\frac{ع}{\sqrt{ن}})$ حيث: $ل = \frac{1}{\alpha}$	بيان الحالة يستحيل معه - نظرياً - استخدام التوزيع المعتاد أو توزيع ت.

* نظرية (١) مذكورة في صفحة ١١٣ ** نظرية (٢) مذكورة في صفحة ١٢٢

التقدير بفترة ثـقة للفرق بين متوسطيْ مجتمعيْن

كثيراً ما نحتاج في حياتنا العملية إلى تقدير الفرق بين متوسطي مجتمعيْن وذلك من واقع عينتيْن تم سحبهما من هذيْن المجتمعيْن. فَعلى سبيل المثال قد نكون معنيين بمعرفة الفرق بين متوسطي درجات مجموعتيْن مختلفتيْن من الطلاب، أو معرفة الفرق بين متوسطي أجور الموظفين في شركتيْن مختلفتيْن وذلك باستخدام عينات من الطلاب أو من موظفي الشركتيْن دون اللجوء إلى دراسة تلك المجتمعات بأكملها.

وكما سبق أن أوضحنا في هذا الفصل أن الوسط الحسابي للعينة ($\bar{س}$) هو تقدير غير متحيز لمتوسط المجتمع ($م$)، لذلك نجد أن الفرق بين متوسطي العينتيْن ($\bar{س}_1 - \bar{س}_2$) هو أيضاً تقدير غير متحيز للفرق بين متوسطي المجتمعيْن المسحوب منهما العينتيْن ($م_1 - م_2$). وذلك لأن متوسط العينة الأولى ($\bar{س}_1$) هو تقدير غير متحيز لمتوسط المجتمع الأول ($م_1$)، ونفس الشيء بالنسبة لـ ($\bar{س}_2$)، ($م_2$). وحيث أن فترة الثـقة لمعلمة المجتمع المجهولة تأخذ الشكل العام الحالي:

المُقدِّر بنقطة للمعلمة $\pm$ الدرجة المعيارية × الانحراف المعياري للمُقدر بنقطة للمعلمة

فإنه يمكننا استنتــاج أن فتــرة الثـقة للفــرق بيــن متوسطي المجتمعيْن ($م_1 - م_2$) يمكن كتابتها على الصورة التالية:

الحد الأدنى لفترة الثـقة =
($\bar{س}_1 - \bar{س}_2$) – الدرجة المعيارية × الخطأ المعياري للفرق بين متوسطي العينتيْن

الحد الأعلى لفترة الثـقة =
($\bar{س}_1 - \bar{س}_2$) + الدرجة المعيارية × الخطأ المعياري للفرق بين متوسطي العينتيْن

أي أن فترة الثـقة للفرق بين متوسطي مجتمعيْن ($م_1 - م_2$) لا تختلف من حيث طريقة إنشائها عن فترة الثـقة لمتوسط المجتمع ($م$) سوى في أننا في هذه الحالة نعتبر($\bar{س}_1 - \bar{س}_2$) بدلا من $\bar{س}$، كما نعتبر الخطأ المعياري للفرق بين متوسطي عينتيْن بدلاً من الخطأ المعياري لمتوسط عينة.

والآن نستطيع أن نتعرض بالدراسة لتقدير فترة الثـقة للفرق بين متوسطي مجتمعيْن ($م_1 - م_2$) وذلك في الحالات المختلفة. وهنا نود أن نشير إلى أن تناولنا لهذا الأمر لن يكون بالتفصيل المتَّبع في حالة تقدير فترة الثـقة لمتوسط المجتمع وذلك منعاً للتكرار وتجنباً لما قد يصيب القارئ من ملل. لذلك فإنه تيسيراً على القارئ سوف نصنف الحالات التي يستخدم فيها التوزيع المعتاد وتوزيع ت إلى قسميْن فقط هما:

- فترات الثـقة للفرق بين متوسطيْ مجتمعيْن باستخدام التوزيع المعتاد.
- فترات الثـقة للفرق بين متوسطيْ مجتمعيْن باستخدام توزيع ت.

وسوف نتناول كلاً من هذيْن القسميْن بشيء من التفصيل وذلك على النحو التالي:

أولاً: فترات الثـقة للفرق بين متوسطيْ مجتمعيْن باستخدام التوزيع المعتاد

يستخدم التوزيع المعتاد في تقدير فترات الثـقة للفرق بين متوسطي مجتمعيْن وذلك في الحالات الموضحة في الجدول (٢ - ٥):

الجدول (٢ - ٥)
الحالات المختلفة لإيجاد فترات الثـقة للفرق بين متوسطي مجتمعيْن ($م_١ - م_٢$) باستخدام التوزيع المعتاد

الحالة	توزيع كل من المجتمعيْن	حجم كل من العينتيْن ($ن_١$، $ن_٢$)	تباين كل من المجتمعيْن ($\sigma_١$، $\sigma_٢$)
١	معتـــاد	كبيـر ($ن_١$، $ن_٢ \geq ٣٠$)	معلوم
٢	معتـــاد	صغـير ($ن_١$، $ن_٢ < ٣٠$)	معلوم
٣	معتـــاد	كبيـــر ($ن_١$، $ن_٢ \geq ٣٠$)	مجهول
٤	غير معتـــاد	كبيـــر ($ن_١$، $ن_٢ \geq ٣٠$)	معلوم
٥	غير معتـــاد	كبيـــر ($ن_١$، $ن_٢ \geq ٣٠$)	مجهول

وهذه الحالات الخمسة هي على الترتيب المناظرة للحالات رقم ١، ٢، ٣، ٥، ٦ في الجدول (٢ - ٤) والمتعلقة بعينة واحدة. ويلاحَظ أن مبررات استخدام التوزيع المعتاد هنا هي نفسها الموضحة في الجـدول (٢ - ٤) آخذين في الاعتبار أننا نتعامل هنا مع عينتيْن ومجتمعيْن لا عينة واحدة ومجتمع واحد.

هذا وتجدر الإشارة إلى أنه في الحالات ١، ٢، ٤ والموضحة في الجدول (٢ - ٥) حيث أن كلاً من المجتمعيْن يتبع توزيعاً معتاداً (أو أن كليهما أو أحدهما لا يتبع التوزيع المعتاد ولكن $ن_١$، $ن_٢ \geq ٣٠$) فإن توزيع المعاينة للفرق بين متوسطي عينتيْن ($\bar{س}_١ - \bar{س}_٢$) يتبع أيضاً التوزيع المعتاد (أو يتبعه تقريباً) وذلك بمتوســط ($م_١ - م_٢$) وخطـأ معياري $\sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}}$ ، هذا مع ملاحظة أنه يجب أن نأخـذ معامل التصحيح في الاعتبار متى كان ذلك ملائماً، وبذلك يمكننا أن نستنتج أن المقدار:

$$ي = \frac{(\bar{س}_1 - \bar{س}_2) - (م_1 - م_2)}{\sqrt{\frac{\sigma_1^2}{ن_1} + \frac{\sigma_2^2}{ن_2}}}$$

هو متغير عشوائي يتبع التوزيع المعتاد المعياري. أي أنه يمكننا – وعلى غرار ما تم في حالة المجتمع الواحد – استنتاج أن الفرق بين متوسطي المجتمعيْن ($م_1 - م_2$) يقع في الفترة:

$$(\bar{س}_1 - \bar{س}_2) \pm ي_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{ن_1} + \frac{\sigma_2^2}{ن_2}} \qquad (٢ - ٩)$$

وذلك بمعامل ثقة ($1 - \alpha$) ودرجة ثقة ($1 - \alpha$) × ١٠٠

أي أن:

الحد الأدنى لفترة الثقة = $(\bar{س}_1 - \bar{س}_2) - ي_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{ن_1} + \frac{\sigma_2^2}{ن_2}}$

الحد الأعلى لفترة الثقة = $(\bar{س}_1 - \bar{س}_2) + ي_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{ن_1} + \frac{\sigma_2^2}{ن_2}}$

حيث: $\bar{س}_1$، $\bar{س}_2$: الوسط الحسابي للعينة الأولى والعينة الثانية،

σ_1^2، σ_2^2: التباين للمجتمع الأول والمجتمع الثنائي على التوالي،

$ن_1$، $ن_2$: حجم العينة الأولى والعينة الثانية،

$ي_{\frac{\alpha}{2}}$ لها نفس التعريف السابق، حيث أن هناك توزيعاً معتاداً معيارياً واحداً يمكن أن تتحول إليه جميع التوزيعات المعتادة مهما اختلفت أوساطها الحسابية أو انحرافاتها المعيارية، أي أن: $ي_{\frac{\alpha}{2}}$ هي قيمة الدرجة المعيارية بحيث $\phi(ي_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$ وتُستخرج قيمتها من جدول المساحات تحت المنحنى المعتاد المعياري.

وأما في الحالتيْن (٣، ٥) والتي يكون فيهما الانحراف المعياري للمجتمع مجهولاً وحجم العينة كبيراً فإنه يمكن استخدام الانحراف المعياري للعينة (ع) بدلاً من الانحراف المعياري للمجتمع (σ) بنفس المنطق المتَّبع في حالة المجتمع الواحد وحيث ع معرَّفة بالصورة (٢ - ٤). أي أن الخطأ المعياري للفرق بين متوسطيْن في هاتيْن الحالتيْن يأخذ الشكل التالي:

$$ع_{(\bar{س}_1 - \bar{س}_2)} = \sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}$$

وهذا يعنى بالتالي أن المتغير (ي) حيث:

$$ي = \frac{(\bar{س}_1 - \bar{س}_2) - (م_1 - م_2)}{\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}}$$

هو أيضاً متغير عشوائي يتبع تقريباً التوزيع المعتاد المعياري. وبذلك يمكننا استنتاج أن فترة الثقة للفرق بين متوسطي مجتمعيْن تتحدد كما يلي:

$$\text{الحد الأدنى لفترة الثقة} = (\bar{س}_1 - \bar{س}_2) - ي_{\frac{\alpha}{2}} \sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}$$

$$\text{الحد الأعلى لفترة الثقة} = (\bar{س}_1 - \bar{س}_2) + ي_{\frac{\alpha}{2}} \sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}$$

وذلك بمعامل ثقة يساوي ($1 - \alpha$) ودرجة ثقة ($1 - \alpha$) × ١٠٠، وحيث ع معرَّفة بالصورة (٢ - ٤).

وفيما يلي بعض الأمثلة التي توضح كيفية إيجاد فترة الثقة للفرق بين متوسطيْ مجتمعيْن ($م_1 - م_2$) والتي تمثل بعض الحالات الخمسة المشار إليها:

مثال (١٧):

لنفترض أنه أُجري اختبار في الذكاء لعينة من الطلاب مختارة عشوائياً من كل من كلية التجارة وكلية التربية بإحدى الجامعات، فإذا كانت نتائج هذا الاختبار كما يلي:

البيان / الكلية	حجم العينة	متوسط الدرجات في العينة	الانحراف المعياري للمجتمع
كلية التجـــــارة	١٠٠	٨٥	١٦
كلية التربيـــــــة	١٥٠	٨٠	١٢

المطلـــوب:

إيجاد فترة ثقة ٩٥٪ للفرق بين متوسطي الذكاء في مجتمعي كلية التجارة وكلية التربية (وذلك بافتراض أن درجات الذكاء في كل من المجتمعيْن تتبع التوزيع المعتاد).

الحـــل:

هذا المثال يمثل الحالة (١) في الجدول (٢ - ٥) حيث يمكن معه استخدام التوزيع المعتاد في عملية التقدير، ومن بيانات المثال نجد أن:

$ن_1 = ١٠٠$ ، $\bar{س}_1 = ٨٥$ ، $\sigma_1 = ١٦$

$ن_2 = ١٥٠$ ، $\bar{س}_2 = ٨٠$ ، $\sigma_2 = ١٢$

وحيث أن درجة الثقة المطلوبة هي ٩٥٪ فإننا نحصل على:

$$\alpha = ٠{,}٠٥ \quad ، \quad ي_{\frac{\alpha}{٢}} = ١{,}٩٦$$

$$\text{الخطأ المعياري} = \sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}} = \sqrt{\frac{(١٦)^٢}{١٠٠} + \frac{(١٢)^٢}{١٥٠}} = \sqrt{٣{,}٥٢} = ١{,}٨٨$$

$$\text{الحد الأدنى لفترة الثقة} = (\bar{س}_١ - \bar{س}_٢) - ي_{\frac{\alpha}{٢}} \sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}}$$

$$= (٨٥ - ٨٠) - ١{,}٩٦ \times ١{,}٨٨$$

$$= ٥ - ٣{,}٦٨ = ١{,}٣٢ \text{ درجة}$$

$$\text{الحد الأعلى لفترة الثقة} = (\bar{س}_١ - \bar{س}_٢) + ي_{\frac{\alpha}{٢}} \sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}}$$

$$= (٨٥ - ٨٠) + ١{,}٩٦ \times ١{,}٨٨$$

$$= ٥ + ٣{,}٦٨ = ٨{,}٦٨ \text{ درجة}$$

أي أن فترة الثقة ٩٥٪ للفرق بين متوسطي ذكاء الطلاب في كليتي التجارة والتربية هي [١,٣٢ ، ٨,٦٨]. وحيث أن الحديْن الأدنى والأعلى لفترة الثقة موجبان، فإنه يمكن القول بدرجة ثقة ٩٥٪ أن متوسط المجتمع الأول ($م_١$) يميل إلى الزيادة عن متوسط المجتمع الثاني ($م_٢$) وذلك بمقدار يتراوح ما بين ١,٣٢ درجة، ٨,٦٨ درجة.

مثـــال (١٨):

في المثال السابق، إذا علمنا أن عدد الطلاب في كل من كليتي التجارة والتربية هو على الترتيب ١٥٠٠ طالب، ٢٠٠٠ طالب، أوجد بفترة ثقة ٩٩٪ الفرق بين متوسطي درجات الذكاء لطلاب الكليتيْن.

الحـــل:

حيث أن درجة الثقة المطلوبة في هذه الحالة هي ٩٩٪ فإننا نحصل على:

$$\alpha = ٠{,}٠١ \quad ، \quad ي_{\frac{\alpha}{٢}} = ٢{,}٥٨$$

ويلاحَظ أنه في المثال السابق حسبنا الخطأ المعياري على أساس أن المجتمع كان غير محدود، وأما في هذا المثال وبالرغم من كبر حجم كل من المجتمعيْن إلا أننا نجد أن ($ن_١ > ٠{,}٠٥\, N_١$ ، $ن_٢ > ٠{,}٠٥\, N_٢$) مما يعنى أنه يجب أن نأخذ في الاعتبار معامل التصحيح عند حسابنا للخطأ المعياري للفرق بين متوسطيْن، أي أن:

$$\text{الخطـأ المعياري} = \sqrt{\frac{\sigma_1^2}{ن_1}\left(\frac{ن_1 - ن_1}{ن_1 - 1}\right) + \frac{\sigma_2^2}{ن_2}\left(\frac{ن_2 - ن_2}{ن_2 - 1}\right)}$$

$$= \sqrt{\frac{(16)^2}{100}\left(\frac{1500-100}{1500-1}\right) + \frac{(12)^2}{150}\left(\frac{2000-150}{2000-1}\right)}$$

$$= \sqrt{2.38 + 0.89} = \sqrt{3.27} = 1.81$$

∴ الحد الأدنى لفترة الثـقة = $(\bar{س}_1 - \bar{س}_2) - ي_{\frac{\alpha}{2}}$ × الخطـأ المعياري

$= (85 - 80) - 2.58 \times 1.81$

$= 5 - 4.67 = 0.33$ درجة

الحد الأعلى لفترة الثـقة = $(\bar{س}_1 - \bar{س}_2) + ي_{\frac{\alpha}{2}}$ × الخطـأ المعياري

$= (85 - 80) + 2.58 \times 1.81$

$= 5 + 4.67 = 9.67$ درجة

أي أن الفرق بين متوسطي درجات الذكاء في كليتي التجارة والتربية يتراوح ما بين ٠,٣٣ و ٩,٦٧ درجة، وثقتنا في ذلك تبلغ ٩٩٪، وأن هذا الفرق هو لصالح طلاب كلية التجارة.

مثـــال (١٩):

سُحبت عينة عشوائية حجمها ٤٠ مفردة ومتوسطها ٢٥ من مجتمع معتاد. كما سُحبت عينة عشوائية أخرى حجمها ٦٠ مفردة ومتوسطها ٣٥ من مجتمع معتاد آخر. فإذا علمت أن الانحراف المعياري للعينة الأولى هو ($ع_1 = 9$) وأن الانحراف المعياري للعينة الثانية هو ($ع_2 = 15$) أوجد فترة الثـقة ٩٠٪ للفرق بين متوسطي المجتمعيْن.

الحـــــل:

الوضع في هذا المثال يناظر الحالة (٣) في الجدول (٢ - ٥) والتي يمكن معها استخدام التوزيع المعتاد في عملية التقدير وذلك على النحو التالي:

حيث أن درجة الثـقة المطلوبة هي ٩٠٪ فإننا نحصل على:

$\alpha = 0.10$ ، $ي_{\frac{\alpha}{2}} = 1.645$

$$\text{الخطـأ المعياري} = \sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}} = \sqrt{\frac{(9)^2}{40} + \frac{(15)^2}{60}} = \sqrt{5.775} = 2.4$$

∴ الحد الأدنى لفترة الثقة = $(\bar{س}_1 - \bar{س}_2) - ي_{\frac{\alpha}{2}} \sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}$

= (٢٥ – ٣٥) – ١,٦٤٥ × ٢,٤

= -١٠ – ٣,٩٥ = - ١٣,٩٥

الحد الأعلى لفترة الثقة = $(\bar{س}_1 - \bar{س}_2) + ي_{\frac{\alpha}{2}} \sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}$

= (٢٥ – ٣٥) + ١,٦٤٥ × ٢,٤

= -١٠ + ٣,٩٥ = - ٦,٠٥

أي أن فتـــــرة الثـــقة ٩٠٪ للفــــــرق بيـــــــــن متوسطي المجتمعيْن ($م_1 - م_2$) هي [-١٣,٩٥ ، - ٦,٠٥].

وحيث أن الحديْن الأدنى والأعلى لفترة الثقة سالبان فإن ذلك يعنى أن الفرق بين المتوسطيْن هو لصالح متوسط المجتمع الثاني والذي يميل إلى الزيادة عن نظيره في المجتمع الأول بمقدار يتراوح ما بين ٦,٠٥، ١٣,٩٥ وذلك بدرجة ثقة ٩٠٪

مثال (٢٠):

المطلوب هو حل المثال السابق بافتراض أن كلاً من المجتمعيْن لا يتبع التوزيع المعتاد.

الحـــل:

باستخدام بيانات المثال السابق وبافتراض أن كلاً من المجتمعيْن لا يتبع التوزيع المعتاد فإن الوضع هنا يناظر الحالة (٥) في الجـــدول (٢ - ٥) والتي يستخدم فيها أيضاً التوزيع المعتاد في إيجاد فترة الثقة للفرق بين متوسطي المجتمعيْن ($م_1 - م_2$). ويأتي استخدام التوزيع المعتاد في هذه الحالة تطبيقاً لنظرية الحد المركزية مع استخدام الانحراف المعياري للعينة (ع) كتقدير للانحراف المعياري للمجتمع (σ) وذلك لكبر حجم العينة. نستنتج من ذلك أنه لا اختلاف بين طريقة الحل في هذا المثال وبين تلك المتبعة في المثال السابق، فالنتائج واحدة في الحالتيْن.

ثانياً: فترات الثقة للفرق بين متوسطيْ مجتمعيْن باستخدام توزيع ت

تناولنا من قبل الشروط التي يجب أن تتوافر حتى يمكننا استخدام توزيع ت في عملية تقدير متوسط مجتمع وذلك باستخدام بيانات عينة مسحوبة من هذا المجتمع. وحيث أن الأمر يتعلق هنا بمجتمعيْن وعينتيْن فإن شروط استخدام توزيع ت هنا هي نفس الشروط السابقة ولكن بالنسبة لكل مجتمع وعينة على حدة. ويُضاف إلى تلك الشروط شرط رابع وهو أن المجتمعيْن

لهما نفس التباين المجهول. هذا وعملية تقدير الفرق بين متوسطي مجتمعيْن تحكمهما النظرية التالية:

نظرية: إذا كان لدينا عينتان صغيرتان مستقلتان مسحوبتان من مجتمعيْن معتاديْن تباينهما مجهولان. فإذا كان الوسط الحسابي والتباين للعينة الأولى هما: $\bar{\text{س}}_{١}$، $\text{ع}_{١}$ بينما الوسط الحسابي والتباين للعينة الثانية هما $\bar{\text{س}}_{٢}$، $\text{ع}_{٢}$ على الترتيب. فإنه بافتراض أن المجتمعيْن لهما نفس التباين المجهول فإن المتغير العشوائي ت حيث:

$$\text{ت} = \frac{\left(\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}\right) - \left(\text{م}_{١} - \text{م}_{٢}\right)}{\sqrt{\text{ع}_{\text{م}}^{٢}\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}} \qquad \textbf{(٢ - ١٠)}$$

يخضع لتوزيع ت بدرجات حرية ($\text{ن}_{١} + \text{ن}_{٢} - ٢$) وحيث:

$\text{م}_{١}$ الوسط الحسابي للمجتمع الأول،

$\text{م}_{٢}$ الوسط الحسابي للمجتمع الثاني،

$\text{ع}_{\text{م}}^{٢}$ التباين المشترك للمجتمعيْن والمرجح بدرجات الحرية، حيث:

$$\text{ع}_{\text{م}}^{٢} = \frac{\left(\text{ن}_{١} - ١\right)\text{ع}_{١}^{٢} + \left(\text{ن}_{٢} - ١\right)\text{ع}_{٢}^{٢}}{\text{ن}_{١} + \text{ن}_{٢} - ٢}$$

وباستخدام الصيغة (٢ -١٠)، وباتباع نفس الأسلوب المستخدم في تقدير فترة الثقة لمتوسط مجتمع، فإنه يمكننا الوصول إلى أن فترة الثقة للفرق بين متوسطي مجتمعيْن في هذه الحالة تأخذ الشكل التالي:

$$\left(\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}\right) \pm \text{ت}_{\frac{\alpha}{٢}} \sqrt{\text{ع}_{\text{م}}^{٢}\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}$$

أي أن:

$$\textbf{الحد الأدنى لفترة الثقة} = \left(\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}\right) - \text{ت}_{\frac{\alpha}{٢}} \sqrt{\text{ع}_{\text{م}}^{٢}\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}$$

$$\textbf{الحد الأعلى لفترة الثقة} = \left(\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}\right) + \text{ت}_{\frac{\alpha}{٢}} \sqrt{\text{ع}_{\text{م}}^{٢}\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}$$

وذلك بمعامل ثقة ($١ - \alpha$) ودرجة ثقة ($١ - \alpha$) × ١٠٠، وحيث $\text{ع}_{\text{م}}^{٢}$ هو التباين المشترك للمجتمعيْن والمعرَّف من قبل مع ملاحظة أن قيمة $\text{ت}_{\frac{\alpha}{٢}}$ تعتمد على درجات الحرية ($\text{ن}_{١} + \text{ن}_{٢} - ٢$).

مثال (٢١):

سُحبت عينة عشوائية حجمها ١٥ مفردة بمتوسط ٤٥ وذلك من مجتمع معتاد مجهولٌ تباينه ووسطه الحسابي. كما سُحبت عينة عشوائية أخرى حجمها ١٠ بمتوسط ٣٧ وذلك من مجتمع معتاد آخر مجهولٌ تباينه ووسطه الحسابي أيضاً. أوجد فترة الثقة ٩٥٪ للفرق بين متوسطي المجتمعيْن إذا علمت أن الانحراف المعياري للعينة الأولى هو ٦ وللعينة الثانية هو ٨ وأنه يُفترض أن المجتمعيْن لهما نفس التباين.

الحـل:

حيث أن كلاً من المجتمعيْن يتبع توزيعاً معتاداً وسطه الحسابي وانحرافه المعياري غير معلوميْن، وحيث أن حجم كل من العينتيْن صغير ($ن_١$، $ن_٢$ < ٣٠) فإن التوزيع الملائم لتقدير فترة الثقة المطلوبة هو توزيع ت وذلك لافتراض أن المجتمعين لهما نفس التباين.

وبما أن درجة الثقة المطلوبة هي ٩٥٪ فإننا نحصل على: α = ٠,٠٥

وحيث أن درجات الحرية تساوي ($ن_١$+ $ن_٢$ – ٢) أي ٢٣ درجة فإنه من جدول ت يمكننا الحصول على: = $ت_{٢٣ ، ٠,٠٢٥}$ = ٢,٠٦٩

هذا ويمكننا حساب الخطأ المعياري للفرق بين متوسطي العينتيْن كما يلي:

$$ع_م^٢ = \frac{(ن_١ - ١)ع_١^٢ + (ن_٢ - ١)ع_٢^٢}{ن_١ + ن_٢ - ٢}$$

$$= \frac{١٤(٦)^٢ + ٩(٨)^٢}{١٥ + ١٠ - ٢} = \frac{١٠٨٠}{٢٣} = ٤٦,٩٦$$

$$\text{الخطأ المعياري} = \sqrt{ع_م^٢\left(\frac{١}{ن_١} + \frac{١}{ن_٢}\right)} = \sqrt{\frac{٤٦,٩٦}{١٥} + \frac{٤٦,٩٦}{١٠}} = \sqrt{٧,٨٣} = ٢,٨$$

∴ الحد الأدنى لفترة الثقة = ($\bar{س}_١$ – $\bar{س}_٢$) – $ت_{\frac{\alpha}{٢}}$ × الخطأ المعياري

= (٤٥ – ٣٧) – ٢,٠٦٩ × ٢,٨

= ٨ – ٥,٧٩ = ٢,٢١

الحد الأعلى لفترة الثقة = ($\bar{س}_١$ – $\bar{س}_٢$) + $ت_{\frac{\alpha}{٢}}$ × الخطأ المعياري

= (٤٥ – ٣٧) + ٢,٠٦٩ × ٢,٨

= ٨ + ٥,٧٩ = ١٣,٧٩

أي أن فترة الثقة ٩٥٪ للفرق بين متوسطي المجتمعيْن ($م_١ - م_٢$) هي [٢,٢١ ، ١٣,٧٩]، وأن هذا الفرق هو لصالح متوسط المجتمع الأول ($م_١$) والذي يتجه إلى الزيادة عن مثيله في المجتمع الثاني ($م_٢$) بمقدار يتراوح ما بين ٢,٢١ و ١٣,٧٩ وذلك بثقة ٩٥٪.

مثال (٢٢):

في مصنع لإنتاج اللمبات الكهربائية أُستخدِمت طريقتان مختلفتان للإنتاج (أ ،ب)، فإذا سحبنا عينة من الوحدات المنتَجة باستخدام كل من الطريقتيْن وكانت النتائج كالتالي:

البيان / الطريقة	حجم العينة	متوسط أعمار المصابيح في العينة (بالساعات)	الانحراف المعياري للعينة
أ	٨	١٢٠٠	٢٥
ب	١٢	١١٠٠	٢٠

المطلـــوب:

إيجاد فترة الثقة ٩٠٪ للفرق بين متوسطي أعمار المصابيح المنتَجة بكل من الطريقتيْن، وذلك بافتراض أن أعمار المصابيح المنتَجة لكل من الطريقتيْن تتبع التوزيع المعتاد.

الحـــــل:

حيث أن كلاً من المجتمعيْن يتبع توزيعاً معتاداً تباينه مجهول، كما أن حجم كل من العينتيْن صغير ($ن_١$، $ن_٢$ < ٣٠) فإن توزيع ت هو التوزيع المناسب لتقدير فترة الثقة للفرق بين متوسطي المجتمعيْن ($م_١ - م_٢$).

وحيث أن درجة الثقة المطلوبة هي ٩٠٪، فان $\alpha = ٠,١٠$

$$\therefore\ ت_{(ن_١+ن_٢-٢)،\frac{\alpha}{٢}} = ت_{(٨+١٢-٢)،٠,٠٥} = ت_{١٨،٠,٠٥}$$

ومن جدول ت نحصل على:

$$ت_{١٨،٠,٠٥} = ١,٧٣٤$$

كما أن:

$$ع^٢_م = \frac{(ن_١-١)ع^٢_١ + (ن_٢-١)ع^٢_٢}{ن_١+ن_٢-٢}$$

$$= \frac{٧(٢٥)^٢ + ١١(٢٠)^٢}{٨+١٢-٢} = \frac{٨٧٧٥}{١٨} = ٤٨٧,٥$$

$$ع_{(\bar{س}_1 - \bar{س}_2)} = \sqrt{\frac{ع_م^2}{ن_1} + \frac{ع_م^2}{ن_2}}$$

$$= \sqrt{\frac{٤٨٧,٥}{٨} + \frac{٤٨٧,٥}{١٢}} = \sqrt{١٠١,٥٦} = ١٠,٠٨$$

∴ الحد الأدنى لفترة الثقة =

$$(\bar{س}_1 - \bar{س}_2) - ت_{\frac{\alpha}{2}} \sqrt{\frac{ع_م^2}{ن_1} + \frac{ع_م^2}{ن_2}}$$

$$= (١٢٠٠ - ١١٠٠) - ١,٧٣٤ \times ١٠,٠٨$$

$$= ١٠٠ - ١٧,٤٨ = ٨٢,٥٢$$ ساعة

الحد الأعلى لفترة الثقة =

$$(\bar{س}_1 - \bar{س}_2) + ت_{\frac{\alpha}{2}} \sqrt{\frac{ع_م^2}{ن_1} + \frac{ع_م^2}{ن_2}}$$

$$= (١٢٠٠ - ١١٠٠) + ١,٧٣٤ \times ١٠,٠٨$$

$$= ١٠٠ + ١٧,٤٨ = ١١٧,٤٨$$ ساعة

أي أن الفرق بين متوسطيْ أعمار المصابيح المنتَجة بكل من الطريقتيْن يتراوح ما بين ٨٢,٥٢ و ١١٧,٤٨ ساعة، وثقتنا في ذلك تبلغ ٩٠٪. ولكون الحديْن الأدنى والأعلى موجبيْن فإن ذلك يشير إلى أن متوسط المجتمع الاول ($م_1$) يتجه إلى الزيادة عن متوسط المجتمع الثاني ($م_2$) بمقدار يتراوح ما بين ٨٢,٥٢ و ١١٧,٤٨ ساعة وذلك بثقة ٩٠٪.

التقدير بفترة ثقة للنسبة في المجتمع

بعد أن تناولنا توزيع المعاينة للوسط الحسابي، ورأينا كيفية إيجاد فترات الثقة للوسط الحسابي للمجتمع وكذلك للفرق بين وسطيْن حسابييْن لمجتمعيْن مختلفيْن، فإنه من الملائم الآن أن نتناول بالدراسة تقدير فترة الثقة للنسبة في المجتمع حيث أن نسبة وجود ظاهرة ما في مجتمع من المجتمعات هو أمر على قدر كبير من الأهمية. ذلك أن النسبة – مثلها مثل الوسط الحسابي – هي إحدى معلمات المجتمع والتي كثيراً ما تكون موضع اهتمام الباحثين في شتى المجالات.

وقبل أن نتناول هذا الموضوع بشيء من التفصيل لعله من الملائم أن نوضح للقارئ ما هو المقصود بالنسبة. فالنسبة هي عدد المفردات التي تملك خاصية معينة منسوباً إلى إجمالي عدد المفردات في المجتمع أو في العينة. ومثال ذلك نسبة الأشخاص المصابين بمرض معين أو نسبة المدخنين في مجتمع من المجتمعات إلى غير ذلك من الأمثلة.

هذا وإذا كنا بصدد دراسة ظاهرة ما في المجتمع فإن نسبة الظاهرة في المجتمع (وسنرمز لها بالرمز: ل) تتمثل في عدد مرات حدوث تلك الظاهرة في المجتمع مقسوماً على حجم المجتمع (ن)، أي أن:

$$\text{ل} = \frac{\text{عدد مرات حدوث الظاهرة في المجتمع}}{\text{حجم المجتمع (ن)}}$$

وفي دراستنا لتقدير النسبة في المجتمع بفترة ثقة فإنه من المناسب أن نبدأ بإيجاد توزيع المعاينة للنسبة وذلك على النحو الموضح فيما يلي:

أولاً: توزيع المعاينة للنسبة

إذا قمنا بسحب عينات عشوائية متكررة من مجتمع ما (أو جميع العينات الممكنة ذات الحجم الواحد)، ثم حسبنا نسبة الظاهرة لكل عينة نقوم بسحبها فإننا سوف نجد أن النسبة تختلف في قيمتها من عينة لأخرى. أي أن النسبة في العينة تُعتبر متغيراً عشوائياً يمكن إيجاد التوزيع الاحتمالي له والذي يعرف عادة باسم "توزيع المعاينة للنسبة". وهذا يعنى أن توزيع المعاينة للنسبة هو التوزيع الاحتمالي لنسب العينات ذات الحجم الواحد والمأخوذة من نفس المجتمع.

هذا ويمكننا إثبات أن نسبة الحدوث في العينة (سنرمز لها بالرمز: $\hat{\text{ل}}$) هي متغير عشوائي له توزيع يقترب من التوزيع المعتاد إذا كان حجم العينة كبيراً كبراً كافياً ($\text{ن} \geq ٣٠$) وكانت النسبة ليست بالكبيرة جداً أو الصغيرة جداً. ويرى بعض الإحصائيين أن التقريب يكون جيداً إذا انحصرت قيمة النسبة في المجتمع ما بين ٠,٠٥، ٠,٩٥. وهذا التوزيع المعتاد له توقُّـع يساوي ل وتباينه يساوي $\frac{\text{ل ك}}{\text{ن}}$، حيث:

ل: نسبة الحدوث في المجتمع،

ن: حجم العينة،

ك: نسبة عدم الحدوث في المجتمع، أي أن: $\text{ك} = ١ - \text{ل}$

وبذلك يمكننا استنتاج ما يلي:

أن نسبة حدوث الظاهرة في العينة ($\hat{\text{ل}}$) هي متغير عشوائي يقترب توزيعه من التوزيع المعتاد (وفقاً للشروط المشار إليها) وذلك بوسط حسابي وانحراف معياري يعرفان على النحو التالي:

$$\text{ت}(\hat{\text{ل}}) = \text{ل} \text{ ،}$$

$$\sigma_{\hat{\text{ل}}} = \sqrt{\frac{\text{ل ك}}{\text{ن}}}$$ ويسمى بالخطأ المعياري للنسبة

حيث: $\hat{\text{ل}}$ النسبة في العينة ، ل النسبة في المجتمع،

$\text{ك} = ١ - \text{ل}$ ، ن: حجم العينة

وبناءً على ذلك، فإنه وفقاً لنظرية الحد المركزية يمكننا استنتاج أن المتغير العشوائي:

$$ي = \frac{\hat{ل} - ل}{\sqrt{\frac{لك}{ں}}}$$

له توزيع يقترب من التوزيع المعتاد المعياري وذلك عندما يكون حجم العينة كبيراً كبراً كافياً ($ں \geq ٣٠$) وكذلك ںل، ںك ≥ ٥ كما يرى بعض الإحصائيين.

هذا وفي الحالات التي يكون فيها ($ں > ٠,٠٥ ن$) فإن الخطأ المعياري يمكن إيجاده آخذين في الاعتبار معامل التصحيح المشار إليه من قبل عند دراستنا لتوزيع المعاينة للوسط الحسابي، أي أن:

$$\sigma_{\hat{ل}} = \sqrt{\frac{لك}{ں}} \sqrt{\frac{ن - ں}{ن - ١}} \qquad \textbf{(٢ - ١١)}$$

وحيث أن ل غير معلومة (إذ أننا معينين أساساً بتقديرها) فإنه يمكن حساب الخطأ المعياري بالاعتماد على بيانات العينة وذلك على النحو التالي:

$$ع_{\hat{ل}} = \sqrt{\frac{\hat{ل}\hat{ك}}{ں}} \qquad \textbf{(٢ - ١٢)}$$

أي أن الخطأ المعياري تتحدد قيمته باستخدام نسب العينة وليست نسب المجتمع (يلاحَظ هنا أن $\hat{ل}$ هي تقدير غير متحيز لنسبة المجتمع ل).

وفي حالة استخدام معامل التصحيح فإن الخطأ المعياري يأخذ الصورة التالية:

$$ع_{\hat{ل}} = \sqrt{\frac{\hat{ل}\hat{ك}}{ں}} \sqrt{\frac{ن - ں}{ن - ١}} \qquad \textbf{(٢ - ١٣)}$$

ولعل المثال التالي يوضح للقارئ بعض النقاط المُشار اليها.

مثال (٢٣):

لنفترض أننا بصدد دراسة المدخـنين في مجتمع مكوَّن من ستة طلاب (أ ، ب ، **جـ** ، ء ، هـ ، و)، فإذا كان معلوماً لدينا أن الطلاب المدخنين هم (أ ، ء ، هـ) وأن الطلاب غير المدخنين في المجتمع هم (ب ، **جـ** ، و).

المطلـــــوب:

إيجاد توزيع المعاينة لنسبة المدخنين في المجتمع وذلك باستخدام عينة حجمها ٤ طلاب.

الحـــــل:

حيث أن هناك طلاب مدخنين من بين الطلاب الستة المكونين للمجتمع فإننا نجد ما يلي:

$$\text{نسبة المدخنين في المجتمع} = \frac{\text{عدد الطلاب المدخنين}}{\text{عدد الطلاب في المجتمع}} = \frac{٣}{٦} = ٠,٥$$

أي أن: $\hat{\text{ل}} = ٠,٥$

هذا ويوضح الجدول (٢ - ٦) جميع العينات الممكنة ذات الحجم ٤ والتي يمكن سحبها من مجتمع الطلاب، كما يوضح نسبة المدخنين في كل عينة وانحراف تلك النسبة عن النسبة الحقيقية في المجتمع (أي: $\hat{\text{ل}} - \text{ل}$).

هذا مع ملاحظة أن الحروف التي تحتها خط تمثل الأشخاص المدخنين.

الجدول (٢ - ٦)

بيانات جميع العينات ذات الحجم ٤ والتي يمكن سحبها من مجتمع حجمه ٦ فيما يتعلق بظاهرة التدخين

رقم العينة	مفردات العينة				نسبة المدخنين ($\hat{\text{ل}}$)	($\hat{\text{ل}} - \text{ل}$)، ل = ٠,٥
١	أ	ب	جـ	د	٠,٥٠	صفر
٢	أ	ب	جـ	هـ	٠,٥٠	صفر
٣	أ	ب	جـ	و	٠,٢٥	-٠,٢٥
٤	أ	ب	د	هـ	٠,٧٥	٠,٢٥
٥	أ	ب	د	و	٠,٥٠	صفر
٦	أ	ب	هـ	و	٠,٥٠	صفر
٧	أ	جـ	د	هـ	٠,٧٥	٠,٢٥
٨	أ	جـ	د	و	٠,٥٠	صفر
٩	أ	جـ	هـ	و	٠,٥٠	صفر
١٠	أ	د	هـ	و	٠,٧٥	٠,٢٥
١١	ب	جـ	د	هـ	٠,٥٠	صفر
١٢	ب	جـ	د	و	٠,٢٥	-٠,٢٥
١٣	ب	جـ	هـ	و	٠,٢٥	-٠,٢٥
١٤	ب	د	هـ	و	٠,٥٠	صفر
١٥	جـ	د	هـ	و	٠,٥٠	صفر

ومن واقع بيانات الجدول (٢ - ٦) يمكننا ايجاد توزيع المعاينة للنسبة في العينة وذلك كما يوضحه الجدول (٢ - ٧). ومن الجدولين يمكننا استنتاج ما يلي:

$$\text{ت}(\hat{\text{ل}}) = \frac{\sum \hat{\text{ل}}}{١٥} = \frac{٧,٥}{١٥} = ٠,٥ \qquad \text{أي أن: } \text{ت}(\hat{\text{ل}}) = \text{ل}$$

الجدول (٢ - ٧)

توزيع المعاينة لنسبة عينة حجهما ٤ مسحوبة من مجتمع حجم ٦

نسبة المدخنين في العينة ($\hat{ل}$)	٠,٢٥	٠,٥٠	٠,٧٥
عدد مرات التكرار	٣	٩	٣
$ع(\hat{ل}) = \frac{\text{عدد مرات التكرار}}{١٥}$	٠,٢	٠,٦	٠,٢

وهذا يعنى أن متوسط نسب العينات يساوي النسبة في المجتمع.

كما يمكننا إثبات ذلك باستخدام توزيع المعاينة للنسبة حيث نجد أن:

$$ت(\hat{ل}) = \sum \hat{ل}\ ع(\hat{ل})$$

$$= ٠,٢٥ \times ٠,٢ + ٠,٥ \times ٠,٦ + ٠,٧٥ \times ٠,٢ = ٠,٥$$

أي أن: $ت(\hat{ل}) = ل$

وهذا يدل على أن نسبة العينة ($\hat{ل}$) هي تقدير غير متحيز للنسبة في المجتمع (ل).

$$\sigma^2_{\hat{ل}} = \sum \left(\hat{ل} - ل\right)^2 ع(\hat{ل})$$

$$= ٠,٢(٠,٢٥ - ٠,٥)^٢ + ٠,٦(٠,٥ - ٠,٥)^٢ + ٠,٢(٠,٧٥ - ٠,٥)^٢$$

$$= ٠,٢ \times ٠,٠٦٢٥ + ٠,٦ \times صفر + ٠,٢ \times ٦٢٥ = ٠,٢٥$$

هذا ويمكننا التحقق من صحة القانون (٢ - ١١) وذلك على النحو الموضح فيما يلي:

حيث أن:

$$\sigma_{\hat{ل}} = \sqrt{\frac{\hat{ل}\hat{ك}}{ن}}\ \sqrt{\frac{ن - ن}{ن - ١}}$$

$$\sigma_{\hat{ل}} = \sqrt{\frac{٠,٥ \times ٠,٥}{٤}}\ \sqrt{\frac{٦ - ٤}{٦ - ١}} = \sqrt{٠,٠٢٥} = ٠,١٥٨$$

أي أن: $\sigma^2_{\hat{ل}} = ٠,٠٢٥$ وهي نفس النتيجة التي حصلنا عليها من قبل باستخدام توزيع المعاينة للنسبة في العينة.

ثانياً: تقدير فترة الثقة للنسبة في المجتمع

سبق أن ذكرنا أن المتغير **ي** حيث: $ي = \frac{\hat{ل} - ل}{\sqrt{\frac{ل ك}{ن}}}$

له توزيع يقترب من التوزيع المعتاد المعياري، لذلك فإنه يمكن تحديد فترات الثقة للنسبة في المجتمع وذلك على النحو التالي:

حيث أن ي متغير عشوائي له توزيع معتاد معياري فإننا نجد أن:

$$ع(-ي_{\frac{\alpha}{2}} \leq \frac{\hat{ل} - ل}{\sqrt{\frac{ل ك}{ن}}} \leq ي_{\frac{\alpha}{2}}) = ١ - \alpha$$

وحيث أن ل، ك عادة ما تكونان مجهولتان فانه يمكن الاستعاضة عنهما في حساب الخطأ المعياري بنسبتي العينة $\hat{ل}$ ، $\hat{ك}$ ، أي أن:

$$ع(-ي_{\frac{\alpha}{2}} \leq \frac{\hat{ل} - ل}{\sqrt{\frac{\hat{ل} \hat{ك}}{ن}}} \leq ي_{\frac{\alpha}{2}}) = ١ - \alpha \qquad (٢ - ١٤)$$

أي أن:

$$ع(-ي_{\frac{\alpha}{2}} \sqrt{\frac{\hat{ل} \hat{ك}}{ن}} \leq \hat{ل} - ل \leq ي_{\frac{\alpha}{2}} \sqrt{\frac{\hat{ل} \hat{ك}}{ن}}) = ١ - \alpha$$

$$\therefore ع(\hat{ل} - ي_{\frac{\alpha}{2}} \sqrt{\frac{\hat{ل} \hat{ك}}{ن}} \leq ل \leq \hat{ل} + ي_{\frac{\alpha}{2}} \sqrt{\frac{\hat{ل} \hat{ك}}{ن}}) = ١ - \alpha$$

ومنه يمكننا استنتاج أن:

الحد الأدنى لفترة الثقة = $\hat{ل} - ي_{\frac{\alpha}{2}} \sqrt{\frac{\hat{ل} \hat{ك}}{ن}}$

الحد الأعلى لفترة الثقة = $\hat{ل} + ي_{\frac{\alpha}{2}} \sqrt{\frac{\hat{ل} \hat{ك}}{ن}}$

حيث: $\hat{ل}$ النسبة في العينة ، $\hat{ك} = ١ - \hat{ل}$،

ن حجم العينة

$ي_{\frac{\alpha}{2}}$ قيمة الدرجة المعيارية ي، بحيث: $\phi(ي) = ١ - \frac{\alpha}{2}$

وحيث تتحد قيمة $ي_{\frac{\alpha}{2}}$ على النحو التالي:

$ي_{\frac{\alpha}{2}} = ١,٦٤٥$ في حالة $\alpha = ٠,١$ ،

$= ١,٩٦$ في حالة $\alpha = ٠,٠٥$ ،

$= ٢,٥٨$ في حالة $\alpha = ٠,٠١$

مثال (٢٤):

في عينة عشوائية تضم ٢٥٠ ناخباً وجد أن هناك ١٥٠ ناخباً يؤيدون مرشحاً معيناً من مرشحي دائرتهم الانتخابية وذلك لعضوية مجلس الشعب. فإذا علمت أن إجمالي عدد الذين سيدلون بأصواتهم في هذه الدائرة الانتخابية يبلغ ١٥٠٠ ناخباً، أوجد بفترة ثقة ٩٥٪ نسبة الناخبين الذين سيدلون بأصواتهم لصالح هذا المرشح.

الحـــل:

نسبة المؤيدين للمرشح في العينة = $\hat{ل} = \frac{١٥٠}{٢٥٠} = ٠,٦$

وحيث أن حجم العينة كبير كما أن $ن\hat{ل}$ ، $ن\hat{ك}$ > ٥ فإنه يمكن استخدام التوزيع المعتاد في إيجاد فترات الثقة للنسبة في المجتمع. أي أن حدي الثقة يعطيان في الصورة التالية:

$$\hat{ل} \pm ي_{\frac{\alpha}{٢}} ع_{\hat{ل}}$$

حيث $ع_{\hat{ل}}$ هي الخطأ المعياري للنسبة في العينة $(\hat{ل})$

وحيث أن: ن > ٠,٠٥ ن فإنه يجب أخذ معامل التصحيح في الاعتبار عند حسابنا للخطأ المعياري، أي أن:

$$\text{الخطأ المعياري} = \sqrt{\frac{\hat{ل}\hat{ك}}{ن}}\sqrt{\frac{ن - ن}{ن - ١}}$$

$$= \sqrt{\frac{٠,٦ \times ٠,٤}{٢٥٠}}\sqrt{\frac{١٥٠٠ - ٢٥٠}{١٥٠٠ - ١}} = ٠,٠٢٨$$

∴ الحد الأدنى لفترة الثقة = ٠,٦ – ١,٩٦ × ٠,٠٢٨

= ٠,٦ – ٠,٠٥٥ = ٠,٥٤٥ ،

الحد الأعلى لفترة الثقة = ٠,٦ + ١,٩٦ × ٠,٠٢٨

= ٠,٦ + ٠,٠٥٥ = ٠,٦٥٥

وهذا يعنى أنه من المتوقَّع أن يحصل هذا المرشح على نسبة تتراوح ما بين ٥٤,٥٪، ٦٥,٥٪ من إجمالي أصوات الناخبين في دائرته الانتخابية، ودرجة ثقتنا في ذلك هي ٩٥٪.

مثال (٢٥):

أُجري استطلاع للآراء بشأن امتحانات نصف العام فوافق عليه ٨٠ طالباً من بين أفراد عينة مكونة من ٤٠٠ طالب. أوجد فترة الثقة ٩٠٪ لنسبة الطلاب الذين يؤيدون عقد امتحانات في نصف العام الدراسي.

الحـــل:

نسبة المؤيدين = $\hat{ل} = \frac{٨٠}{٤٠٠} = ٠,٢$

وحيث أن: $ن\hat{ل}$ ، $ن\hat{ك} > ٥$

فإنه يمكن إيجاد فترة الثقة على النحو التالي:

حيث أن درجة الثقة المطلوبة هي ٩٠٪ فإننا نجد أن:

$\alpha = ٠,١$ ، $ي_{\frac{\alpha}{٢}} = ١,٦٤٥$

الخطأ المعياري = $\sqrt{\frac{\hat{ل}\hat{ك}}{ن}} = \sqrt{\frac{٠,٢ \times ٠,٨}{٤٠٠}} = ٠,٠٢$

∴ الحد الأدنى لفترة الثقة = $\hat{ل} - ي_{\frac{\alpha}{٢}} \sqrt{\frac{\hat{ل}\hat{ك}}{ن}}$

$= ٠,٢ - ١,٦٤٥ \times ٠,٠٢ = ٠,١٦٧$

الحد الأعلى لفترة الثقة = $\hat{ل} + ي_{\frac{\alpha}{٢}} \sqrt{\frac{\hat{ل}\hat{ك}}{ن}}$

$= ٠,٢ + ١,٦٤٥ \times ٠,٠٢ = ٠,٢٣٣$

أي أن نسبة الطلاب الذين يؤيدون إجراء امتحانات في نصف العام تتراوح ما بين ١٦,٧٪، ٢٣,٣٪ وذلك بدرجة ثقة ٩٠٪.

التقدير بفترة ثقة للفرق بين نسبتيْن

يمكن تقدير فترات الثقة للفرق بين نسبتيْن في مجتمعيْن مختلفيْن وذلك على النحو السابق شرحه حيث يلزم في البداية إيجاد توزيع المعاينة للفرق بين نسبتيْن. وحيث أنه من المعروف أن $\hat{ل}$ هي تقدير غير متحيز للنسبة في المجتمع (ل) فإنه يمكننا استنتاج أن $\hat{ل}_{١}$ هي تقدير غير متحيز للنسبة في المجتمع الأول ($ل_{١}$)، كما أن $\hat{ل}_{٢}$ هي تقدير غير متحيز للنسبة في المجتمع الثاني ($ل_{٢}$). وعليه فإننا نجد أن ($\hat{ل}_{١} - \hat{ل}_{٢}$) هي أيضاً تقدير غير متحيز للفرق بين النسبتيْن ($ل_{١} - ل_{٢}$) أي أن:

$$ت(\hat{ل}_{١} - \hat{ل}_{٢}) = ل_{١} - ل_{٢}$$

وبصفة عامة - وكما سبق أن أشرنا - فإن تقدير فترة الثقة لمعلمة المجتمع إنما يقوم أساساً على تقدير هذه المعلمة من العينة وعلى الخطأ المعياري لهذا التقدير وأخيراً على درجة الثقة المطلوبة. أي أن فترة الثقة للفرق بين النسبتيْن ($ل_{١} - ل_{٢}$) تأخذ الصورة التالية:

$$(\hat{ل}_{١} - \hat{ل}_{٢}) \pm ي_{\frac{\alpha}{٢}} \, ع_{(\hat{ل}_{١} - \hat{ل}_{٢})}$$

حيث $ع_{(\hat{ل}_{١} - \hat{ل}_{٢})}$ هي الخطأ المعياري للفرق بين النسبتيْن $(\hat{ل}_{١} - \hat{ل}_{٢})$

ويأخذ الخطأ المعياري للفرق بين النسبتيْن $(\hat{ل}_{١} - \hat{ل}_{٢})$ الصورة التالية:

$$ع_{(\hat{ل}_{١} - \hat{ل}_{٢})} = \sqrt{ع^{٢}_{\hat{ل}_{١}} + ع^{٢}_{\hat{ل}_{٢}}}$$

$$= \sqrt{\frac{\hat{ل}_{١}\hat{ك}_{١}}{ن_{١}} + \frac{\hat{ل}_{٢}\hat{ك}_{٢}}{ن_{٢}}} \qquad (٢ - ١٥)$$

هذا مع ملاحظة استخدام معامل التصحيح متى كان ذلك ملائماً. بالإضافة إلى أننا نفترض في هذه الحالة أن العينتيْن المسحوبتيْن مستقلتان.

وهكذا فإن فترة الثـقة للفرق بين النسبتيْن $(ل_{١} - ل_{٢})$ تتحدد كما يلي:

$(\hat{ل}_{١} - \hat{ل}_{٢}) \pm ي_{\frac{\alpha}{٢}}$ × الخطأ المعياري (كما هو معرف في ٢ - ١٥)

وهذا يعنى ما يلي:

الحد الأدنى لفترة الثـقة = $(\hat{ل}_{١} - \hat{ل}_{٢}) - ي_{\frac{\alpha}{٢}}$ × الخطأ المعياري

الحد الأعلى لفترة الثـقة = $(\hat{ل}_{١} - \hat{ل}_{٢}) + ي_{\frac{\alpha}{٢}}$ × الخطأ المعياري

وذلك بمعامل الثـقة $(١ - \alpha)$ ودرجة ثـقة $(١ - \alpha) \times ١٠٠$.

حيث:

$\hat{ل}_{١}$ النسبة في العينة الأولى، $\hat{ك}_{١} = ١ - \hat{ل}_{١}$،

$\hat{ل}_{٢}$ النسبة في العينة الثانيـــة، $\hat{ك}_{٢} = ١ - \hat{ل}_{٢}$،

$ن_{١}$ حجم العينة الأولى،

$ن_{٢}$ حجم العينة الثانية،

$ي_{\frac{\alpha}{٢}}$ الدرجة المعيارية (ي) حيث: $\phi(ي) = ١ - \frac{\alpha}{٢}$

مثال (٢٦):

أُجري استطلاع لآراء طلاب الفرقة الرابعة في كلٍّ من كليتي التجارة والتربية في إحدى الجامعات وذلك بشأن القيام برحلة مشتركة إلى الأقصر وأسوان خلال الإجازة الصيفية. فإذا

افترضنا أنه في عينة حجمها ٢٥٠ طالباً من طلاب الفرقة الرابعة بكلية التجارة وافق ١٧٥ طالباً على القيام بهذه الرحلة، في حين أنه في عينة حجمها ٩٠ طالباً من طلاب الفرقة الرابعة بكلية التربية وافق ٥٤ طالباً فقط على القيام بهذه الرحلة.

المطلــوب:

إيجاد فترة الثقة ٨٥٪ للفرق بين نسبة الموافقين على القيام بهذه الرحلة في الكليتيْن.

الحــل:

حيث أن:

$$ن_١ = ٢٥٠ \quad ، \quad \hat{ل}_١ = \frac{١٧٥}{٢٥٠} = ٠,٧ ،$$

$$ن_٢ = ٩٠ \quad ، \quad \hat{ل}_٢ = \frac{٥٤}{٩٠} = ٠,٦ ،$$

وحيث أن:

$$ع_{(\hat{ل}_١ - \hat{ل}_٢)} = \sqrt{\frac{\hat{ل}_١ \hat{ك}_١}{ن_١} + \frac{\hat{ل}_٢ \hat{ك}_٢}{ن_٢}}$$

$$= \sqrt{\frac{٠,٧ \times ٠,٣}{٢٥٠} + \frac{٠,٦ \times ٠,٤}{٩٠}} = ٠,٠٥٩$$

وحيث أن كلاً من:

$ن_١\hat{ل}_١$ ، $ن_١\hat{ك}_١$ ، $ن_٢\hat{ل}_٢$ ، $ن_٢\hat{ك}_٢$ أكبر من ٥

فإنه يمكن استخدام التوزيع المعتاد في إيجاد فترة الثقة المطلوبة وذلك على النحو الموضح فيما يلي:

بما أن درجة الثقة المطلوبة هي ٨٥٪ فإننا نجد أن: $\alpha = ٠,١٥$،

$$\phi(ي_{\frac{\alpha}{٢}}) = ١ - \frac{\alpha}{٢} = ١ - \frac{٠,١٥}{٢} = ٠,٩٢٥$$

وبالبحث عن قيمة (ي) المناظرة للمساحة (الاحتمال) ٠,٩٢٥ نجد أن:

$$ي_{\frac{\alpha}{٢}} = ١,٤٤$$

∴ الحد الأدنى لفترة الثقة = $(\hat{ل}_١ - \hat{ل}_٢) - ي_{\frac{\alpha}{٢}} \times$ الخطأ المعياري

$$= (٠,٧ - ٠,٦) - ١,٤٤ \times ٠,٠٥٩ = ٠,٠١٥$$

الحد الأعلى لفترة الثقة = $(\hat{ل}_١ - \hat{ل}_٢) + ي_{\frac{\alpha}{٢}} \times$ الخطأ المعياري

$$= (٠,٧ - ٠,٦) + ١,٤٤ \times ٠,٠٥٩ = ٠,١٨٥$$

أي أن الفرق بين نسبة الموافقين على القيام بالرحلة المشتركة في الكليتيْن يتراوح ما بين ١,٥٪ و ١٨,٥٪ وذلك بدرجة ثقة ٨٥٪. وبما أن الحديْن الأدنى والأعلى موجبان فإنه يمكن تفسير هذا الفرق لصالح $ل_١$. أي أن نسبة الموافقين في كلية التجارة يميل إلى الزيادة على مثليتها في كلية التربية وهذه الزيادة تتراوح ما بين ١,٥٪ و ١٨,٥٪ وذلك بدرجة ثقة ٨٥٪.

مثال (٢٧):

المطلوب هو حل المثال السابق وذلك بافتراض أن عدد الطلاب الفرقة الرابعة في كلٍّ من كليتي التجارة والتربية هو على الترتيب ٩٠٠ طالب، ٢٥٠٠ طالب.

الحـــل:

حيث أن:

$$ن_١ = ٢٥٠ \quad ، \quad ن_١ = ٩٠٠$$

$$\therefore ن_١ > ٠,٠٥ ن_١$$

لذلك فإنه يجب أن نأخذ معامل التصحيح في الاعتبار عند حساب الخطأ المعياري للنسبة في العينة المسحوبة من كلية التجارة. أي أن:

$$ع_١^٢ = \left(\frac{\hat{ل}_١ \hat{ك}_١}{ن_١}\right)\left(\frac{ن_١ - ن_١}{ن_١ - ١}\right)$$

$$= \frac{٠,٧ \times ٠,٣}{٢٥٠}\left(\frac{٩٠٠ - ٢٥٠}{٩٠٠ - ١}\right) = ٠,٠٠٠٦١$$

وحيث أن: $ن_٢ = ٩٠ \quad ، \quad ن_٢ = ٩٠٠$

$$\therefore ن_٢ < ٠,٠٥ ن_٢$$

لذلك فإنه يمكن إهمال معامل التصحيح عند حساب الخطأ المعياري للنسبة في العينة المسحوبة من كلية التربية. وهذا يعنى حصولنا على النتيجة الموضحة في مثال (٢٦).

أي أن:

$$ع_١^٢ = \frac{٠,٦ \times ٠,٤}{٩٠٠} = ٠,٠٠٢٦٧$$

مما تقدم يمكننا حساب الخطأ المعياري للفرق بين النسبتيْن ($ل_١ - ل_٢$) على النحو التالي:

$$ع_{(\hat{ل}_١ - \hat{ل}_٢)} = \sqrt{٠,٠٠٠٦١ + ٠,٠٠٢٦٧} = \sqrt{٠,٠٠٣٢٨} = ٠,٠٥٧$$

بعد ذلك يمكن إيجاد فترة الثقة للفرق بين النسبتيْن ($ل_١ - ل_٢$) وذلك على النحو المتبع في مثال (٢٦).

أي أن:

الحد الأدنى لفترة الثقة = (٠,٧ – ٠,٦) – ١,٤٤ × ٠,٠٥٧ = ٠,٠١٨

الحد الأعلى لفترة الثقة = (٠,٧ – ٠,٦) + ١,٤٤ × ٠,٠٥٧ = ٠,١٨٢

أي أن الفرق بين نسبة الموافقين على القيام بالرحلة المشتركة في الكليتيْن يتراوح ما بين ١,٧٩٪ و ١٨,٢٪ وذلك بدرجة ثقة ٨٥٪. وهنا أيضاً تتجه نسبة الموافقين على القيام بالرحلة المشتركة في كلية التجارة إلى الزيادة عن مثيلتها في كلية التربية، وهذه الزيادة تتراوح ما بين ١,٨٪ و ١٨,٢٪ وذلك بدرجة ثقة ٩٥٪.

تحديد حجم العينة

إن تحديد حجم العينة هو أحد الأمور الهامة في مجال التحليل الإحصائي. إذ أن الاختيار المناسب للعينة العشوائية (بسيطة كانت أم طبقية أم غير ذلك) لا يكفي وحده للحصول على نتائج موثوق بها وعلى درجة عالية من الدقة. فالحجم الملائم للعينة المختارة هو أيضاً على قدر كبير من الأهمية وذلك إذا ما كان هدفنا هو الحصول على نتائج تعكس - إلى حد كبير - الواقع الحقيقي للمجتمع الذي سحبت منه العينة. فَعلى سبيل المثال لو كان حجم العينة أصغر مما ينبغي فإن ذلك قد يؤدى إلى نتائج تنقصها الدقة، وبالتالي فإن تعميم نتائج العينة على المجتمع الذي سحبت منه يصبح أمراً مشكوكاً في صحته. وفي الجانب الآخر فإنه بالرغم من أن دقة النتائج تزداد بزيادة حجم العينة لا يمكننا - من الناحية العملية - زيادة هذا الحجم بغير حدود أو تحفظات. ولو لم يكن الأمر كذلك لما كانت هناك مشكلة من أي نوع تتعلق باختيار العينة كمّاً أو كيفاً، أو لنقل لما كانت هناك حاجة أساساً لاختيار العينة طالما كان في إمكاننا إجراء الدراسة على جميع مفردات المجتمع (الحصر الشامل) دون محاذير أو مشكلات. ولكن الواقع يشير إلى أن اختيار عينة كبيرة جدا فيه تبديد للموارد، كما يتطلب وقتاً وجهداً كبيريْن. هذا بالإضافة إلى أن الدقة في النتائج المتحصل عليها قد لا تعادل في قيمتها مقدار ما بُذل من جهد ووقت ومال.

وعلى ضوء هذه الاعتبارات تصبح عملية تحديد الحجم الملائم للعينة أمراً بالغ الأهمية عند إجراء بحث من البحوث. إذ أنه من الضروري أن نوازن دائماً ما بين الجهد والوقت والتكلفة في جانب وبين درجة الدقة المطلوبة وما يترتب عليها في عملية اتخاذ القرارات في جانب آخر. وحيث أن مقدار الجهد والوقت والتكلفة يتوقف في مُجمله على حجم العينة فإنه يمكننا القول إن الحجم الملائم للعينة إنما يتحدد وفقا لدرجة الدقة المطلوبة في النتائج، أي وفقاً لمقدار الخطأ المسموح به والذي يُتوقع أن نقع فيه عند تقديرنا لمعلمة المجتمع المختلفة. وهذا الخطأ يسمى بـ "خطأ المعاينة".

هذا وفي تحديدنا للحجم الملائم للعينة سوف نأخذ في اعتبارنا الحالتيْن التاليتيْن:

- تحديد حجم العينة الملائم عند تقدير الوسط الحسابي للمجتمع.
- تحديد حجم العينة الملائم عند تقدير النسبة في المجتمع.

وسوف نتناول كلاً من هاتيْن الحالتيْن بشيء من التفصيل وذلك على النحو الموضَّح فيما يلي:

أولاً: تحديد حجم العينة في حالة تقدير متوسط المجتمع

لعله من الملائم هنا أن نضرب مثالاً نوضح فيه بعض الاعتبارات التي تحكم عملية تحديد الحجم الملائم للعينة في حالة تقديرنا لمتوسط المجتمع (μ). لو افترضنا أننا نريد أخذ عينة من أوراق إجابات طلاب الثانوية العامة في مادة اللغة الانجليزية وذلك بهدف التعرف على مستوى إجابات مجتمع الطلاب في هذه المادة والتي يرى المسئولون أن الامتحان فيها جاء فوق مستوى الطالب المتوسط. فإذا كان الأمر هكذا فإن ذلك سوف يؤخذ في الاعتبار عند تصحيح أوراق إجابات الطلاب. لنفترض أننا نرغب في أن يكون المتوسط الحقيقي لدرجات مجتمع الطلاب (أي جميع طلاب الثانوية العامة الذين أدوا هذا الامتحان) لا يختلف في قيمته عن متوسط الدرجات في العينة التي نود اختيارها بأكثر من درجتيْن، وأن نثق في ذلك بدرجة ٩٥٪ أي أن: $\mu = \bar{س} \pm ٢$ بدرجة ثقة ٩٥٪

لنفترض أيضاً أنه من واقع سجلات سابقة وتحت نفس الظروف وجدنا أن الانحراف المعياري لدرجات طلاب الثانوية العامة في مادة اللغة الإنجليزية يقترب كثيراً من ٢٥ درجة.

وهنا يبرز التساؤل التالي: ما هو حجم العينة اللازم للحصول على درجة الدقة المطلوبة وذلك بافتراض أن إجمالي عدد الطلاب الذين أدَّوْا امتحان اللغة الانجليزية هو خمسون ألف طالب (أي أن: $ن = ٥٠٠٠٠$)؟

يمكننا الإجابة على هذا التساؤل كما يلي: حيث أن فترة الثقة لمتوسط المجتمع (μ) تُعرف وفقاً للصورة التالية:

$$\bar{س} \pm ي_{\frac{\alpha}{٢}} \, \sigma_{\bar{س}}$$

وحيث أن درجة الثقة المطلوبة هي ٩٥٪، فإن: $ي_{\frac{\alpha}{٢}} = ١,٩٦$، أي أن فترة الثقة لمتوسط درجات المجتمع تتحدد كما يلي:

$$\bar{س} \pm ١,٩٦ \, \sigma_{\bar{س}}$$

وهذا يعنى أن فترة الثقة لمتوسط درجات مجتمع الطلاب لابد وأن تضم متوسط الدرجات بالعينة ناقصاً أو زائداً ١,٩٦ خطأ معياري. وحيث أننا لا نريد لأيٍّ من حدي الثقة أن يبتعد عن متوسط العينة بأكثر من درجتيْن، فإن ذلك يعنى ما يلي:

$$٢ = ١,٩٦ \, \sigma_{\bar{س}}$$

والمقدار ٢ هنا هو حجم الخطأ المسموح به (أي خطأ المعاينة)، وهو يمثل مقدار اختلاف متوسط المجتمع (م) - زيادة أو نقصاناً - عن متوسط العينة ($\bar{س}$) وسوف نرمز لهذا الخطأ بالرمز (خ) حيث:

$$خ = ي_{\frac{\alpha}{2}} \, \sigma_{\bar{س}} \qquad (٢ - ١٦)$$

وحيث أن الخطأ المعياري يأخذ إحدى الصورتيْن (٢ - ١) أو (٢ - ٣) فإننا نحصل على:

$$خ = ي_{\frac{\alpha}{2}} \, \frac{\sigma}{\sqrt{ں}} \qquad (٢ - ١٧)$$

أو: $$خ = ي_{\frac{\alpha}{2}} \, \frac{\sigma}{\sqrt{ں}} \sqrt{\frac{ن - ں}{ن - ١}} \qquad (٢ - ١٨)$$

وهنا تلزم الإشارة إلى أننا نستخدم الصورة (٢ - ١٧) في تحديد حجم العينة وذلك في حالة المجتمع غير المحدود أو في حالة اختيار مفردات العينة مع الالتزام بإعادة المفردة المسحوبة في كل مرة قبل سحب المفردة التالية وهو ما يطلق عليه بـ "السحب مع الإرجاع". وأما الصورة (٢ - ١٧) فإنها تستخدم في تحديد حجم العينة وذلك إذا كان المجتمع محدوداً أو في حالة اتباع أسلوب "السحب بدون إرجاع"، أي عدم إعادة المفردة المسحوبة في كل مرة. وبطبيعة الحال فإنه لا يجوز الحديث عن ظروف استخدام كل من الصيغتيْن فيما يتعلق بعلاقة حجم العينة ں بحجم المجتمع ن. أي هل تبلغ ں مقداراً يقل أو يزيد على ٥٪ من حجم المجتمع ن حتى يتسنى استخدام أي من هاتيْن الصيغتيْن، وهو ما سبق الإشارة إليه عند تناول تقدير فترات الثقة. ويرجع ذلك إلى أن هذا هو أمر لا يمكن تحديده حيث أنه لم يتم تحديد حجم العينة بعد. وهنا أيضاً، يجب أن نشير إلى أن هاتيْن الصيغتيْن لتحديد حجم العينة سوف تؤديان إلى نتيجة واحدة تقريباً وذلك في حالة ما إذا كان حجم المجتمع كبيراً. إذ أنه في حالة كبر حجم المجتمع ن فإنه معامل التصحيح تقترب قيمته من الواحد الصحيح بدرجة كافية.

وباستخدام الصورة (٢ - ١٧) يمكننا حساب حجم العينة ں على النحو التالي:

حيث أن: $خ = ي_{\frac{\alpha}{2}} \, \frac{\sigma}{\sqrt{ں}}$

وبتربيع الطرفيْن نحصل على:

$خ^{٢} = ي^{٢}_{\frac{\alpha}{2}} \, \frac{\sigma^{٢}}{ں}$ ، أي أن:

$$ں = \frac{ي^{٢} \sigma^{٢}}{خ^{٢}} \qquad (٢ - ١٩)$$

حيث: $\sigma^{٢}$ تباين المجتمع،

خ حجم الخطأ المسموح به (خطأ المعاينة)،

ي الدرجة المعيارية (ي) حيث: $\phi(\text{ي}) = ١ - \frac{\alpha}{٢}$

وهذا يعني أنه في حالة المجتمع غير المحدود أو السحب مع الإرجاع فإنه يمكننا تحديد الحجم الملائم للعينة باستخدام الصورة (٢ - ١٩) وذلك بعد معرفتنا لدرجة الثقة المطلوبة وحجم خطأ المعاينة المسموح به في عملية التقدير بالإضافة إلى تباين المجتمع ($\sigma^{٢}$).

وأما في حالة المجتمع المحدود أو السحب بدون إرجاع فإنه يمكن تحديد الحجم الملائم للعينة على النحو التالي:

حيث أن: $\text{خ} = \text{ي}_{\frac{\alpha}{٢}} \frac{\sigma}{\sqrt{\text{ں}}} \sqrt{\frac{\text{ن} - \text{ں}}{\text{ن} - ١}}$

فإنه بتربيع الطرفيْن نحصل على:

$$\text{خ}^{٢} = \text{ي}^{٢}_{\frac{\alpha}{٢}} \left(\frac{\sigma^{٢}}{\text{ں}}\right)\left(\frac{\text{ن} - \text{ں}}{\text{ن} - ١}\right)$$

ومنها نحصل على:

$$\text{ں} = \frac{\text{ن}}{١ + \frac{\text{خ}^{٢}(\text{ن} - ١)}{\sigma^{٢}\text{ي}^{٢}}} \qquad \text{(٢ - ٢٠)}$$

هذا وإذا كانت ن كبيرة فإنه تبسيطاً لعملية الحسابات يمكننا اعتبار ن بدلاً من ن – ١ في مقام الكسر المعطى في الصيغة (٢ - ٢٠)، أي أن:

$$\text{ں} = \frac{\text{ن}}{١ + \frac{\text{خ}^{٢}\text{ن}}{\sigma^{٢}\text{ي}^{٢}}} \qquad \text{(٢ - ٢١)}$$

علماً بأن المقصود بـ ي في الصيغتيْن (٢-٢٠)، (٢ - ٢١) وما يليهما من صيغ هو $\text{ي}_{\frac{\alpha}{٢}}$.

وبالرجوع إلى بيانات مثالنا الحالي للإجابة على التساؤل المشار إليه سابقاً نجد أن:

خ = ٢ ، ي = ١,٩٦ ، σ = ٢٥ ، ن = ٥٠٠٠٠

وحيث أن المجتمع محدود في هذا المثال (وإن كانت ن كبيرة جداً) فإنه يمكننا استخدام الصورة (٢ - ٢٠) أو (٢ - ٢١) في تحديد الحجم الملائم للعينة. ونظراً لأن حجم المجتمع كبير فإنه من المتوقَّع ألا يكون هناك اختلاف يذكر بين حجمي العينة المحسوبين باستخدام الصورتيْن (٢ - ٢٠)، (٢ - ٢١). أي أن:

$$ن = \frac{ن}{1 + \frac{خ^{2}(ن - 1)}{\sigma^{2}ي^{2}}}$$

$$= \frac{50000}{1 + \frac{4 \times 49999}{625 \times (1.96)^{2}}} = \frac{50000}{84.297} = 594 \text{ طالباً}$$

أي أن حجم العينة الذي يُتوقع معه ألا يزيد الاختلاف ما بين متوسط درجات مجتمع طلاب الثانوية العامة في مادة اللغة الإنجليزية وبين متوسط درجات الطلاب في العينة المسحوبة من هذا المجتمع عن درجتيْن فقط هو ٥٩٤ طالباً، وثقتنا في ذلك تبلغ ٩٥٪.

ملحوظة: هنا تجدر الإشارة إلى أنه يجب تقريب حجم العينة المتحصَّل عليه بأي صيغة إلى العدد الصحيح التالي وليس وفقاً لقواعد التقريب المعروفة. هذا إن لم يكن تم الحصول على عدد صحيح عند حسابه. وفي حالتنا هذه تم تقريب العدد ٥٩٣,٤ ليكون ٥٩٤.

وإذا أردنا إيجاد حجم العينة باستخدام الصورة (٢ - ٢١) فإننا نحصل على:

$$ن = \frac{ن}{1 + \frac{خ^{2}ن}{\sigma^{2}ي^{2}}}$$

$$ن = \frac{50000}{1 + \frac{4 \times 49999}{625 \times (1.96)^{2}}} = \frac{50000}{84.299} = 594 \text{ طالباً}$$

وهي نفس النتيجة التي حصلنا عليها من قبل.

ويرجع تطابق النتيجتيْن إلى ما سبق الإشارة إليه وهو كبر حجم المجتمع.

وأما إذا اعتبرنا مجتمع الطلاب مجتمعاً لا نهائياً (حيث أن ن كبيرة جداً)، فإنه يمكننا تحديد حجم العينة باستخدام الصورة (٢ - ١٩) وذلك على النحو التالي:

$$ن = \frac{ي^{2}\sigma^{2}}{خ^{2}}$$

$$= \frac{(1.96)^{2} \times 625}{4} = 601 \text{ طالب}$$

أي أن حجم العينة لا يختلف كثيراً في هذه الحالة عن ذلك الحجم المحسوب في حالة اعتبار المجتمع غير محدود، وتفسير ذلك أيضاً هو أن حجم المجتمع في هذا المثال كبير جداً.

وهنا وقبل الخوض في مزيد من التفصيلات تلزم الاشارة إلى نقطة هامة وهي أنه بإيجاد حجم العينة وفقاً للصيغ المختلفة التي تعرضنا لها في حالة تقدير متوسط مجتمع وكذلك الصيغ

التي سوف نتناولها في حالة تقدير النسبة في المجتمع، فإن هذا الحجم إنما يمثل الحد الأدنى الذي يحقق الشروط المطلوبة فيما يتعلق بدرجة الثقة وحجم الخطأ المسموح به في التقديـر. إذ أن أي نقص في حجم العينة عن هذا الحد الأدنى سوف يؤثر سلباً على أي من درجة الثقة أو حجم الخطأ المسموح به. ويتمثل هذا التأثير إما في تقليل درجة الثقة عن الحد المطلوب وذلك بافتراض ثبات حجم الخطأ المسموح به، أو في زيادة حجم الخطأ المسموح عن الحد المطلوب بافتراض ثبات درجة الثقة. ومن الناحية العملية ليس هناك ما يمنع الباحث من أن يأخذ عينة يزيد حجمها عن الحد الأدنى المتحصَّل عليه باستخدام أي من الصيغ المختلفة، إذ أن ذلك سوف يؤثر إيجابياً على أي من درجة الثقة أو حجم الخطأ المسموح به. وهذا يعنى انه إذا أُخذت عينة من مجتمع يزيد حجمها عن الحد الأدنى فإن ذلك يؤدى إلى زيادة درجة الثقة عن الحد المطلوب إذا ظل حجم الخطأ المسموح به على ما هو عليه، أو إلى نقص حجم الخطأ المسموح به عن الحد المطلوب وذلك إذا ظلت درجة الثقة بدون تغيير.

وفي النهاية، فإن الحد الأدنى لحجم العينة هو المؤشر الذي لا يجب على الباحث أن يقل عنه في اختياره لحجم العينة وذلك حتى يضمن تحقق الشروط المطلوبة في عملية التقدير. وإن للباحث أن يزيد حجم العينة عن الحد الأدنى متى كان هذا ممكناً وذلك تحقيقاً لشروط أفضل في عملية التقدير.

وفيما يتعلق بالعلاقة ما بين حجم العينة وبين كل من درجة الثقة وحجم الخطأ المسموح به، تجدر الاشارة إلى ما يلي:

١- إذا كان حجـم المجتمـع كبـيراً جدا فإن الصور الثلاثة (٢ - ١٩)، (٢ - ٢٠)، (٢ - ٢١) تعطي جميعها نتائج متقاربة وذلك فيما يتعلق بحجم العينة. هذا وحجم العينة في حالة السحب مع الإرجاع يكون أكبر منه في حالة السحب مع عدم الإرجاع وذلك في حالة إذا لم يكن حجم المجتمع كبيراً بشكل كاف.

٢- في حالة عدم معرفتنا بأن المجتمع يتبع التوزيع المعتاد فإنه يجب علينا أن نرفع حجم العينة إلى ٣٠ على الأقل إذا نقص حجم العينة المحسوب عن هذا المقدار، وذلك حتى يمكننا تبرير استخدام التوزيع المعتاد في عملية تقدير حجم العينة.

٣- أن جميع الصيغ السابقة لتحديد حجم العينة تعاني من مشكلة عدم معرفة الانحراف المعياري للمجتمع والذي عادة ما يكون مجهولاً. وللتغلب على ذلك يمكن اتباع أحد الأسلوبيْن التاليين:

أ- اختيار عينة عشوائية مبدئية من المجتمع يزيد حجمها عن ٣٠ مفردة ثم استخدام الانحراف المعياري المحسوب من العينة كتقدير للانحراف المعياري للمجتمع.

ب- تقدير المدى الذي تتراوح خلاله قيم الظاهرة موضوع البحث، حيث أن قيمة هذا المدى تساوي – تجريبياً – أربعة أمثال قيمة الانحراف المعياري للمجتمع. أي أن:

المدى = 4σ

وبذلك نجد أن: $\sigma \cong 0.25$ المدى

حيث يمكن استخدام ربع قيمة المدى كتقريب للانحراف المعياري للمجتمع.

٤- من الصور (٢ - ١٩)، (٢ - ٢٠)، (٢ - ٢١) يمكننا استنتاج أن هناك علاقة عكسية بين حجم العينة ومقدار الخطأ المسموح به (خطأ المعاينة). وهذا أمر منطقي، إذ أنه كلما قل حجم خطأ المعاينة زاد حجم العينة والعكس صحيح. وتبلغ هذه العلاقة ذروتها في حالة اتباعنا لأسلوب الحصر الشامل (أي أن: ٮ = ں). عندها نجد أن خطأ المعاينة يتلاشى تماماً أي يساوي الصفر. ويمكن للقارئ استنتــاج ذلك بسهولة إذا وضع ٮ = ں في الصورة (٢ - ١٨).

ويمكننا التحقق من ذلك عملياً إذا ما رجعنا إلى المثال المتعلق بدرجات طلاب الثانوية العامة في مادة اللغة الإنجليزية. فقد وجدنا أن حجم العينة الذي يحقق خطأ معاينة مقداره درجتان هو ٥٩٤ طالباً وذلك بدرجة ثقة ٩٥٪. فإذا افترضنا أن مقدار الخطأ المسموح به هو الآن درجة واحدة فقط وليس درجتيْن. فإنه من المتوقَّع - وباستخدام نفس درجة الثقة - أن يزيد حجم العينة في هذه الحالة عن ٥٩٤ طالباً حيث نجد أن حجم العينة يتحدد كما يلي:

$$ٮ = \frac{ں}{1 + \frac{خ^2 ں}{\sigma^2 ي^2}}$$

$$= \frac{50000}{1 + \frac{1 \times 50000}{625 \times (1.96)^2}} = \frac{50000}{21.825} = 2291 \text{ طالباً}$$

وهو حجم يزيد كثيراً عن نظيره في الحالة التي كان فيها خطأ المعاينة مساو للقيمة ٢. أي أنه كلما قل حجم خطأ المعاينة المطلوب كلما زاد حجم العينة والعكس صحيح.

٥- هناك علاقة طردية بين درجة الثقة المطلوبة وحجم العينة. ويمكننا استنتاج ذلك بسهولة من الصور (٢ - ١٩)، (٢ - ٢٠)، (٢ - ٢١) وهذا أيضاً أمر منطقي إذ أنه من المعروف أنه كلما زاد حجم العينة كلما زادت ثقتنا في صحة النتائج. وتبلغ هذه العلاقة ذروتها في حالة استخدامنا لأسلوب الحصر الشامل، وهو الوضع الذي لا تكون فيه نتائجنا موضع أدنى شك (أي أن درجة الثقة تساوي ١٠٠٪)، اللهم إلا إذا كانت البيانات نفسها هي مصدر الشك.

وللتحقق من ذلك دعنا أيضاً نستخدم بيانات المثال المُشار إليه. فلو افترضنا أن درجة الثقة المطلوبة هي ٩٠٪ وليست ٩٥٪ فإن حجم العينة يتحدد كما يلي:

حيث أن درجة الثقة المطلوبة هي ٩٠٪ فإننا نجد أن:

$$\alpha = \text{٠,١} \quad ، \quad \text{ي}_{\frac{\alpha}{\text{٢}}} = \text{١,٦٤٥}$$

$$\text{ن} = \frac{\text{ن}}{\text{١} + \dfrac{\text{خ}^{\text{٢}}\,\text{ن}}{\text{ي}^{\text{٢}}\sigma^{\text{٢}}}}$$

$$= \frac{\text{٥٠٠٠٠}}{\text{١} + \dfrac{\text{٤} \times \text{٥٠٠٠٠}}{(\text{١,٦٤٥})^{\text{٢}} \times \text{٦٢٥}}} = \frac{\text{٥٠٠٠٠}}{\text{١١٩,٢٥٥}} = \text{٤٢٠ طالب}$$

أي أن إنقاص درجة الثقة المطلوبة من ٩٥٪ إلى ٩٠٪ أدى إلى إنقاص حجم العينة من ٥٩٤ طالبا إلى ٤٢٠ طالبا فقط، وهذا يؤكد صحة العلاقة المشار إليها.

هذا ولعل الامثلة التالية تلقى مزيدا من الضوء على ما سبق أن تناولناه فيما يتعلق بتحديد حجم العينة، مع ملاحظة أنه ما لم يذكر حجم المجتمع صراحة فسوف نعتبره مجتمعا غير محدود.

مثال (٢٨):

ترغب إدارة كلية التجارة بإحدى الجامعات في معرفة متوسط أعمار الطلاب فيها وذلك في حدود نصف سنة زيادة أو نقصاناً عن المتوسط الحقيقي للأعمار وبدرجة ثقة ٩٩٪. فإذا كانت سجلات الكلية لسنوات مضت تفيد بأن أعمار الطلاب فيها تتبع توزيعاً معتاداً بانحراف معياري مقداره ثلاث سنوات. احسب الحد الأدنى لحجم العينة المطلوب إذا علمت أن إجمالي عدد الطلاب في هذه الكلية يبلغ ٩٠٠٠ طالب.

الحـــل:

حيث أن: $\text{ن} = \text{٩٠٠٠} \quad ، \quad \alpha = \text{٣} \quad ، \quad \text{خ} = \text{٠,٥}$

درجة الثقة المطلوبة هي ٩٩٪، أي أن:

$$\alpha = \text{٠,١} \quad ، \quad \text{ي}_{\frac{\alpha}{\text{٢}}} = \text{٢,٥٨}$$

$$\text{ن} = \frac{\text{ن}}{\text{١} + \dfrac{\text{خ}^{\text{٢}}\,\text{ن}}{\sigma^{\text{٢}}\text{ي}^{\text{٢}}}}$$

$$= \frac{\text{٩٠٠٠}}{\text{١} + \dfrac{(\text{٠,٥})^{\text{٢}}(\text{٩٠٠٠})}{(\text{٣})^{\text{٢}}(\text{٢,٥٨})^{\text{٢}}}} = \frac{\text{٩٠٠٠}}{\text{٣٨,٥٥٨}} = \text{٢٣٤ طالباً}$$

مثال (٢٩):

المطلوب هو حل المثال السابق باستخدام درجة ثقة ٩٥٪ وخطأ معاينة مسموح به مقداره ٠,٣٨ سنة. قارن نتائجك في هذا المثال بنتائج المثال السابق مُبدياً ما تراه مناسباً من تعليق.

الحـل:

حيث أن: ن = ٩٠٠٠ ، σ = ٣ ، خ = ٠,٣٨

درجة الثقة المطلوبة هي ٩٥٪، أي أن:

$$\alpha = ٠,٠٥ \quad ، \quad ي_{\frac{\alpha}{٢}} = ١,٩٦$$

$$ن = \frac{ن}{١ + \frac{خ^{٢}ن}{\sigma^{٢}ي^{٢}}}$$

$$= \frac{٩٠٠٠}{١ + \frac{(٠,٣٨)^{٢}(٩٠٠٠)}{(٣)^{٢}(١,٩٦)^{٢}}} = \frac{٩٠٠٠}{٣٨,٥٨٩} = ٢٣٤ \text{ طالباً}$$

وهي نفس النتيجة التي حصلنا عليها في المثال السابق.

وهنا يلزم التنويه إلى أن حجم العينة في المثاليْن واحد وذلك بالرغم من اختلاف درجة الثقة وخطأ المعاينة في المثاليْن. ويمكن تفسير ذلك بأن إنقاص درجة الثقة من ٩٩٪ إلى ٩٥٪ له أثر طردي على حجم العينة، بمعنى أنه يؤدى إلى نقص حجم العينة في مثال (٢٩) عن ٢٣٤ طالب (يلاحظ ان العلاقة ما بين درجة الثقة وحجم العينة هي علاقة طردية). كما أن إنقاص مقدار خطأ المعاينة من ٠,٥ سنة إلى ٠,٣٨ له أثر عكسي على حجم العينة، بمعنى أنه يؤدى إلى زيادة حجم العينة في مثال (٢٩) عن ٢٣٤ طالب (يُلاحَظ ان العلاقة ما بين مقدار خطأ المعاينة المسموح به وحجم العينة هي علاقة عكسية). وثبات حجم العينة في المثاليْن يشير إلى أن الأثر الأول المتمثل في نقصان حجم العينة والأثر الثاني المتمثل في زيادة حجم العينة هما في هذه الحالة أثران متساويان وفي اتجاهيْن متضاديْن ويمكننا استنتاج ذلك إذا ما لاحظنا أن خ، σ هما العاملان اللذان اختلفت قيمة كل منهما في المثاليْن وأنه:

في مثال (٢٨): $\frac{خ^{٢}}{ي^{٢}} = \frac{(٠,٥)^{٢}}{(٢,٥٨)^{٢}} = ٠,٠٣٨$

في مثال (٢٩): $\frac{خ^{٢}}{ي^{٢}} = \frac{(٠,٣٨)^{٢}}{(١,٩٦)^{٢}} = ٠,٠٣٨$

أي أن النسبة: $\frac{خ^{٢}}{ي^{٢}}$ والمستخدمة في مقام الصورة (٢ - ٢١) هي واحدة في المثاليْن. وهذا يفسر النتائج التي حصلنا عليها.

مثال (٣٠):

ترغب شركة ما لإنتاج المعدات الإلكترونية في تقدير متوسط عدد ساعات التشغيل لنوع معين من الصمامات وذلك في حدود ١٥ ساعة تشغيل (زيادة أو نقصاناً) وبدرجة ثقة ٩٥٪. فإذا كان لدى الشركة معلومات سابقة تفيد بأن الانحراف المعياري لعدد ساعات تشغيل هذا النوع من الصمامات يساوي ٣٠ ساعة، احسب حجم العينة الواجب أخذه في عملية التقدير.

الحـل:

حيث درجة الثقة المطلوبة هي ٩٥٪، نجد أن:

$$\alpha = ٠,٠٥ \quad ، \quad ي_{\frac{\alpha}{٢}} = ١,٩٦$$

وحيث أن:

$$خ = ١٥ \quad ، \quad \sigma = ٣٠$$

$$ن = \frac{ي^{٢}\sigma^{٢}}{خ^{٢}}$$

$$= \frac{(١,٩٦)^{٢}(٣٠)^{٢}}{(١٥)^{٢}} = ١٦ \text{ صماماً}$$

وحيث أنه ليس معلوماً لدينا أن عدد ساعات التشغيل يتبع التوزيع المعتاد، فإنه يلزم زيادة حجم العينة إلى ٣٠ صماماً وذلك حتى يُبرّز استخدامنا لهذه الصيغة في تحديد حجم العينة.

أي أن: $ن = ٣٠$ صماماً.

ثانياً: في حالة تقدير النسبة في المجتمع

سوف نوضح فيما يلي كيفية تحديد الحجم الملائم للعينة وذلك في حالة إيجاد تقدير للنسبة في المجتمع متَّبعين في ذلك نفس المنهج المستخدم في حالة تقدير متوسط المجتمع. وتيسيراً على القارئ سوف نستعين في عرضنا لهذا الموضوع بالمثال المشار إليه من قبل والمتعلق بدرجات طلاب الثانوية العامة في مادة اللغة الانجليزية.

إذا افترضنا أن المطلوب هو تقدير نسبة نجاح الطلاب في المجتمع (أي نسبة النجاح بين جميع الطلاب الذين أدوا الامتحان في هذه المادة) وعددهم ٥٠ ألف طالب وذلك في حدود ٥٪ زيادة أو نقصاناً عن النسبة الحقيقية بين طلاب الثانوية العامة ككل وبدرجة ثقة ٩٥٪.

والسؤال الآن هو: ما هو الحد الأدنى لحجم العينة إذا كانت السجلات السابقة تشير بأن نسبة النجاح في هذه المادة تساوي ٦٥٪ تقريباً؟

ويمكننا الإجابة على هذا السؤال وذلك على النحو الموضح فيما يلي:
حيث أن فترة الثقة لنسبة المجتمع (ل) تعرف وفقاً للصورة:

$$\hat{ل} \pm ي_{\frac{\alpha}{٢}} \sigma_{\hat{ل}}$$

وحيث أن درجة الثقة المطلوبة ٩٥٪، فإن: $ي_{\frac{\alpha}{٢}} = ١,٩٦$. أي أن فترة الثقة للنسبة في المجتمع تتحدد كما يلي:

$$\hat{ل} \pm ١,٩٦ \sigma_{\hat{ل}}$$

وهذا يعنى أن فترة الثقة لنسبة نجاح الطلاب في المجتمع لابد وأن تضم نسبة النجاح في العينة ناقصة أو زائدة ١,٩٦ خطأ معياري. وحيث أننا لا نريد لأي من حدى الثقة أن يبتعد عن نسبة العينة بأكثر من ٠,٠٥، فإن ذلك يعنى ما يلي: $٠,٠٥ = ١,٩٦ \sigma_{\hat{ل}}$

والمقدار ٠,٠٥ هنا هو حجم خطأ المعاينة المسموح به، وهو يمثل مقدار اختلاف نسبة المجتمع (ل) - زيادة أو نقصاناً - عن نسبة العينة ($\hat{ل}$)، أي أن:

$$خ = ي_{\frac{\alpha}{٢}} \sigma_{\hat{ل}}$$

وحيث أن الخطأ المعياري يأخذ إحدى الصورتيْن (٢ - ١٢) أو (٢ - ١٣) فإننا نحصل على ما يلي:

$$خ = ي_{\frac{\alpha}{٢}} \sqrt{\frac{لك}{ن}} \qquad (٢ - ٢٢)$$

أو: $$خ = ي_{\frac{\alpha}{٢}} \sqrt{\frac{لك}{ن}} \sqrt{\frac{ن - ن}{ن - ١}} \qquad (٢ - ٢٣)$$

حيث تستخدم الصيغتيْن (٢ - ٢٢)، (٢ - ٢٣) وفقاً للشروط المشار إليها سابقاً عند تناولنا لتحديد حجم العينة في حالة تقدير متوسط المجتمع.
وباستخدام الصورة (٢ - ٢٢) يمكننا حساب حجم العينة على النحو التالي:

حيث أن: $خ = ي_{\frac{\alpha}{٢}} \sqrt{\frac{لك}{ن}}$

فإنه بتربيع الطرفيْن نحصل على:

$خ^{٢} = ي^{٢}_{\frac{\alpha}{٢}} \frac{لك}{ن}$ ، أي أن:

$$ن = \frac{ي^{٢} ل ك}{خ^{٢}} \qquad (٢ - ٢٤)$$

حيث: ل النسبة في المجتمع ، ك = ١ – ل
خ حجم الخطأ المسموح به (خطأ المعاينة)،
ي الدرجة المعيارية (ي) حيث: ϕ(ي) = ١ – $\frac{\alpha}{٢}$

وهذا يعنى أنه في حالة المجتمع غير المحدود أو السحب مع الإرجاع فإنه يمكننا تحديد الحجم الملائم للعينة باستخدام الصورة (٢ - ٢٤) وذلك بعد معرفتنا لدرجة الثقة المطلوبة ومقدار خطأ المعاينة المسموح به في عملية التقدير، هذا بالإضافة إلى معرفة تقدير تقريبي للنسبة في المجتمع.

وأما في حالة المجتمع المحدود أو السحب بدون إرجاع فإنه يمكن تحديد الحجم الملائم للعينة على النحو التالي:
حيث أن:

$$خ = ي_{\frac{\alpha}{٢}} \sqrt{\frac{ل ك}{ن}} \sqrt{\frac{ن - ن}{ن - ١}}$$

فإنه بتربيع الطرفيْن نحصل على:

$$خ^{٢} = ي^{٢}_{\frac{\alpha}{٢}} \frac{ل ك}{ن} \left(\frac{ن - ن}{ن - ١}\right)$$

ومنها نحصل على:

$$ن = \frac{ن}{١ + \frac{خ^{٢}(ن - ١)}{ل ك ي^{٢}}} \qquad (٢ - ٢٥)$$

وإذا كانت ن كبيرة فإنه يمكن وضع الصورة (٨ - ٢٥) في شكل أبسط كما يلي:

$$ن = \frac{ن}{١ + \frac{خ^{٢} ن}{ل ك ي^{٢}}} \qquad (٢ - ٢٦)$$

هذا ويمكننا ملاحظة أن الصور (٢ - ٢٤)، (٢ - ٢٥)، (٢ - ٢٦) هي على الترتيب نفس الصور (٢ - ١٩)، (٢ - ٢٠)، (٢ - ٢١) ولكن بعد استبدال $\sigma^{٢}$ بـ ل ك. ولعل هذه الملاحظة تكون أكثر وضوحاً إذا ما أدركنا أن الخطأ المعياري لكل من الوسط الحسابي والنسبة يأخذ على الترتيب الصورة:

$$\sqrt{\frac{\sigma^{٢}}{ن}} \ ، \ \sqrt{\frac{ل ك}{ن}}$$

وبالرجوع إلى بيانات المثال الحالي للإجابة على السؤال المُشار إليه نجد أن:

$$خ = ٠,٠٥ \quad ، \quad ي^{٢}_{\frac{\alpha}{٢}} = ١,٩٦ ،$$

$$ل = ٠,٦٥ \quad ، \quad ك = ٠,٣٥ \quad ، \quad ن = ٥٠٠٠٠$$

وحيث أن المجتمع محدود في هذا المثال، كما أن ن كبيرة فإنه يمكننا استخدام الصورة (٢ - ٢٦) في تحديد حجم العينة، أي أن:

$$\text{ـم} = \frac{ن}{١ + \frac{خ^{٢} ن}{ل ك ي^{٢}}}$$

$$= \frac{٥٠٠٠٠}{١ + \frac{(٠,٠٥)^{٢} \times ٥٠٠٠٠}{٠,٦٥ \times ٠,٣٥ \times (١,٩٦)^{٢}}} = \frac{٥٠٠٠٠}{١٤٤,٠٢٦} = ٣٤٨ \text{ طالباً}$$

ويمكن للقارئ التحقق من أننا سوف نحصل على نفس النتيجة لو أننا استخدمنا الصورة (٢ - ٢٥) بدلاً من الصورة (٢ - ٢٦). كما أن حجم العينة سيصبح ٣٥٠ طالباً لو اعتبرنا أن المجتمع غير محدود واستخدمنا الصورة (٢ - ٢٤)، وهو لا يختلف كثيراً عن ٢٤٨ وذلك نظراً لكبر حجم المجتمع.

وفي هذا الشأن تجدر الإشارة إلى الملاحظات التالية:

١- إذا كان حجم المجتمع كبيراً جداً فإن الصور (٢ - ٢٤)، (٢ - ٢٥)، (٢ - ٢٦) تعطي جميعها نتائج متقاربة إن لم تكن واحدة.

٢- في حالة عدم معرفتنا بأن المجتمع يتبع التوزيع المعتاد فإنه يجب علينا أن نرفع حجم العينة إلى ٣٠ على الأقل وذلك إذا نقص حجم العينة المحسوب عن هذا المقدار حتى يمكننا تبرير استخدام التوزيع المعتاد في عملية تقدير حجم العينة.

٣- إذا كانت النسبة في المجتمع (ل) غير معلومة، وحيث أننا بصدد تحديد حجم العينة فإنه لا يمكننا استخدام النسبة في العينة ($\hat{ل}$) كتقدير للنسبة في المجتمع (ل) وذلك لأننا لم نقم بعد بأخذ العينة من المجتمع حتى يمكننا معرفة النسبة فيها. وللتغلب على تلك المشكلة يمكننا تقدير الحد الأعلى لحجم العينة (ـم) مهما كانت القيمة الحقيقية للنسبة في المجتمع (ل) وذلك للحصول على وضع أفضل فيما يتعلق بدرجة الدقة المطلوبة ودرجة الثقة المرغوب فيها. ويتحقق ذلك باعتبار أن ل = ك = ٠,٥، حيث يبلغ المقدار (ل ك) أكبر قيمة ممكنة لـه. ويسمى حجم العينة في هذه الحالة بـ "Conservative sample size". وحيث أن هناك علاقة طردية ما بين حجم العينة وبين المقدار (ل ك) {راجع الصيغ (٢ - ٢٤)، (٢ - ٢٥)، (٢ - ٢٦)} فإنه يمكن القول إنه كلما اقتربت قيمة ل من النصف

كلما زاد حجم العينة، وأن حجم العينة (ن) يصل إلى أكبر قيمة ممكنة له في ظل درجة ثـقة معينة في الحالة التي يكون فيها: ل = ك = ٠,٥

ويمكننا التحقق من ذلك باستخدام بيانات المثال الحالي على النحو التالي:

إذا اعتبرنا أن: ل = ك = ٠,٥ فإنه يمكننا الحصول على قيمة ن كما يلي:

$$\text{ن} = \frac{\text{ن}}{١ + \frac{\text{خ}^{٢}\,\text{ن}}{\text{لك}\,\text{ي}^{٢}}}$$

$$= \frac{٥٠٠٠٠}{١ + \frac{(٠,٠٥)^{٢} \times ٥٠٠٠٠}{(٠,٥)(٠,٥)(١,٩٦)^{٢}}} = \frac{٥٠٠٠٠}{١٣١,١٥٤} = ٣٨٢ \text{ طالباً}$$

وهذا يعنى أنه باستخدام درجة ثـقة ٩٥٪ وخطأ مسموح به مقداره (٠,٥) فإن حجم العينة وفقاً للصيغة المستخدمة في إيجاده لا يمكن أن يزيد بحال من الأحوال عن ٣٨٢ طالباً (مهما كانت قيمة ل أو ك). يُلاحَـظ أن حجـم العينــة كان ٣٤٨ طالباً في الحالة التي كــان فيهــا ل = ٠,٦٥.

٤- هناك علاقة عكسية ما بين مقدار الخطأ المسموح به (خطأ المعاينة) وبين حجم العينة، كما أن هناك علاقة طردية ما بين حجم العينة ودرجة الثـقة المطلوبة، وذلك على النحو الموضح في الملاحظتيْن ٤، ٥ عند تناولنا لتحديد حجم العينة في حالة تقدير متوسط المجتمع.

مثال (٣١):

ترغب هيئة حزب "مستقبل وطن" في تقدير نسبة الناخبين الذين يتوقع أن يعطوا أصواتهم لمرشح الحزب في إحدى الدوائر الانتخابية لانتخابات مجلس الشعب وذلك في حدود خطأ في التقدير لا يتجاوز ± ٠,٠٨، ما هو الحد الأدنى لحجم العينة التي يجب استخدامها إذا كانت استطلاعات أخرى تشير إلى أن نسبة المُصوِّتين لهذا المرشح تبلغ تقريبا ٧٠٪ (استخدام درجة الثـقة ٩٩٪).

الحــــل:

حيث أن:

خ = ٠,٠٨ ، ل = ٠,٧ ، ك = ٠,٣

وحيث أن درجة الثـقة المطلوبة هي ٩٩٪، فإن:

$$\alpha = ٠,٠١ \quad ، \quad \text{ي}^{٢}_{\frac{\alpha}{٢}} = ٢,٥٨$$

ونظراً لأن المجتمع غير محدود فإنه يمكننا استخدام الصورة (٢ - ٢٤) في تحديد حجم العينة على النحو التالي:

$$\text{ن} = \frac{\text{ي}^{٢}\text{لك}}{\text{خ}^{٢}} = \frac{(٢,٥٨)^{٢}(٠,٧)(٠,٣)}{(٠,٠٨)^{٢}} = ٢١٩ \text{ ناخباً}$$

مثال (٣٢):

في المثال السابق، إذا كانت نسبة المصّوتين في المجتمع لصالح المرشح المذكور موضع شك أو غير معروفة، احسب الحد الأقصى لحجم العينة على ضوء الظروف المشار اليها.

الحــــل:

حيث أن النسبة في المجتمع غير معلومة فإننا - متَّخذين جانب الجذر - سوف نعتبر أن $\text{ل} = \text{ك} = ٠,٥$ وذلك حتى تبلغ ن أكبر قيمة لها، أي أن:

$$\text{ن} = \frac{\text{ي}^{٢}\text{لك}}{\text{خ}^{٢}}$$

$$= \frac{(٢,٥٨)^{٢}(٠,٥)(٠,٥)}{(٠,٠٨)^{٢}} = ٢٦١ \text{ ناخباً}$$

وهو أكبر من حجم العينة في المثال السابق. أي أنه باستخدام درجة ثقة ٩٩٪ وخطأ معاينة مقداره (± ٠,٠٨) فإن أكبر حجم يمكن أن تصل إليه العينة في ظل الشروط المطلوبة هو ٢٦١ ناخباً.

وبصفة عامة إذا كانت: $\text{ل} = \text{ك} = ٠,٥$ وكان المجتمع غير محدود يمكننا حساب حجم العينة باستخدام الصورة التالية:

$$\text{ن} = \left(\frac{\text{ي}}{\text{٢خ}}\right)^{٢} \qquad (٢ - ٢٧)$$

أي ان الصورة (٢ - ٢٧) هي حالة خاصة من الصورة (٢ - ٢٤) والتي يكون فيها: $\text{ل} = \text{ك} = ٠,٥$

مثال (٣٣):

ترغب إدارة إحدى الكليات في تقدير نسبة النجاح في مرحلة البكالوريوس وذلك في حدود خطأ لا يتجاوز ٠,٠٥ (زيادة أو نقصاناً) وبدرجة ثقة ٩٥٪. ما هو الحد الأدنى لحجم العينة الذي يجب أخذه إذا كانت السجلات السابقة للكلية تشير إلى أن نسبة النجاح في مرحلة البكالوريـوس تساوي ٧٥٪ تقريباً؟ وذلك إذا كان إجمالي عدد الطلاب في مرحلة البكالوريوس هو ٣٥٠٠ طالباً.

الحــــل:

حيث أن:

خ = ٠,٠٥ ، ل = ٠,٧٥ ، ك = ٠,٢٥

ن = ٣٥٠٠ ، $\text{ي}^{2}_{\frac{\alpha}{2}}$ = ١,٩٦

وحيث أن المجتمع محدود وكبير فانه يمكننا تحديد حجم العينة ما يلي:

$$\text{ﻥ} = \frac{\text{ن}}{1 + \frac{\text{خ}^{2}\,\text{ن}}{\text{ل}\,\text{ك}\,\text{ي}^{2}}}$$

$$= \frac{\text{٣٥٠٠}}{1 + \frac{(\text{٣٥٠٠})(\text{٠,٠٥})^{2}}{(\text{٠,٧٥})(\text{٠,٢٥})(\text{١,٩٦})^{2}}} = \frac{\text{٣٥٠٠}}{\text{١٣,١٤٨}} = \text{٢٦٧ طالباً}$$

ويمكن للقارئ أن يتحقق من أنه سوف يحصل على نفس النتيجة إذا استخدم الصورة (٢ - ٢٥) بدلاً من الصورة (٢ - ٢٦).

مثال (٣٤):

يرغب صاحب مصنع للمصابيح الكهربائية في تقدير نسبة المصابيح المعيبة في إنتاج مصنعه وذلك في حدود خطأ مقداره ± ٠,٠٦ عن النسبة الحقيقية في إنتاج المصنع. احسب بدرجة ثقة ٩٠٪ الحد الأدنى لحجم العينة المطلوب إذا علمت أن نسبة المصابيح المعيبة في إنتاج مصنع مشابه لمصنعه تبلغ ١٥٪ تقريباً.

الحــــل:

حيث أن:

خ = ٠,٠٦ ، ل = ٠,١٥ ،

ك = ٠,٨٥ ، $\text{ي}^{2}_{\frac{\alpha}{2}}$ = ١,٨٥

وحيث أن المجتمع غير محدود فانه يمكن تحديد حجم العينة باستخدام الصورة (٢ - ٢٤) وذلك على النحو التالي:

$$\text{ﻥ} = \frac{\text{ي}^{2}\,\text{ل}\,\text{ك}}{\text{خ}^{2}}$$

$$= \frac{(\text{١,٦٤٥})^{2}(\text{٠,١٥})(\text{٠,٨٥})}{(\text{٠,٠٦})^{2}} = \text{٩٦ مصباحاً}$$

مثال (٣٥):

في المثال السـابق، إذا لم تتوافر بيانات عن نسـبة المعيب في إنتاج هذا النوع من المصـابيح، احسـب حجم العينـة المطلوب؟ قـارن بين نتـائجـك في المثـالين مع إبـداء التعليق المناسب.

الحـل:

حيث أن النسبة في المجتمع غير معروفة فإنه - أخذاً بجانب الحذر - سوف نعتبر أن: ل = ك = ٠,٥ وذلك حتى نحصل على أكبر حجم للعينة يحقق على الأقل درجة الدقة المطلوبة، أي أن:

$$ن = \left(\frac{ي}{٢خ}\right)^{٢}$$

$$= \left(\frac{١,٦٤٥}{٢ \times ٠,٠٦}\right)^{٢} = ١٨٨ \text{ مصباحاً}$$

ويلاحَظ في هذا المثال أن حجم العينة يبلغ تقريباً ضعف حجم العينة في مثال (٣٤). ويرجع هذا الاختلاف الكبير في حجم العينة إلى أننا اعتبرنا أن ل = ك = ٠,٥ في مثالنا هذا مما جعل المقدار (ل ك) يبلغ حده الأعلى. هذا في الوقت الذي كانت فيه قيمة ل (وبالتالي ك) تبتعد كثيرا عن ٠,٥ في مثال (٣٤). وهو الأمر الذي يجعل المقدار (ل ك) يقترب من حده الأدنى في هذه الحالة. ولعل الصورة تكتمل وضوحاً إذا ما أدركنا أن حجم العينة ن يتناسب طردياً مع المقدار (ل ك).

وفي نهاية هذا الفصل نود أن نشير إلى أننا سوف نكتفي بهذا القدر في دراستنا لنظرية التقدير الإحصائي تاركين في هذا الشأن جوانب عديدة إلى مراحل متقدمة في دراستنا لعلم الإحصاء إن شاء الله.

تمارين
الفصل الثاني

١- مجتمع مكون من أربعة طلاب، فإذا كانت درجاتهم في مادة المحاسبة هي: ١٨، ١١، ١٠، ١٧ المطلوب:

أ- حساب الوسط الحسابي والانحراف المعياري للمجتمع.

ب- إيجاد توزيع المعاينة لمتوسط عينة حجمها طالبان مسحوبة من هذا المجتمع.

ج- حساب الوسط الحسابي والخطأ المعياري لتوزيع المعاينة للمتوسط.

٢- إذا كان مجتمع البحث يتكون من ٣٥٠٠ مفردة بمتوسط ١٥٠ وانحراف معياري ٩٠، احسب التوقُّع والخطأ المعياري لتوزيع المعاينة للوسط الحسابي للعينة عندما يكون حجم العينة:

أ- ١٢٠ مفردة ب- ٢٥٠ مفردة

٣- إذا كان متوسط نسبة الذكاء لطلاب إحدى الكليات هو ١٢٠ درجة والانحراف المعياري ٢٠ درجة. فإذا سحبت عينة عشوائية مكونة من ٢٥٦ طالباً من هذه الكلية المطلوب:

أ- إيجاد الوسط الحسابي والخطأ المعياري لتوزيع المعاينة للمتوسط.

ب- احسب احتمال ألا تقل نسبة الذكاء لأفراد العينة عن ١١٠ درجة.

٤- مجتمع معتاد وسطه الحسابي ٢٠٠ وانحرافه المعياري ٥٠، احسب احتمال أن متوسط عينة حجمها ٢٥ مفردة مسحوبة من هذا المجتمع سوف يزيد عن ٢١٥.

٥- مستخدماً بيانات التمرين (٢)، احسب احتمال أن يقل متوسط العينة عن ١٢٠.

٦- سُحبت عينة عشوائية مكونة من ٣٦ مصباحاً كهربائياً من إنتاج مصنع ما للمصابيح الكهربائية. فإذا علمت أن عمر التشغيل للمصابيح الكهربائية المنتَجة في المصنع يتبع توزيعاً معتاداً متوسط ٣٠٠ ساعة وانحرافه المعياري ٤٥ ساعة، المطلوب:

أ- احتمال أن يزيد متوسط العينة عن ٢٨٠ ساعة.

ب- احتمال أن يتراوح متوسط العينة ما بين ٢٥٠، ٣١٠ ساعة.

٧- إذا كان توزيع أجور العمال في أحد المصانع يتبع توزيعاً معتاداً انحرافه المعياري ١٥ جنيهاً. ويأخذ عينة عشوائية من عمال المصانع والبالغ عددهم ١٢٠٠ عاملاً وجد أن متوسط أجور العمال في العينة يساوي ٢٥٠ جنيهاً. أوجد بفترة ثقة ٩٥٪ متوسط أجور العمال في المصنع إذا كان:

أ- حجم العينة ٨٠ عاملاً ب- حجم العينة ٢٠ عاملاً

٨- تهتم كلية التجارة بإحدى الجامعات بدراسة متوسط عدد الساعات التي يقضيها الطالب في مذاكرته لدروسه. ولدراسة هذه الظاهرة تم اختيار عينة من طلاب الكلية. وبتسجيل الوقت الذي يقضيه الطالب في المذاكرة وجد أن متوسط عدد ساعات المذاكرة أسبوعياً هـــو ٢٤ ساعة، كما أن الانحراف المعياري هو ٤ ساعات. فإذا علمت أن الوقت الذي يقضيه الطلاب في مذاكرتهم لدروسهم يتبع توزيعاً معتاداً مجهولةٌ معلماته، أوجد فترة الثـقة ٩٩٪ لمتوسط ساعات المذاكرة أسبوعياً وذلك إذا كان:

أ- حجم العينة مائة طالب **ب-** حجم العينة ٢٥ طالب

٩- في دراسة إحصائية لمعرفة معدل استهلاك نوع معين من السيارات للوقود تم اختيار عشر سيارات من هذا النوع من السيارات وتم تشغيلها وسجلت المسافة التي تقطعها كل سيارة لكل لتر من البنزين. فإذا كانت المسافات التي قطعتها السيارات العشر كما يلي:

١٤ ٩ ١٥ ١١ ١٣ ١٢ ١٣ ٨ ١٥ ١٠

المطلوب:

إيجاد فترة الثـقة ٩٥٪ لمتوسط المسافة التي تقطعها السيارة المنتجة من هذا النوع لكل لتر واحد من البنزين وذلك بافتراض أن معدل استهلاك هذا النوع من السيارات يتبع التوزيع المعتاد.

١٠- إذا كان متوسط مرتبات الموظفين لعينة مكونة من ٦٠ موظفاً في إحدى الشركات هو ٣٥٠ جنيهاً وكان الانحراف المعياري لمرتبات الموظفين في هذه الشركة هو ٣٠ جنيهاً. أوجد فترة الثـقة لمتوسط مرتبات الموظفين في هذه الشركة مستخدماً في ذلك:

أ- فترة ثـقة ٩٥٪ **ب-** فترة ثـقة ٩٩٪

١١- في دراسة لتقدير متوسط دخل الأسرة اليومي في إحدى المحافظات أُختيرت عينة مكونة من ٢٠٠ أسرة من هذه المحافظة فوجد أن الوسط الحسابي والانحراف المعياري لدخول الأسر في العينة هما ٢٥٠ جنيهاً، ٢٠ جنيهاً على الترتيب. أوجد فترة ثـقة ٩٥٪ لمتوسط دخل الأسرة اليومي في المحافظة.

١٢- في التمرين (٧)، إذا افترضنا أن توزيع أجور العمال في المصنع لا يتبع التوزيع المعتاد، أوجد فترة الثـقة لمتوسط أجور العمال في المصنع مستخدماً في ذلك:

أ- فترة ثـقة ٩٥٪ **ب-** فترة ثـقة ٩٠٪

١٣- سُحبت عينة عشوائية حجمها ٦٠ مفردة ومتوسطها ٣٠ من مجتمع معتاد انحرافه المعياري ١٢. كما سحبت عينة عشوائية أخرى حجمها ٢٠ مفردة ومتوسطها ٣٤ من مجتمع معتاد آخر انحرافه المعياري ١٠. أوجد فترة الثـقة ٩٠٪ للفرق بين متوسطي المجتمعيْن ثم فسِّر النتيجة التي تحصل عليها.

١٤- أُخذت عينة مكوَّنة من ٢٠٠ مصباح كهربائي من النوع (أ) فكان الوسط الحسابي والانحراف المعياري لأعمار المصابيح هما ١٥٠٠ ساعة، ١٠٠ ساعة على الترتيب. كما سحبت عينة أخرى حجمها ١٦٠ مصباحا كهربائيا من النوع (ب) فكان الوسط الحسابي والانحراف المعياري لأعمار المصابيح في هذه العينة هما على الترتيب ١١٠٠ ساعة، ٩٠ ساعة. قدر بفترة ثقة ٩٩٪ الفرق بين الوسطيْن الحسابييْن لأعمار المصابيح في النوعيْن أ، ب إذا كان من المعروف أن عمر المصابيح يتبع التوزيع المعتاد في كل منهما.

١٥- إذا افترضنا في التمرين السابق أن عمر المصابيح لا يتبع التوزيع المعتاد، قدر بفترة ثقة ٩٥٪ الفرق بين الوسطيْن الحسابييْن لأعمار المصابيح في النوعين أ، ب.

١٦- سحبت عينة عشوائية حجمها ١٠ مفردات بمتوسط ٦٠ وذلك من مجتمع معتاد مجهول تباينه ووسطه الحسابي. كما سحبت عينة عشوائية اخرى حجمها ١٤ مفردة بمتوسط ٥٧ وذلك من مجتمع معتاد آخر مجهول أيضا تباينه ووسطه الحسابي. أوجد فترة الثقة ٩٥٪ للفرق بين متوسطي المجتمعيْن إذا علمت ان الانحراف المعياري للعينة الأولى هو ٩ وللعينة الثانية هو ٦. ثم فسِّر ما تحصل عليه من نتائج.

١٧- المطلوب حل التمرين (١٤) إذا كان حجم كل من العينتيْن هو ١٢، ١٥ مصباحاً على الترتيب.

١٨- أُختيرت عينة حجمها ٨٠ طالباً من بين طلاب إحدى كليات التجارة، أوجد احتمال أن تزيد نسبة الطلاب المدخنين في العينة عن ١٥٪ إذا علمت أن نسبة المدخنين بين طلاب هذه الكلية تبلغ ١٧٪ تقريباً.

١٩- في دراسة لمعرفة نسبة المصابين بمرض معين في إحدى المحافظات أُخذت عينة مكونة من ٥٠٠ شخص فوجد أن بينهم ٢٠ شخصاً مصاباً بهذا المرض. أوجد فترتي الثقة ٩٥٪، ٩٩٪ لنسبة المصابين بهذا المرض في المحافظة ككل.

٢٠- ترغب هيئة لاستطلاع الرأي العام في أن تقدر بدرجة ثقة ٩٥٪ نسبة الناخبين في المجتمع الذين يُتوقع أن يُعطوا أصواتهم لمرشح معين. فقامت بأخذ عينة حجمها ١٢٠ شخصاً حيث وجدت أن منهم ٨٤ شخصاً يؤيدون هذا المرشح. أوجد فترة الثقة لنسبة الناخبين المؤيدين لهذا المرشح في المجتمع.

٢١- يرغب صاحب مصنع للمصابيح الكهربائية في تقدير نسبة المصابيح المعيبة في إنتاج مصنعه فقام باختيار عينة حجمها ٥٩٠ مصباحاً كان من بينها ثلاثة مصابيح معيبة. أوجد بفترة ثقة ٩٩٪ نسبة المصابيح المعيبة في إنتاج المصنع ككل.

٢٢- أُخذت عينة عشوائية حجمها ١٠٠ طالب من كل من كليتي التجارة والآداب بإحدى الجامعات فوجد أن نسبة الطلاب الحاصلين في العام الدراسي السابق على تقدير عام جيد فأقل هي ٩٠٪ في عينة كلية التجارة و٨٥٪ في عينة كلية الآداب. أوجد بفترة ثقة ٩٥٪ تقديرا للفرق بين نسبة الطلاب الحاصلين على تقدير عام جيد فأقل في مجتمعي الكليتيْن.

٢٣- يرغب أحد الباحثين في تقدير متوسط الأجر اليومي للعمال في إحدى شركات القطاع الخاص وذلك في حدود ٢٥ جنيهاً زيادة أو نقصاناً عن المتوسط الحقيقي للأجور اليومية للعمال في هذه الشركة. فلو افترضنا أن الباحث يعلم من خبرته الشخصية أن توزيع الأجور في الشركة يتبع توزيعاً معتاداً انحرافه المعياري ١٥ جنيهاً. أوجد الحد الادنى لحجم العينة الواجب اختيارها وذلك بدرجة ثقة ٩٥٪.

٢٤- في التمرين السابق، إذا افترضنا عدم توافر بيانات عن الانحراف المعياري للأجور في الشركة وأن البيانات المتاحة كانت على النحو التالي: أعلى أجر = ٢٥٠ جنيها، أدنى أجر = ١٣٠ جنيها.
وضح كيف يمكن استخدام البيانات المتاحة في تحديد حجم العينة الملائم.

٢٥- ترغب شركة في تقدير متوسط عدد ساعات التشغيل لنوع معين من المصابيح الكهربائية وذلك في حدود ١٥ ساعة زيادة أو نقصاناً وبدرجة ثقة ٩٠٪. فإذا عُلِم أن الانحراف المعياري لعدد ساعات التشغيل هو ٢٥ ساعة، أوجد حجم العينة الملائم الواجب اختياره.

٢٦- يرغب صاحب مزرعة لتربية الدواجن في تقدير متوسط وزن الدجاجة في مزرعته وذلك في حدود ١٢٥ جم زيادة أو نقصاناً عن المتوسط الحقيقي لوزن الدجاجة في المزرعة ككل. فإذا افترضنا أنه معلومٌ لدى صاحب المزرعة من واقع خبرته السابقة ان الانحراف المعياري لوزن الدجاجة في المزرعة يبلغ تقريبا ٠,٣ كجم. أوجد الحد الأدنى لحجم العينة الواجب اختيارها وذلك بدرجة ثقة ٩٥٪.

٢٧- المطلوب إعادة حل التمرين السابق وذلك باعتبار كل من الحالتيْن التاليتيْن على حدا وكذلك باعتبارهما معاً:

أ- ليست هناك معلومات متاحة لدى صاحب المزرعة عن الانحراف المعياري لوزن الدجاجة في مزرعته وإنما تمثلت المعلومات المتاحة فيما يلي:
أكبر وزن للدجاجة = ٢,٥ كجم، أصغر وزن للدجاجة = ١,٢ كجم.

ب- إجمالي عدد الدجاج في مزرعته يبلغ ٥٠٠٠ دجاجة.

٢٨- يرغب صاحب مصنع للملابس الجاهزة في تقدير نسبة الوحدات المعيبة في إنتاج مصنعه وذلك في حدود ± ٠,٠٥ وبدرجة ثقة ٩٩٪. أوجد الحد الأدنى لحجم العينة اللازم أخذه إذا كانت الخبرة السابقة لصاحب المصنع تشير إلى أن نسبة المعيب في إنتاج الملابس الجاهزة في مصنع مشابه تبلغ تقريبا ٤٪.

٢٩- يرغب مركز الدراسات السياسية والاقتصادية بمؤسسة الأهرام في تقدير نسبة الناخبين الذين يُتوقَّع أن يعطوا أصواتهم لصالح أحد الأحزاب في إحدى المحافظات على ألا يتجاوز حجم الخطأ ٠,٠٢ زيادة أو نقصاناً عن النسبة الحقيقية في المحافظة. ما هو الحد الأدنى لحجم العينة الواجب اختيارها إذا علمت ان الدراسات الأولية تفيد بأن نسبة المؤيدين لهذا الحزب في المحافظة تبلغ تقريباً ٤٥٪، استخدم في ذلك درجة ثقة ٩٥٪.

٣٠- في التمرين السابق، إذا علمت أن إجمالي عدد الناخبين في المحافظة يبلغ ١٥ ألف ناخب، أوجد بدرجة ثقة ٩٠٪ الحد الأدنى لحجم العينة الواجب اختيارها.

٣١- ترغب إدارة البحوث بإحدى الشركات في عمل دراسة لتحديد نسبة الذين يفضلون منتَجاً معيناً من منتجات هذه الشركة وذلك في حدود خطأ لا يتجاوز حجمه ± ٠,٠٣ عن النسبة الحقيقية في المجتمع وبدرجة ثقة ٩٥٪. أوجد الحد الأدنى لحجم العينة الواجب اختيارها إذا كان معلوما لدى إدارة البحوث أن هذه النسبة في المجتمع تبلغ تقريباً ٣٨٪. وبافتراض عدم توافر معلومات عن هذه النسبة في المجتمع أوجد الحد الأقصى لحجم العينة اللازم. قارن بين نتائجك في الحالتيْن مُبدِياً ما تراه مناسباً من تعليق.

٣٢- في التمرين ٢٩، وبافتراض عدم توافر أية معلومات عن نسبة المؤيدين للحزب في المحافظة، ما هو حجم العينة الواجب اختيارها في هذه الحالة؟ قارن بين النتيجة في الحالتيْن مع إبداء التعليق المناسب.

الفصل الثالث

اختبار الفرضيات الإحصائية

مقدمـــــة:

تناولنا في الفصل السابق من هذا الكتاب أحد أساليب الاستدلال الإحصائي وهو التقدير الإحصائي. وفي هذا الفصل سوف نتعرض لأسلوب آخر يحتل مكانة هامة في عملية اتخاذ القرارات المبنية على تطبيق الأساليب الإحصائية والتي بدونها لا نستطيع اتخاذ قرار موثوق فيه حول أمر من الأمور. وهذا الأسلوب هو اختبار الفرضيات. واختبار الفرضيات - شأنه في ذلك شأن التقدير الإحصائي - إنما يتعلق باتخاذ قرارات حول معلمات المجتمعات موضع الدراسة وذلك من واقع النتائج المتعلقة بالعينات المسحوبة من تلك المجتمعات.

وسوف نرى في هذا الفصل أن معظم المفاهيم المستخدمة في عملية التقدير الإحصائي سوف تُستخدم في اختبار الفرضيات، مثل توزيع المعاينة والخطأ المعياري والاحتمالات المرتبطة بفترة الثقة، إلى غير ذلك من المفاهيم. وفي واقع الأمر، فإن اختبار الفرضيات والتقدير الإحصائي يؤديان إلى نتيجة واحدة، إلا أن اختبار الفرضيات هو الأكثر استخداماً والأعم تطبيقاً في مجال اتخاذ القرارات.

بعض المفاهيم والمصطلحات في اختبار الفرضيات:

عند تناولنا لبعض المفاهيم والمصطلحات المستخدمة في اختبار الفرضيات والتي لم يتم التعرض لها من قبل، دعنا ندعم ذلك بمثال قد يجعل القارئ أكثر فهماً وأعمق إدراكاً لتلك المفاهيم والمصطلحات.

مثال:

لنفترض أن جهة حكومة تلقت عدداً من شكاوى المواطنين فحواها أن وزن رغيف الخبز المنتَج بواسطة أحد المخابز يختلف - من وجهة نظر البعض - عن الوزن المقرر له وهو ١٨٠ جراماً. كما يرى البعض الآخر أن وزن الرغيف يقل عن الوزن المقرر له. وللتحقق من شكاوى المواطنين المستهلِكين لهذا النوع من الخبز، قامت الجهة الحكومية بأخذ عينة من إنتاج هذا المخبز حجمها ١٠٠ رغيف فوجدت أن متوسط وزن الرغيف في العينة هو ١٧٨ جراماً وأن الانحراف المعياري لوزن الرغيف هو ٩ جرامات. كيف يمكن للجهة الحكومية التحقق من صحة شكاوى المواطنين على أن تكون واثقة من قرارها الذي سوف تتخذه بنسبة ٩٥٪؟

والآن لعله من الملائم أن نتناول بشيء من التفصيل تلك المفاهيم والمصطلحات وذلك تطبيقاً على مثالنا الحالي.

الفرضية الإحصائية (فرضية العدم): Null Hypothesis

نظراً لعدم المعرفة الحقيقية لقيم معلمات المجتمع والتي نكون بصدد اتخاذ قرارات بشأنها فإننا نفترض قيمة لمعلمة المجتمع. وهذا الافتراض يحتمل الصواب كما يَحتِمل الخطأ. وللوصول إلى قرار بشأن هذا الافتراض فإنه عادة ما يتم أخذ عينة من هذا المجتمع. وبناءً على نتائج هذه العينة يمكننا الحكم على صحة أو خطأ هذا الافتراض وذلك في ظل درجة ثقة مطلوبة. هذا ويجب أن تكون الفرضية منطقية وقابلة للاختبار الإحصائي.

وفي مثالنا الحالي فإن الفرضية هي: أن متوسط وزن الرغيف في إنتاج المخبز هو ١٨٠ جراماً. وهذه الفرضية تسمى بفرضية العدم أو الأوّلية أو الصفرية. ومن باب التيسير على القارئ سوف يقتصر تناولنا لمسمى هذه الفرضية على فرضية العدم ويرمز لها بالرمز ف$_{٠}$.

وبشكل عام، تأخذ فرضية العدم الصورة:

ف$_{٠}$: م = م$_{٠}$ حيث م هي قيمة مُفترَضة لمتوسط المجتمع في فرضية العدم.

وفي المثال الحالي تأخذ هذه الفرضية الصورة التالية:

ف$_{٠}$: م = ١٨٠ جراما

وفي حالة قبول هذه الفرضية مستخدمين في ذلك النتائج المتحصَّل عليها من العينة، فإن قرار الجهة الحكومية هو أن متوسط وزن الرغيف في إنتاج هذا المخبز هو ١٨٠ جراماً وذلك بثقة ٩٥٪، وأن شكاوى المواطنين هي محض ادعاء باطل.

الفرضية البديلة: Alternative Hypothesis

في حالة رفض فرضية العدم فإن ذلك يعني قبولنا لفرضية بديلة مصاحبة لها يرمز لها بالرمز (ف$_{١}$). وهي الفرضية التي تُقبل إذا رُفضت فرضية العدم. ويمكن أن تُصاغ الفرضية البديلة بإحدى طريقتيْن: غير متجهة Non-directional أو متجهة Directional. والفرضية البديلة غير المتجهة تفيد بأنه ما لم تكن معلمة المجتمع مساوية للقيمة المفترضة في فرضية العدم (ولتكن م$_{٠}$ على سبيل المثال) فـإن قيمة معلمة المجتمع تختلف عن م$_{٠}$ بغض النظر عن كونها أصغر أو أكبر منها. وأما الفرضية البديلة المتجهة فتهتم بكون قيمـــة معلمة المجتمع أصغر أو أكبر من القيمة التي تفترضها فرضية العدم ويمكن التعبير عن ذلك كما يلي:

الفرضية البديلة غير المتجهة: ف$_{١}$: م ≠ م$_{٠}$

وهي تعني أن متوسط المجتمع لا يساوي (يختلف عن) القيمة م$_{٠}$.

الفرضية البديلة المتجهـــــة: ف$_{١}$: م < م$_{٠}$

وهي تعني أن متوسط المجتمع أقل من القيمة م$_{٠}$.

أو ف$_١$: م > م.

وهي تعني أن متوسط المجتمع أكبر من القيمة م.

وفي مثالنا الحالي فإن الفرضية البديلة غير المتجهة تتمثل في حالة شكوى بعض المواطنين من أن متوسط وزن رغيف الخبز يختلف عن ١٨٠ جراماً.

أي أن: ف$_١$: م ≠ ١٨٠ جراماً

وأما الفرضية البديلة المتجهة فتتمثل في حالة شكوى البعض الآخر من المواطنين والتي يدَّعون فيها بأن متوسط وزن الرغيف يقل عن ١٨٠ جرماً.

أي أن: ف$_١$: م < ١٨٠ جراماً

وهنا تجدر الإشارة إلى أن الفرضية البديلة لا تخضع للاختبار الإحصائي إذ أن فرضية العدم هي وحدها التي تخضع للاختبار الإحصائي. وتكمن أهمية الفرضية البديلة في أنها تحدد نوع الاختبار الإحصائي من حيث كونه بذيل واحد One - Tailed Test أو بذيلَيْن Two - Tailed Test.

هذا وفي حالة الفرضية البديلة غير المتجهة يكون الاختبار الإحصائي بذيلَيْن وأما في حالة الفرضية البديلة المتجهة فإن الاختبار الإحصائي يكون بذيل واحد. وسوف تتضح هذه النقطة بشكل أكبر فيما بعد.

المختبِر الإحصائي: Test Statistic

يهدف الاختبار الإحصائي إلى اتخاذ قرار حول ما إذا كانت فرضية العدم مقبولة أم مرفوضة. ويُستخدم الاختبار الإحصائي في ذلك ما يسمى بالمختبِر الإحصائي (أو إحصاء الاختبار) والذي يختلف في تحديده باختلاف معلمة المجتمع موضع الاختبار. والمختبِر الإحصائي هو متغير عشوائي ذو توزيع احتمالي، وهو يصف العلاقة بين القيم النظرية لمعلمات المجتمع وقيم تلك الإحصاءات Statistics المناظرة المحسوبة من العينات المأخوذة من المجتمع. ويأخذ المختبِر الإحصائي الصورة التالية:

$$\textbf{المُختبِر الإحصائي} = \frac{\text{قيمة إحصاء العينة – قيمة معلمة المجتمع في فرضية العدم}}{\text{الانحراف المعياري لإحصاء العينة}}$$

حيث يتم حساب قيمة المختبِر الإحصائي بافتراض صحة فرضية العدم.

ملحوظة: معلمة المجتمع Parameter يناظرها إحصاء العينة Statistic. على سبيل المثال، الوسط الحسابي للمجتمع (م) هو معلمة من معلمات المجتمع، بينما متوسط العينة ($\bar{س}$) يسمى "إحصاء العينة"

هذا ويتم اتخاذ القرار بمقارنة قيمة المختبِر الإحصائي المحسوبة من العينة بقيمته المستخرجة من توزيعه الاحتمالي والذي عادة ما يتبع توزيعاً معروفاً مثل التوزيع المعتاد أو توزيع ت، إلى غير ذلك من التوزيعات المختلفة.

الخطأ من النوع الأول والخطأ من النوع الثاني وقوة الاختبار:

Type One and Type Two Errors and Power of the Test

إن اتخاذ قرار بشأن قبول أو رفض فرضية العدم ليس أمر قاطعاً لا يقبل الشك. وذلك لأن اتخاذ مثل هذا القرار إنما يكون بدرجة ثـقة معينــة مرغوبٌ فيها. ولا يمكن لهذه الدرجة أن تبلغ حد الكمــال (أي ١٠٠٪) طالما أن إجراءات اتخاذ اتخاذ القرار مبنيَّة على نتائج عينة مسحوبة من المجتمع وليس على دراسة المجتمع بأكمله. ومن هنا، فإن عدم رفض فرضية العدم لا يعني بالضرورة أنها صحيحة، فقد يرجع ذلك إلى عدم وجود الأدلة الكافية من بيانات العينة لرفض فرضية العدم. وعلى الجانب الآخر، فإن رفض فرضية العدم لا يعني بالضرورة أنها خاطئة.

مما تقدم يتضح لنا أن اتخاذ قرار بشأن فرضية العدم يتمثل في واحدة من الحالات الأربع التالية:

- عدم رفض فرضية العدم وهي صحيحة، وهو قرار صائب.
- أن تُرفض فرضية العدم في حين أنها صحيحة. وهنا يقع متخذ القرار في ما يُعرف بـ **الخطأ من النوع الأول** ويُرمز له بالرمز α.
- عدم رفض فرضية العدم في حين أنها خاطئة. وهنا يقع متخذ القرار في ما يُعرف بـ **الخطأ من النوع الثاني** ويرمز له بالرمز β.
- أن تُرفض فرضية العدم وهي خاطئة، وهو قرار صائب.

ويمكن تلخيص تلك الحالات في الجدول (٣ - ١).

الجدول (٣ - ١)

الخطأ من النوع الأول والخطأ من النوع الثاني

حالة فرضية العدم / القرار	صحيحة	خاطئة
رفـــــض	خطأ من النوع الأول (α)	قرار صائب
قبـــــول	قرار صائب	خطأ من النوع الثاني (β)

هذا وتمثل α الحد الأقصى لاحتمال ارتكاب الخطأ من النوع الأول، وهي نفسها التي يُشار إليها بمستوى المعنوية. ويُقصد بكلمة "معنوية" أن الفرق بين القيمة النظرية للمعلمة في المجتمع وبين القيمة الناتجة والمحسوبة لهذه المعلمة من العينة هو فرق جوهري وحقيقي ولا يرجع إلى عوامل الصدفة Chance. إذ أنه من الناحية العملية، وفي الأمور المتعلقة بالمقارنات، فإننا نلاحظ أن ظهور الفروق هو الأمر المألوف. ولكن من وجهة النظر الإحصائية البحتة فإننا كثيراً ما نجد أن مثل تلك الفروق إنما نشأت نتيجة لعوامل الصدفة وأنها فروق ليست جوهرية أو حقيقية. ومن هنا تأتى أهمية اختبار الفرضيات لكي تحسم لنا مثل هذه الأمور، ولكي توضح ما إذا كانت تلك الفروق هي فروق حقيقية واقعية أم غير ذلك. ومستوى المعنوية α يحدده الباحث لنفسه قبل عملية جمع البيانات من عينة الدراسة. ومن القيم الشائعة لمستوى المعنوية α: ٠,٠٥، ٠,٠١ وأما المكمّل لمستوى المعنوية ١ – α فهو ما يُسمَّى بـ "معامل الثقة" والذي أشرنا اليه من قبل عند دراستنا لنظرية التقدير في الفصل السابق من هذا الكتاب. وكما سنرى فيما بعد في هذا الفصل فإن ١ – α تمثل منطقة عدم رفض فرضية العدم بينما تمثل α منطقة رفض هذه الفرضية. وهنا يلزم التنويه إلى أنه يمكن التعبير عن مستوى المعنوية في صورة نسبة أو نسبة مئوية (مثلاً: ٠,٠٥ أو ٥٪).

وقد تعرضنا في الفصل السابق إلى مفهوم α، ومع ذلك فإنه من الملائم أن نتناول هذا المفهوم هنا من وجهة نظر عملية اختبار الفرضيات. إذ أن مستوى المعنوية ٠,٠٥، على سبيل المثال، يعني أنه إذا تكررت التجربة عدداً كبيراً جداً من المرات فإنه من المحتمل أن نرفض فرضية العدم وهي في الواقع صحيحة وذلك بمعدل ٥ مرات في كل ١٠٠ مرة. أي أن احتمال رفضنا لفرضية العدم وهي صحيحة هو ٠,٠٥ واحتمال عدم رفضنا لها، أي اتخاذ قرار صائب بعدم رفضها هو ٠,٩٥.

وأما الخطأ من النوع الثاني فتكون القيمة القصوى لاحتمال ارتكابه مساوية β (وتُقرأ بيتا). ويختلف هذا النوع من الخطأ عن الخطأ من النوع الأول في أنه يمكن للباحث أن يحدد لنفسه قيمة α قبل عملية جمع البيانات وهو ما لا يحدث عادة بالنسبة لـ β. وتجدر الإشارة هنا إلى أن العلاقة بين الخطأ من النوع الأول والخطأ من النوع الثاني هي علاقة عكسية، أي كلما قلّت α زادت β (ولكن ليس بنفس المقدار)، والعكس صحيح. وهنا، يحاول الباحث أن يحدد القيمة المناسبة لـ α والتي تجعل قيمة β أقل ما يمكن والحالة الوحيدة التي يمكن فيها إنقاص كل من α و β هي زيادة حجم العينة محل الدراسة.

ومن الواضح هنا أنه كلما زاد حجم العينة كلما قلت قيمة كل من α و β وأن الحالة القصوى هي أن حجم العينة يزداد حتى يصل إلى حجم المجتمع (أي يتحول أسلوب الدراسة من المعاينة إلى الحصر الشامل). وهنا - وبطبيعة الحال - فإن: α = β = صفر، وهو أمر واقعي إذ أن دراسة المجتمع بأكمله لا تعتريها أية شكوك (أو عدم ثقة) فيما يتعلق بعملية

اتخـــاذ القرارات، اللهم إلا إذا كان مصدر الشك هو البيانات نفسها. هذا وكيفية الموازنة بين قيمتي α و β الملائمتيْن هو أمر يخرج عن نطاق هذا الكتاب. وهنا تجب الإشارة إلى أن قيمة α يتم تحديدها بافتراض صحة فرضية العدم، وأما قيمة β فيتم تحديدها بافتراض صحة الفرضية البديلة.

وأما **قوة الاختبار** فهي قدرة الاختبار على رفض فرضية العدم عندما تكون في حقيقة الأمر خاطئة. ويمكن التعبير عنها في صورة احتمال تعتمد قيمته وبشكل صريح على احتمال ارتكاب الخطأ من النوع الثاني وذلك كما يلي:

قوة الاختبار = ١ – احتمال الخطأ من النوع الثاني

= ١ – β

منطقة الرفض ومنطقة عدم الرفض:
Rejection and Nonrejection Regions

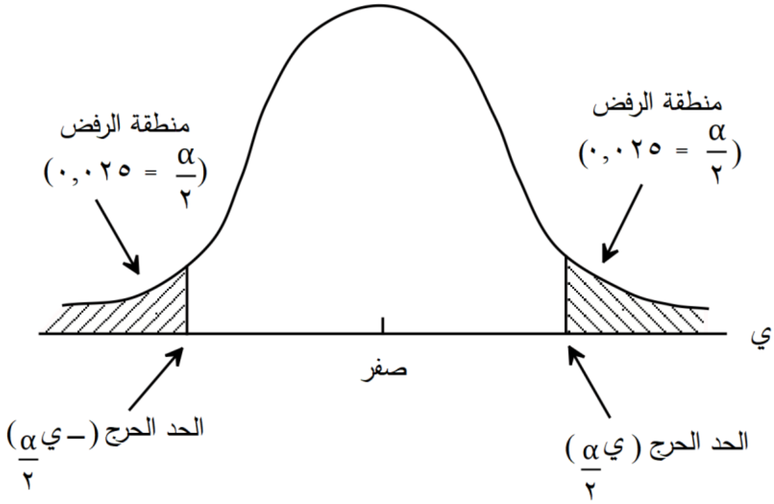
بدايةً تلزم الإشارة إلى أن الحديث عن الرفض أو عدم الرفض إنما يتعلق بفرضية العدم وليست الفرضية البديلة. وأن رفض أو عدم رفض الفرضية البديلة إنما يأتي نتيجة لرفض أو عدم رفض فرضية العدم. بمعنى أن رفض فرضية العدم هو في واقع الأمر قبول للفرضية البديلة والعكس صحيح.

هذا وبناء على مستوى المعنوية المحدَّد من قِبَل الباحث، وبالاستفادة من توزيع المعاينة للمختبر الإحصائي، فإنه يمكن أن نعيِّن حدّاً حرجاً لرفض أو عدم رفض فرضية العدم آخذين في الاعتبار الفرضية البديلة. ذلك أن الفرضية البديلة هي التي تحدد لنا منطقة الرفض أو عدم الرفض للفرضية العدمية.

وفيما يلي نوضح كيفية تحديد منطقتي الرفض وعدم الرفض وذلك فيما يتعلق بالصور المختلفة للفرضية البديلة:

- **في حالة الفرضية البديلة غير المتجهة (الاختبار ذو الذيليْن):**

إذا كانت: $ف_{١}$: $م \neq م_{٠}$ حيث $م_{٠}$ قيمة مفترضة لمتوسط المجتمع (م)، فإن مستوى المعنوية (α) يُقسَّم إلى نصفيْن متساوييْن على كلٍّ من ذيلي منحنى توزيع المعاينة ($\frac{\alpha}{٢}$ عند كل ذيل). وعلى سبيل المثال، يوضِّح الشكل (٣ - ١) هذا الأمر في حالة مستوى المعنوية ٠,٠٥. وفي هذه الحالة تكون منطقة عدم رفض فرضية العدم في الوسط ومنطقة رفضها عند الذيليْن.



الشكل (٣ - ١)

مناطق الرفض وعدم الرفض عندما يكون الاختبار ذو ذيليْن

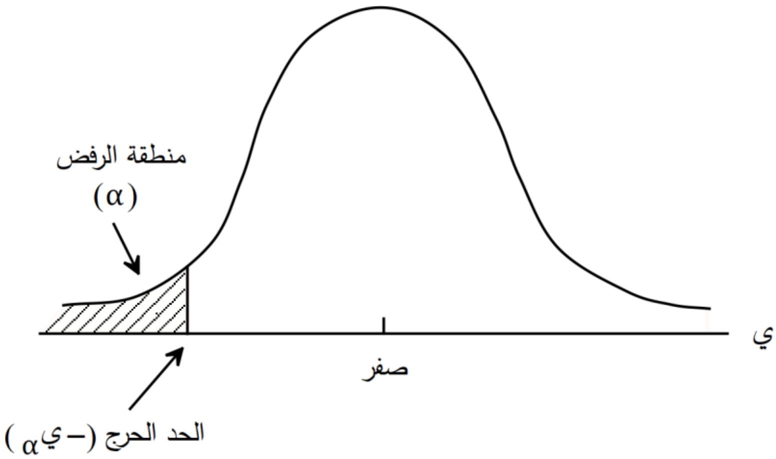
وفي هذه الحالة تكون منطقة عدم رفض فرضية العدم في الوسط ومنطقة رفضها عند الذيلين كما في الشكل (٣ - ١).

- **في حالة الفرضية البديلة المتجهة (الاختبار ذو الذيل الواحد):**

  وهنا تأخذ الفرضية البديلة إحدى الصورتيْن:

  أ- $ف_١$**:** م **<** $م_٠$ **ب-** $ف_١$**:** م **>** $م_٠$

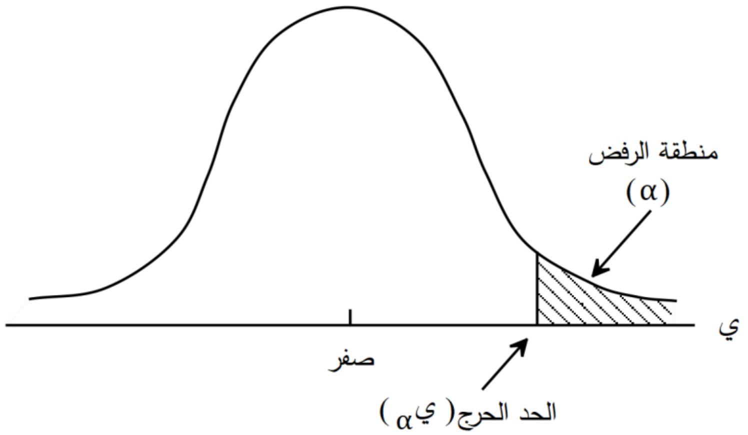
  وتتحدد منطقة الرفض في **الحالة (** أ **)** بالمساحة المظللة الواقعة في الجهة اليسرى (الذيل الأيسر) كما هي في الشكل (٣ - ٢).



الشكل (٣ - ٢)

مناطق الرفض وعدم الرفض عندما تكون الفرضية البديلة متجهة جهة القيمة الصغرى

وأمـا في **الحالة (ب)** فتكون منطقة الرفض في الذيل الأيمن كما فــي الشكــل (٣ - ٣).



الشكل (٣ - ٣)

مناطق الرفض وعدم الرفض عندما تكون الفرضية البديلة متجهة جهة القيمة الكبرى

ومن خلال تناولنا لاختبار الفرضيات فإنه يمكننا إدراك أن هناك تشابهاً ملحوظاً بين إجراءات القضاء لإدانة أو تبرئة متهم معين وبين اختبار الفرضيات الإحصائية لرفض أو عدم رفض فرضية العدم. وتبرز أوجه التشابه من خلال ما يلي:

- **فـــي القضـــــــاء:** المتهم بريء إلى أن تثبُت إدانته.

  في اختبار الفرضيات: فرضية العدم صحيحة إلى أن يثبت عدم صحتها.
- **فـــي القضـــــــاء:** يقوم الادعاء بجمع الأدلة لإثبات أن المتهم مذنب.

  في اختبار الفرضيات: يقوم الباحث بأخذ عينة من المجتمع ودراستها في محاولة للحصول على البيانات (الأدلة) والتي من شأنها أن تثبت صحة أو عدم صحة فرضية العدم. ودور الباحث هنا يناظر دور كل من الادعاء والدفاع في القضاء.
- **فـــي القضـــــــاء:** تُعرض الأدلة على القاضي كي يصدر قراره بإدانة المتهم أو عدم إدانته. وتتمثل أدوات القاضي في المواد والأحكام القانونية.

  في اختبار الفرضيــات: تُعرض البيانات على الإحصائي ليقرر بناء عليها رفض أو عدم رفض فرضية العدم. وتتمثل أدوات الإحصائي هنا أساساً في المختبِر الإحصائي. هذا والإحصائي هنا يناظر القاضي في القضاء.
- **في القضاء واختبار الفرضيات الإحصائية:**

  - إذا كان المتهم بريئاً فعلاً ولم تكن الأدلة كافيةً لإظهار براءته وتمت إدانته. فإن ذلك يناظر الوقوع في الخطأ من النوع الأول (α) في اختبار الفرضيات وهو رفض فرضية العدم (ف.) في حين أنها صحيحة.

- إذا كان المتهم مذنباً بالفعل وتمت عدم إدانته لقصور في حيثيات الدفاع، فإن ذلك يناظر الوقوع في خطأ من النوع الثاني (β) وهو عدم رفض فرضية العدم (ف.) في حين أنها غير صحيحة.
- قدرة القضاء على إدانة المذنبين تناظر قوة الاختبار (١ – β) والتي تتمثل في رفض فرضية العدم وهي خاطئة.

خطوات اختبار الفرضيات:

فيما يلي خطوات اختبار فرضية العدم وذلك وفقاً لتسلسلها المنطقي:

١- تحديد نوع توزيع المجتمع:

كثيراً ما يتطلب إجراء الاختبار حول فرضية العدم شروطاً يجب توافرها في المجتمع المسحوب منه العينة كأن تتخذ المشاهدات في المجتمع شكل التوزيع المعتاد. وفي مثل هذه الحالات يكون الاختبار معلمياً Parametric. أي أن الاختبارات المعلمية - وهي موضوع هذا الفصل - تتطلب توافر شروط معينة في المجتمع موضع البحث ولعل أهم هذه الشروط هي أن يتبع المجتمع توزيعاً معتاداً.

وأما الاختبارات اللامعلمية Non-Parametric فتُستخدم في الحالات التي لا يتبع فيها المجتمع توزيعاً معتاداً. وإجراء مثل تلك الاختبارات لا يتطلب أية شروط حول مجتمع البحث. لذلك فإنه في حالة عدم توافر المعلومات الكافية حول التوزيع في المجتمع، أو في حالة عدم إمكانية تحقق أن التوزيع النظري للمجتمع هو توزيع معتاد، فإن الاختبارات اللامعلمية هي الملائمة للاستخدام في هذه الحالات. وهذا النوع من الاختبارات يخرج عن نطاق هذا الكتاب.

٢- تحديد فرضية العدم (ف.) والفرضية البديلة (ف$_١$):

وعادة ما تُوضَّح فرضية العدم في صورة مبسطة (ف.: م = م. مثلاً)، أي في صورة تساوي وذلك لسهولة حساب الخطأ من النوع الأول (α) في هذه الحالة. كما تُحدَّد الفرضية البديلة بحيث يكون رفض فرضية العدم هو عدم رفض للفرضية البديلة، حيث أن ذلك يساعد على تحديد مناطق القبول والرفض.

٣- تحديد مستوى المعنوية (α) في هذه الحالة:

وهو عادة ما يكون: α = ٠,٠١ أو α = ٠,٠٥ أو غير ذلك حسب ما يراه الباحث ملائماً للحالة التي يقوم بدراستها.

٤- تحديد المختبِر الإحصائي Test Statistic المناسب وإيجاد قيمته باستخدام القيمة النظرية المفترضة لمعلمة المجتمع في فرضية العدم وقيمة المعلمة المحسوبة من العينة (الإحصاء) وأخيراً الانحراف المعياري (الخطأ المعياري) لإحصاء العينة.

٥- **معرفة التوزيع الاحتمالي** للمختبر الإحصائي.

٦- يمكن **اتخاذ القرار** بشأن قبول أو رفض فرضية العدم وذلك بمقارنة قيمة المختبِر الإحصائي المحسوبة في (٤) بالقيمة الحرجة لهذا المختبر والمحدَّدة باستخدام التوزيع النظري له. مع الأخذ في الاعتبار أن القيمة الحرجة لهذا المختبر تتحدد وفقاً للصورة التي تكون عليها الفرضية البديلة، وأن القيمة الحرجة هي التي تحدِّد بدورها مناطق الرفض وعدم الرفض لفرضية العدم. وبمقارنة القيمتيْن المشار إليهما يمكننا رفض أو عدم رفض فرضية العدم وذلك وفقاً للمنطقة التي تقع فيها قيمة المختبِر المحسوبة في (٤) من حيث كونها منطقة رفض أو عدم رفض. وسوف نوضح تلك النقطة في أمثلة قادمة.

وفي النهاية يلزم التنويه إلى أن مبرِّرات استخدام التوزيع المعتاد أو توزيع ت في عملية اختبار الفرضيات هي نفس المبررات التي سقناها في دراستنا لنظرية التقدير في الفصل السابق من هذا الكتاب. وأنه في دراستنا للاستدلال الإحصائي بشقيْه المعروفيْن (نظرية التقدير واختبار الفرضيات) سوف يقتصر استخدامنا للتوزيعات الاحتمالية على التوزيع المعتاد وتوزيع ت وذلك وفقاً لما يقتضيه المحتوى العلمي لهذا الكتاب. وعليه فإن المختبِر الإحصائي المستخدَم في عمليات اختبار الفرضيات في هذا الفصل لا يخرج عن كونه (ي) أو (ت) حيث يتوقف تحديد كل منهما على معلمات المجتمع المراد إجراء اختبار إحصائي عنها. أضف إلى ذلك أننا لسنا في حاجة هنا لأن نسوق مبرِّرات استخدام توزيع معين حيث تمت تغطية هذا الأمر في الفصل السابق.

وسوف يقتصر اهتمامنا في هذا الأمر على تناول اختبار الفرضيات على النحو التالي:

- اختبار الفرضيات حول متوسط مجتمع (م).
- اختبار الفرضيات حول النسبة في مجتمع (ل).
- اختبار الفرضيات حول الفرق بين متوسطيْ مجتمعيْن ($م_١ - م_٢$).
- اختبار الفرضيات حول الفرق بين نسبتيْ مجتمعيْن ($ل_١ - ل_٢$).

وسوف نتناول كلاًّ منها بشيء من التفصيل.

أولاً: اختبار الفرضيات الإحصائية حول متوسط مجتمع

يقوم اختبار الفرضيات في هذه الحالة على إنه إذا قمنا بسحب عينة عشوائية حجمها ن من مجتمع معتاد متوسطه (م) غير معلوم. فإن توزيع المعاينة للوسط الحسابي للعينة ($\bar{س}$) يتبع توزيعاً معتاداً بغض النظر عن حجم العينة صغيراً كان أم كبيراً. وأما إذا كان حجم العينة كبيراً فإنه باستخدام نظرية الحد المركزية - وعلى ضوء ما سبق تناوله في الفصل السابق - نجد أن توزيع المعاينة للوسط الحسابي للعينة $\bar{س}$ يقترب من التوزيع المعتاد وذلك بغض النظر عن توزيع المجتمع معتاداً كان أم غير معتاد.

وفي حالة اختبار ما إذا كان متوسط المجتمع يساوي قيمة معينة $م_.$، وكان تباين المجتمع σ^2 معلوماً، فإن المختبِر الإحصائي يأخذ الصورة:

$$ي = \frac{\bar{س} - م_.}{\frac{\sigma}{\sqrt{ن}}}$$

وهو يتبع التوزيع المعتاد المعياري. وأما إذا كان تباين المجتمع غير معروف فإننا نستخدم ع بدلاً من σ. وفي هذه الحالة فإن المختبِر الإحصائي يتبع توزيع ت في حالة العينات الصغيرة ($ن < ٣٠$) حيث نجد أن المختبِر الإحصائي يأخذ الصورة:

$$ي = \frac{\bar{س} - م_.}{\frac{ع}{\sqrt{ن}}}$$

وفي حالة العينات الكبيرة ($ن \geq ٣٠$) فإن توزيع ت يقترب من التوزيع المعتاد المعياري حيث نجد أن توزيع ت يقترب كثيراً من التوزيع المعتاد المعياري عندما تكون $ن \geq ٣٠$.

والآن لعله من المناسب للقارئ أن ننحو به بعيداً عن خِضَمّ تفصيلات نظرية قد تصيبه بالملل او قد تنتهي به إلى حيث بدأ، متجهين في ذلك إلى تطبيق تلك الخطوات على أمثلة عملية ترسّخ في ذهن القارئ كيفية إجراء فعلى لاختبار إحصائي. ونبدأ ذلك بالمثال الذي سقناه من قبل والذي يتعلق بشكوى المواطنين من عدم مطابقة رغيف الخبز المنتَج في أحد المخابز للمواصفات المطلوبة المتعلقة بالوزن. حيث أن هذا المثال يدخل ضمن نطاق اختبار الفرضيات حول متوسط مجتمع. وسوف تكون نقطة البدء في ذلك هي صياغة المثال المشار اليه في قالب إحصائي على النحو الموضح في مثال (١). آخذين في الاعتبار أن الحسابات المتعلقة بجميع الأمثلة التي نتناولها مبنية على أساس افتراض صحة فرضية العدم.

مثال (١):

أُخذت عينة حجمها ١٠٠ رغيف من إنتاج أحد المخابز حيث وجد أن الوسط الحسابي والانحراف المعياري لوزن الرغيف في العينة هما ١٧٨ جم، ٩ جم على الترتيب. اختبر صحة الفرضيات التالية بافتراض أن وزن الرغيف في المخبز يتبع التوزيع المعتاد.

أولاً: **ف.**: $م = ١٨٠$ جم ، **ف١**: $م \neq ١٨٠$ جم

ثانياً: **ف.**: $م = ١٨٠$ جم ، **ف١**: $م < ١٨٠$ جم

الحــــل:

أولاً: بتطبيق خطوات إجراء الاختبار الإحصائي على مثال (١) نجد ما يلي:

١- التوزيــع المستخــدم هنا هو التوزيع المعتاد وذلك وفقاً لما نص عليه المثال. وحتى لو لم ينص المثال على ذلك فإنه يمكننا استخدام التوزيع المعتاد وذلك لكبر حجم العينة ($ن$ = ١٠٠ > ٣٠).

٢- فرضية العدم والفرضية البديلة تم عرضهما في هذا المثال بصورة واضحة وصريحة كما هو وارد في أولاً وثانياً. وهنا تجدر الإشارة إلى أنه في كثير من الحالات يكون مطلوباً من الباحث صياغة هاتيْن الفرضيتيْـن بشكل يتلاءم مع الأمر المراد بحثه (يُلاحَظ أنه عند عرض المثال في صورته الأولى لم تكن الفرضيات مُصاغة كما هي عليه في هذا المثال).

٣- لنفترض هنا أننا بصدر استخدام مستوى معنوية ٠,٠٥ .

٤- المختبِر الإحصائي في هذه الحالة هو **ي** حيث تعرف كما يلي:

$$ي = \frac{\bar{س} - م_.}{\frac{ع}{\sqrt{ن}}}$$

$$= \frac{١٧٨ - ١٨٠}{\frac{٩}{\sqrt{١٠٠}}} = \frac{-٢}{٠,٩} = -٢,٢$$

٥- التوزيع الاحتمالي للمختبِر الإحصائي **ي** هو التوزيع المعتاد كما سبق الإشارة إلى ذلك في الفصل السابق.

٦- فيما يتعلق بالاختبار الإحصائي في أولاً:

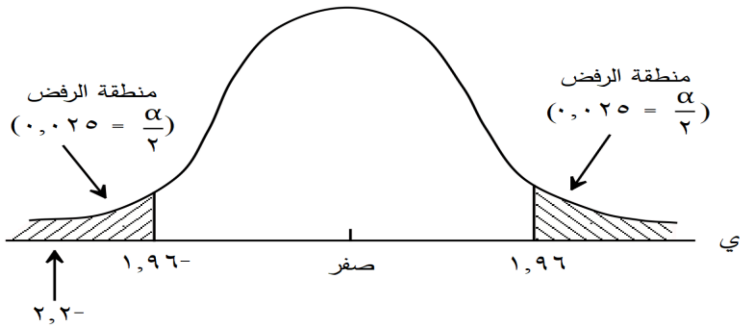
ف.: م = ١٨٠ جم ، **ف.**: م ≠ ١٨٠ جم

وحيث أن الفرضية البديلة هي التي تحدد منطقة الرفض فإن الاختبار هنا ذو ذيليْن وتقع منطقة الرفض عند الذيليْن. وهنا يمكن رفض أو عدم رفض فرضية العدم وذلك على النحو الموضح من قبل.

حيث أن مستوى المعنوية المستخدم هو:

α = ٠,٠٥ فإن $ي_{\frac{\alpha}{٢}}$ = ١,٩٦

وبذلك تتحدد مناطق الرفض وعدم الرفض بالقيم الحرجة -١,٩٦، ١,٩٦ .



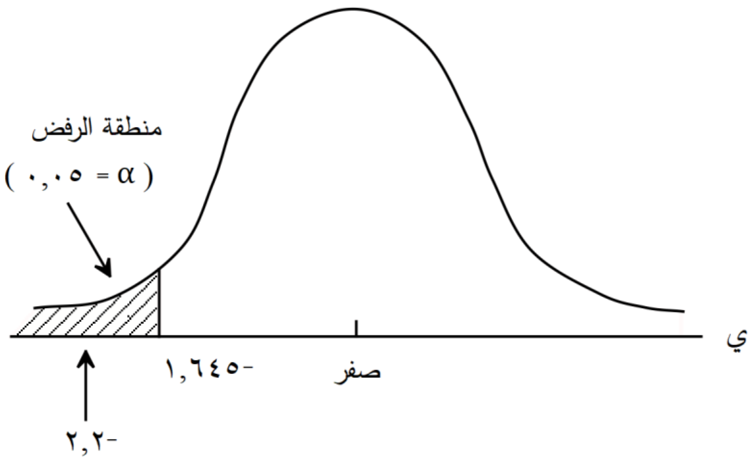
وحيث أن قيمة **ي** المحسوبة هي -٢,٢٢ فإنه يمكننا استنتاج أن هذه القيمة تقع في منطقة الرفض عند الذيل الأيسر. أي أننا نرفض فرضية العدم القائلة بأن متوسط وزن رغيف الخبز المنتَج في هذا المخبز هو ١٨٠ جـم. وبذلك نقبل الفرضية البديلة القائلة بأن هذا المتوسط لا يساوي ١٨٠ جم.

ثانياً: فيما يتعلق بالاختبار الإحصائي في ثانياً:

ف.: م = ١٨٠ جم ، **ف.**: م < ١٨٠ جم

منطقة الرفض هنا تحددها أيضاً الفرضية البديلة حيث تقع عند الذيل الأيسر.

ومن جدول المساحة تحت المنحنى المعتاد المعياري نجد أن الحد الحرج هو **ي** = - ١,٦٤٥.



وبذلك يمكننا استنتاج أن قيمة **ي** المحسوبة (-٢,٢) تقع في منطقة الرفض. أي أننا نرفض فرضية العدم ونقبل الفرضية البديلة القائلة بأن متوسط وزن الرغيف يقل عن ١٨٠ جم. وهذا يؤكد صحة شكاوى المواطنين في هذا الشأن.

وقبل أن نستطرد في سرد المزيد من الأمثلة نوضح أنه عند رفض أو عدم رفض فرضية العدم يقوم الباحث بتحديد الحدود الحرجة للمختبِر الإحصائي وذلك بناءً على توزيع المعاينة النظري له. ثم يتخذ الباحث قراره بعد ذلك برفض أو عدم رفض فرضية العدم حسبما تقع القيمة المحسوبة للمختبِر في منطقة الرفض أو منطقة عدم الرفض. هذا وفي حالة استخدام التوزيع المعتاد، وحتى لا يجد القارئ نفسه مضطراً لاستخدام جداول المساحة تحت المنحنى المعتاد المعياري في كل مرة يستخدم فيها مستويات المعنوية ٠,٠١ ، ٠,٠٥ ، ٠,١٠ فإننا - ومن باب التيسير على القارئ - نعرض الحدود الحرجة لمناطـق الرفـض وعدم الرفض في هذه الحالات على النحو الموضح في الجدول (٣ - ٢):

الجدول (٣ - ٢)

الحدود الحرجة لمناطق الرفض وعدم الرفض عند استخدام مستويات المعنوية

α = ٠,٠١ ، α = ٠,٠٥ ، α = ٠,١٠ (في حالة التوزيع المعتاد)

نوع الاختبار / مستوى المعنوية	اختبار ذو ذيلَيْن		اختبار ذو ذيل واحد	
	منطقة الرفض عند الذيل الأيمن	منطقة الرفض عند الذيل الأيسر	منطقة الرفض عند الذيل الأيمن	منطقة الرفض عند الذيل الأيسر
α = ٠,٠٥	١,٩٦	١,٩٦-	١,٦٤٥	١,٦٤٥-
α = ٠,٠١	٢,٥٨	٢,٥٨-	٢,٣٣	٢,٣٣-
α = ٠,١٠	١,٦٤٥	١,٦٤٥-	١,٢٨	١,٢٨-

وأما فيما يتعلق بالحدود الحرجة في حالة استخدام اختبار ت فإنه وكما عرفنا في الفصل السابق يختلف شكل توزيع ت باختلاف درجات الحرية. وعليه فإنه في كل مرة يستخدم فيها توزيع ت نجد أن هناك حدوداً حرجة تتوقف في قيمتها على درجات الحرية المستخدمة.

والآن، وعلى ضوء ما سبق، يمكننا تلخيص الخطوات التي يجب أن تُتَخذ عند إجراء اختبار إحصائي متعلق بمتوسط المجتمع - سواء تم استخدام اختبار **ي** أو اختبار ت - وذلك على النحو الموضَّح في الجدول (٣ - ٣) والجدول (٣ - ٤).

وفيما يلي بعض الأمثلة التي توضح كيفية إجراء اختبار الفرضيات حول متوسط المجتمع الواحد وذلك في الحالات المختلفة.

مثال (٢):

سُحبت عينة عشوائية مكونة من ٢٠٠ طالب من طلاب الفرقة الثانية بإحدى كليات التجارة حيث وُجد أن متوسط درجاتهم في مادة مبادئ الإحصاء هو ٦٥ درجة والانحراف المعياري ٨ درجات. هل يدل ذلك على أن متوسط درجات طلاب الفرقة الثالثة ككل هو أقل من ٦٧ درجة وذلك عند مستوى معنوية ٠,٠٥؟

الحـــــل:

$\bar{س}$ = ٦٥ ، ع = ٨ ، ن = ٢٠٠ ، $م_٠$ = ٦٧

فرضية العــدم ($ف_٠$): م = ٦٧

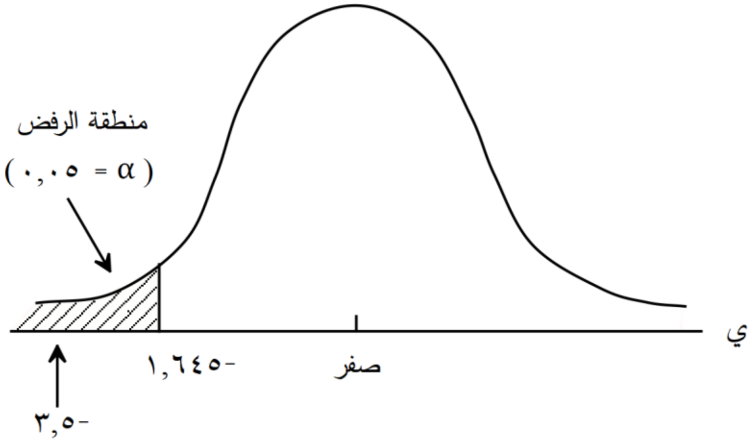
الفرضية البديلة ($ف_١$): م < ٦٧

من بيانات العينة، وبافتراض صحة فرضية العدم نجد أن:

المختبِر الإحصائي:

$$ي = \frac{\bar{س} - م_{٠}}{\frac{ع}{\sqrt{ن}}} = \frac{٦٥ - ٦٧}{\frac{٨}{\sqrt{٢٠٠}}} = \frac{-٢,٠}{٠,٥٧} = - ٣,٥$$

منطقة الرفض: حيث أن الاختبار هنا ذو ذيل واحد ومنطقة الرفض تقع عند الذيل الأيسر ويحددها الحد الحرج (**ي** < - **ي**$_{٠,٠٥}$)، حيث **ي**$_{٠,٠٥}$ تساوي ١,٦٤٥.



وحيث أن قيمة ى المحسوبة (**ي** = - ٣,٥) أقل من - ١,٦٤٥، أي أن قيمة **ي** المحسوبة تقع في منطقة الرفض، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة. مما يدل على أن متوسط درجات طلاب الفرقة الثانية في مادة مبادئ الإحصاء يقل عن ٦٧ درجة وذلك عند مستوى معنوية ٠,٠٥.

مثال (٣):

بالرجوع إلى بيانات المثال (٢)، هل ستختلف النتيجة التي تم التوصُّل إليها في حالة استخدام مستوى المعنوية ٠,٠١ بدلا من ٠,٠٥؟

الحل:

الإختلاف الوحيد هنا عن الوضع في المثال (٢) هو قيمة الحد الحرج والذي يساوي في حالتنا هذه -٢,٣٣ بدلا من -١,٦٤٥.

وحيث أن قيمة ى المحسوبة (**ي** = - ٣,٥) أقل من - ٢,٣٣، أي أن قيمة **ي** المحسوبة لا زالت تقع في منطقة الرفض، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة. مما يدل على أن متوسط درجات طلاب الفرقة الثالثة في مادة مبادئ الإحصاء يقل عن ٦٧ درجة وذلك عند مستوى معنوية ٠,٠١ أيضاً.

الجدول (٣ - ٣)

اختبار الفرضيات الإحصائية حول متوسط مجتمع باستخدام التوزيع المعتاد (اختبار ي)

نوع الاختبار / بيان	اختبار ذو ذيل واحد (الفرضية البديلة متجهة)	اختبار ذو ذيلين (الفرضية البديلة غير متجهة)
فرضية العدم والفرضية البديلة	$ف_٠$: م = م٠ $ف_١$: م < م٠ أو $ف_١$: م > م٠ حيث: م٠ هي قيمة معلمة المجتمع في فرضية العدم.	$ف_٠$: م = م٠ $ف_١$: م ≠ م٠ حيث: م٠ هي القيمة التي تأخذها معلمة المجتمع في فرضية العدم.
المختبِر الإحصائي	$ي = \frac{\bar{س} - م_٠}{\frac{\alpha}{\sqrt{ن}}}$ أو باستخدام ع كتقريب لـ σ إذا كانت σ غير معروفة	
منطقة الرفض	في حالة $ف_١$: م < م٠ ي < - $ي_{\alpha}$ في حالة $ف_١$: م > م٠ ي > $ي_{\alpha}$ حيث $ي_{\alpha}$ تتحدد قيمتها بحيث: $\phi(ي_{\alpha}) = ١ - \alpha$	ي < - $ي_{\frac{\alpha}{٢}}$ أو: ي > $ي_{\frac{\alpha}{٢}}$ حيث $ي_{\frac{\alpha}{٢}}$ تتحدد قيمتها بحيث: $\phi(ي_{\frac{\alpha}{٢}}) = ١ - \frac{\alpha}{٢}$
شروط التطبيق	**في حالة العينات الكبيرة (ن ≥ ٣٠):** ليست هناك شروط حول توزيع المجتمع، ويمكن استخدام التوزيع المعتاد تطبيقاً لنظرية الحد المركزية. **في حالة العينات الصغيرة (ن < ٣٠):** توزيع المجتمع معتاد وتباينه معلوم.	

الجدول (٣ - ٤)

اختبار الفرضيات الإحصائية حول متوسط مجتمع باستخدام توزيع ت (اختبار ت)

نوع الاختبار / بيان	اختبار ذو ذيل واحد (الفرضية البديلة متجهة)	اختبار ذو ذيلين (الفرضية البديلة غير متجهة)
فرضية العدم والفرضية البديلة	ف$_{.}$: م = م$_{.}$ ف$_{1}$: م < م$_{.}$ أو ف$_{1}$: م > م$_{.}$ حيث: م$_{.}$ هي قيمة معلمة المجتمع في فرضية العدم	ف$_{.}$: م = م$_{.}$ ف$_{1}$: م ≠ م$_{.}$ حيث: م$_{.}$ هي قيمة معلمة المجتمع في فرضية العدم
المختبِر الإحصائي	ت = $\dfrac{\bar{س} - م_{.}}{\dfrac{ع}{\sqrt{ن}}}$ أو باستخدام ع كتقريب لـ σ إذا كانت σ غير معروفة.	
منطقة الرفض	في حالة ف$_{1}$: م < م$_{.}$ ت < -ت$_{\alpha}$ في حالة ف$_{1}$: م > م$_{.}$ ت > ت$_{\alpha}$ حيث ت$_{\alpha}$ تتوقف في قيمتها على (ن – ١) درجات حرية.	ت < -ت$_{\frac{\alpha}{2}}$ أو: ت > ت$_{\frac{\alpha}{2}}$ حيث ت$_{\frac{\alpha}{2}}$ تتوقف في قيمتها على (ن – ١) درجات حرية.
شروط التطبيق	توزيع المجتمع المعتاد، العينة صغيرة (ن < ٣٠)، تباين المجتمع مجهول.	

مثال (٤):

تنتج آلة ألواحاً من الصلب سمك الواحد منها يساوي ٠,٥ بوصة بانحراف معياري ٠,١ بوصة. وبعد إجراء تعديلات معينة على هذه الآلة، أُخذت عينة من إنتاجها مكونة من ٦٤ لوحاً من الصلب حيث تبيَّن أن متوسط سمك اللوح هو ٠,٥٣ بوصة.

هل يمكن القول بأن إجراء التعديلات أدى إلى زيادة متوسط سمك اللوح وذلك عند مستوى معنوية ٠,٠١؟

الحـــل:

$\bar{س}$ = ٠,٥٣ ، σ = ٠,١ ، ن = ٦٤ ، م$_{.}$ = ٠,٥

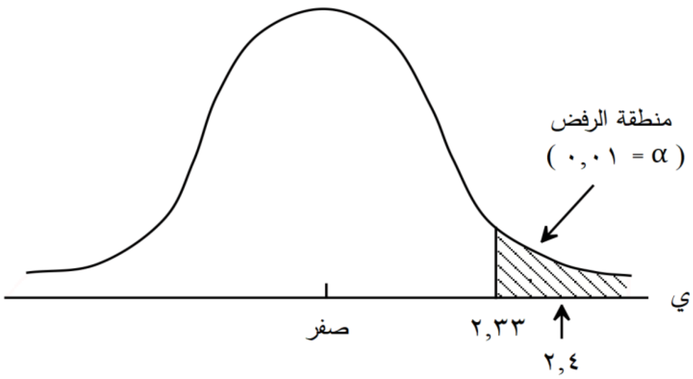
فرضية العـــدم: $ف_٠ = ٠,٥$

الفرضية البديلة: $ف_١ > ٠,٥$

وبافتراض صحة فرضية العدم، ومن بيانات العينة نجد أن:

المختبِر الإحصائي هو: $ي = \dfrac{\bar{س} - م_٠}{\frac{\sigma}{\sqrt{ن}}}$

$$\therefore\ ي = \frac{٠,٥٣ - ٠,٥}{\frac{٠,١}{\sqrt{٦٤}}} = ٢,٤$$



وحيث أن الاختبار هنا ذو ذيل واحد، وأن منطقة الرفض تقع عند الذيل الأيمن ($ي > ي_{١ ,٠١}$)، حيث $ي_{١ ,٠١} = ٢,٣٣$، فإننا نجد أن قيمة ي المحسوبة ($ي = ٢,٤$) هي أكبر من ٢,٣٣ وتقع في منطقة الرفض.

أي أننا نرفض فرضية العدم ونقبل الفرضية البديلة القائلة بأن إجراء التعديلات أدى إلى زيادة متوسط سمك اللوح وذلك عند مستوى معنوية ٠,٠١.

مثال (٥):

يرى مدير إحدى المدارس الثانوية الخاصة بأن متوسط أداء طلاب مدرسته في شهادة الثانوية العامة لا يختلف عن المتوسط العام لطلاب مصر ككل. فإذا علمت أن عدد طلاب الشهادة الثانوية في مدرسته يساوي ١٢٠ طالباً وأن متوسط معدلاتهم كان ٦٤، وأن متوسط المعدلات لطلاب الشهادة الثانوية في المدارس ككل كان ٦٥ بانحراف معياري يساوي ٨. هل تدعم هذه البيانات صحة ما يراه المدير؟ استخدم مستوى معنوية ٠,٠٥.

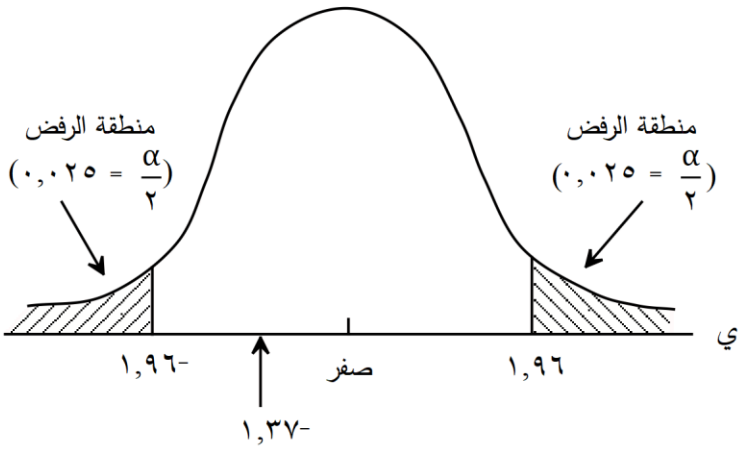
الحـــــل:

$\bar{س} = ٥٤$ ، $ع = ٨$ ، $ن = ١٢٠$ ، $م_٠ = ٦٥$

$ف_٠$: $م = ٦٧$ ، $ف_١$: $م \neq ٦٧$

من بيانات العينة، وبافتراض صحة فرضية العدم نجد أن:

$$ي = \frac{\bar{س} - م_{.}}{\frac{ع}{\sqrt{ن}}} = \frac{٦٤ - ٦٥}{\frac{٨}{\sqrt{١٢٠}}} = \frac{-١}{٠,٧٣} = -١,٣٧$$



الاختبار هنا ذو ذيليْن فإن: $ي_{\frac{\alpha}{٢}} = ي_{٠,٠٢٥} = ١,٩٦$.

وحيث أن منطقة الرفض تقع عند الذيلين ($ي < -ي_{\frac{\alpha}{٢}}$ و $ي > ي_{\frac{\alpha}{٢}}$) لذلك نجد أن قيمة ي المحسوبة (ي = - ١,٣٧) أكبر من - ١,٩٦ وهي تقع في منطقة عدم الرفض. أي أننا نقبل فرضية العدم والتي تدعم رأي مدير المدرسة حول مستوى طلاب مدرسته مقارناً بطلاب المدارس ككل.

مثال (٦):

إذا كانت أوزان مجموعة كبيرة من الأشخاص تتبع توزيعاً معتاداً بمتوسط ٨٣ كجم. وأنه أُخذت عينة مكونة من ٢٥ شخصاً من هذه المجموعة حيث اتبعت نظاماً غذائياً خاصاً لإنقاص الوزن. وبعد ثلاثة شهور وُجد أن متوسط أوزان هذه المجموعة هو ٧٧ كجم. إلى أي مدى يمكننا أن نعزي هذا النقص في الوزن إلى اتباع النظام الغذائي الخاص إذا علمت أن الانحراف المعياري للعينة هو ١٠ كجم. استخدم مستوى معنوية ٠,٠١.

الحــل:

$\bar{س} = ٧٧$ ، $ع = ١٠$ ، $ن = ٢٥$ ، $م_{.} = ٨٣$

$ف_{.}$: $م = ٨٣$ ، $ف_{١}$: $م < ٨٣$

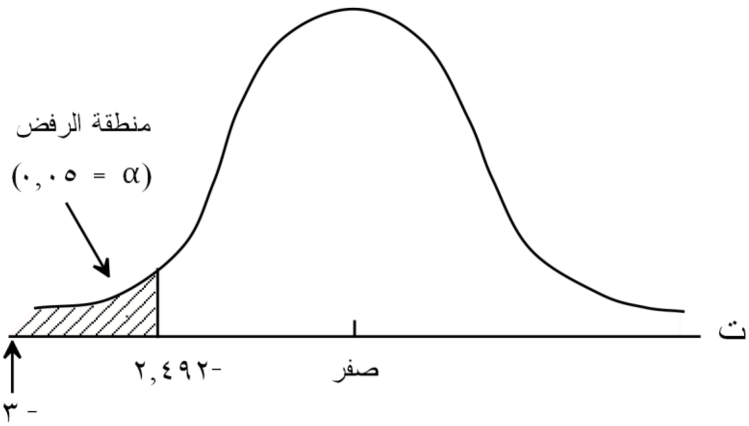
وحيث أن:

- توزيع المجتمع معتاد
- تباين المجتمع مجهول
- حجم العينة صغير (ن < ٣٠)

فإن المختبِر الإحصائي الملائم هو

$$\text{ت} = \frac{\bar{\text{س}} - \text{م}_0}{\frac{\text{ع}}{\sqrt{\text{ن}}}} = \frac{٧٧ - ٨٣}{\frac{١٠}{\sqrt{٢٥}}} = \frac{-٦}{٢} = -٣$$

هذا والاختبار هنا ذو ذيل واحد حيث تقع فيه منطقة الرفض عند الذيل الأيسر (ت < -ت_α)، وبإيجاد قيمة $\text{ت}_{٠,٠١}$ عند درجة حرية (ن – ١ = ٢٥ – ١ = ٢٤) وذلك من جدول ت نجد أن: $\text{ت}_{(٠,٠١ ، ٢٤)}$ تساوي ٢,٤٩٢.



وحيث أن قيمة ت المحسوبة (ت = - ٣) أقل من قيمة - $\text{ت}_{٠,٠١}$ فإننا نرفض فرضية العدم ونقبل الفرضية البديلة والتي تفيد بأن النقص الحادث في وزن الأشخــاص يمكــن إرجاعه إلى النظام الغذائي الخاص وذلك عند مستــوى معنويــة (α = ٠,٠١).

ثانياً: اختبار الفرضيات الإحصائية حول النسبة في مجتمع

كثيراً ما تواجهنا في حياتنا العملية الحاجة إلى الحكم على مجتمعات فيما يتعلق بالنسب وذلك باستخدام نتائج عينات مسحوبة من تلك المجتمعات.

فعلى سبيل المثال:

- هل يمكن القول بأن نسبة الأمية في مصر تختلف عن مثيلتها في بقية الدول العربية؟
- هل يمكن لصاحب مصنع معين أن يدّعى بأن نسبة التالف في إنتاج مصنعه تقل عن مثيلتها في مصانع أخرى مشابهة؟

- هل تختلف نسبة النجاح في إحدى المدارس عن نسبة النجاح في جميع المدارس المماثلة ككل؟
- كيف يمكن البت في صحة ادعاء إحدى الشركات الصناعية بشأن أن إجراء تعديلات في الآلات يؤدى إلى زيادة نسبة الإنتاج المطابق للمواصفات المطلوبة؟

إلى غير ذلك من الأسئلة العديدة والتي تتطلب الإجابة عليها اللجوء إلى استخدام العينات وتطبيق اختبار الفرضيات الإحصائية المتعلقة بالنسب وذلك من أجل الوصول إلى حقائق الأمور بالسرعة والدقة المطلوبتيْن وبأقل التكاليف الممكنة. ذلك أن النسب في المجتمعات عادة ما تكون مجهولة أو على أقل تقدير غير معروفة بدقة، وأنه عادة ما نلجأ إلى اتخاذ قرار بشأن هذه النسب مستعينين في ذلك بدراسة عينات صغيرة نسبية مأخوذة من تلك المجتمعات. واختبار الفرضيات للنسب هو واحد من أهم الأساليب الإحصائية التي تمكننا من تحقيق ذلك.

وهنا، وفي هذا الشأن، يجب أن نميِّز بين حالتيْن:

أولاهما: عندما تكون العينات صغيرة الحجم، وهي الحالة التي يمكن معها استخدام توزيع "ذو الحديْن" في عمليات الرفض او القبول.

ثانيهمـا: عندما تكون العينات الكبيرة، والتي يمكن في حالتها استخدام التوزيع المعتاد في عملية اتخاذ القرارات بشأن رفض أو قبول الفرضيات.

ونظراً لأن الدراسات والأبحاث المتعلقة بالنسب في المجتمعات عادةً ما تتطلب عينات كبيرة الحجم، فإن تناولنا لاختبار الفرضيات حول النسب سوف يكون مقتصراً على الحالات التي تكون فيها العينات كبيرة الحجـم ($ن \geq ٣٠$). ومن ثم يمكن معها استخدام التوزيع المعتاد (اختبار ي) في عمليات الرفض أو القبول.

ويقوم اختبار الفرضيات حول النسبة في مجتمع ما على أنه إذا أردنا اختبار أن النسبة في المجتمع ل تأخذ قيمة معينة ل. مثلاً، فإننا نقوم بسحب عينة من هذا المجتمع حجمها ن ثم نوجد النسبة في العينة $\hat{ل}$ وهي تقدير غير متحيز للنسبة في المجتمع ل. هذا وتوزيع المعاينة للنسبـة في العينة $\hat{ل}$ هو توزيع معتاد بمتوسط ل وتباين $\frac{ل(١ - ل)}{ن}$.

ولإجراء اختبار أن النسبة في المجتمع تساوي قيمة معينة $ل_{.}$، وبافتراض صحة فرضية العدم، فإنه إذا كان حجم العينة ن كبيراً كبراً كافياً ($ن \geq ٣٠$) فإن المختبِر الإحصائي في هذه الحالة يأخذ الصورة:

$$ي = \frac{\hat{ل} - ل_{.}}{\sqrt{\frac{ل_{.}(١ - ل_{.})}{ن}}}$$

وهو متغير عشوائي يتبع التوزيع المعتاد المعياري.

ولعله من الملائم هنا استعراض الحالات المختلفة لاختبار الفرضيات حول النسبة في المجتمع وذلك على النحو الموضح في الجدول (٣ - ٥).

الجدول (٣ - ٥)
اختبار الفرضيات حول النسبة في المجتمع (ل)

بيان / نوع الاختبار	اختبار ذو ذيل واحد (الفرضية البديلة متجهة)	اختبار ذو ذيلين (الفرضية البديلة غير متجهة)
فرضية العدم والفرضية البديلة	**$ف_{.}$: ل = $ل_{.}$** **$ف_{١}$: ل < $ل_{.}$** **أو $ف_{١}$: ل > $ل_{.}$** **حيث: $ل_{.}$ هي قيمة معلمة المجتمع في فرضية العدم.**	**$ف_{.}$: ل = $ل_{.}$** **$ف_{١}$: ل ≠ $ل_{.}$** **حيث: $م_{.}$ هي قيمة معلمة المجتمع في فرضية العدم.**
المختبِر الإحصائي	**$ي = \frac{\hat{ل} - ل_{.}}{\sigma_{\hat{ل}}}$** **حيث: $\hat{ل}$ النسبة في العينة ،** **$\sigma_{\hat{ل}} = \sqrt{\frac{ل_{.}(١ - ل_{.})}{ن}}$**	
منطقة الرفــض	**في حالة $ف_{١}$: ل < $ل_{.}$** **$ي < - ي_{\alpha}$** **في حالة $ف_{١}$: ل > $ل_{.}$** **$ي > ي_{\alpha}$** **حيث $\phi(ي_{\alpha}) = ١ - \alpha$**	**$ت < - ت_{\frac{\alpha}{٢}}$** **أو:** **$ت > ت_{\frac{\alpha}{٢}}$** **حيث $\phi(ت_{\frac{\alpha}{٢}}) = ١ - \frac{\alpha}{٢}$**
شروط التطبيــق	**أن يكون حجم العينـة كبيراً ($ن \geq ٣٠$)، حتى يمكن اســتخدام اختبار (ي).** **٢- أن لا تشمل فترة الثـــقة $\hat{ل} \pm ٣\sigma_{\hat{ل}}$ أي من القيمتيْن صفر أو واحد.**	

والآن نتناول بعض الأمثلة التي تغطى الحالات المختلفة التي أشرنا إليها في هذا الجدول.

مثال (٧):

أُخذت عينة حجمها ٥٠٠ مولود من مواليد إحدى المستشفيات فوجد أن من بينها ٢٦٠ مولوداً ذكراً. اختبر صحة الفرضية القائلة بأن المواليد الذكور يمثلون نصف إجمالي المواليد وذلك بمستوى معنوية ٠,٠٥

الحـــل:

بافتراض أن ل هي نسبة المواليد الذكور إلى إجمالي المواليد في المجتمع ككل. وحيث أن القيمة المُفترَضة لهذه النسبة في فرضية العدم هي ٠,٥ أي أن: $\text{ل}_{\circ} = ٠,٥$

ومن بيانات المثال نجد أن:

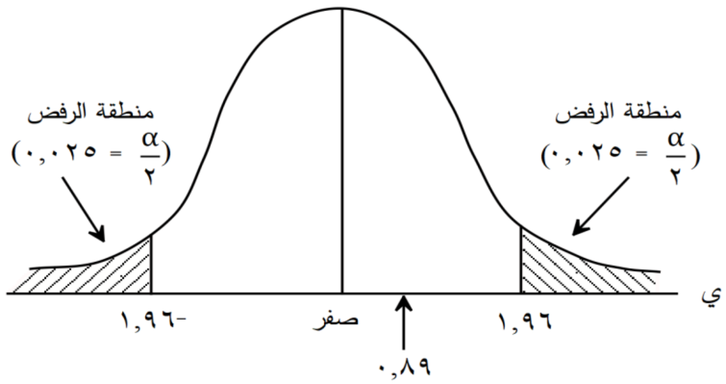
$$\hat{\text{ل}} = \frac{٢٦٠}{٥٠٠} = ٠,٥٢ \quad ، \quad \text{ن} = ٥٠٠ \quad ، \quad \text{ل}_{\circ} = ٠,٥$$

فرضيـــة العدم ($\text{ف}_{\circ}$): $\text{ل} = ٠,٥$

الفرضية البديلة ($\text{ف}_{١}$): $\text{ل} \neq ٠,٥$

$$\text{المختبِر الإحصائي (ي)} = \frac{\hat{\text{ل}} - \text{ل}_{\circ}}{\sqrt{\frac{\text{ل}_{\circ}(١ - \text{ل}_{\circ})}{\text{ن}}}}$$

$$= \frac{٠,٥٢ - ٠,٥٠}{\sqrt{\frac{٠,٥ \times ٠,٥}{٥٠٠}}} = \frac{٠,٠٢}{٠,٠٢٢} = ٠,٨٩$$



$\text{ي}_{\frac{\alpha}{٢}} = \text{ي}_{٠,٠٢٥} = ١,٩٦$ (الاختبار هنا ذو ذيليْن)

وحيــث أن إحدى منطقتي الرفــض يحددهـــا ($\text{ي} > \text{ي}_{\frac{\alpha}{٢}}$)، أي ($\text{ي} > ١,٩٦$)، فإننـــا نجد أن القيمة المحسوبــة للمختبِر الإحصائي ($\text{ي} = ٠,٨٩$) تقع في منطقة عدم الرفض. أي أننا نقبل الفرضية القائلة بأن المواليد الذكور يمثلون ٥٠٪ من إجمالي المواليد.

مثال (٨):

تدَّعي إحدى شركات الأدوية بأن الدواء المنتَج في مصانعها له القدرة على شفاء ٩٠٪ من المرضى الذين يتناولونه. وللتأكد من صحة ادعاء هذه الشركة، أُختيرت عينة من ١٢٠ مريضاً تناولوا هذا الدواء وجرت مراقبتهم في إحدى المستشفيات فوُجد أن ١٠٠ مريض منهم قد شُفي فعلاً بينما ظل الباقون على حالهم.

هل تدعم هذه البيانات صحة ادعاء الشركة؟ مستخدماً في ذلك مستوى معنوية ٠,٠١.

الحـــل:

من بيانات المثال نجد أن:

$$\hat{ل} = \frac{١٠٠}{١٢٠} = ٠,٨٣٣ \quad ، \quad ن = ١٢٠ \quad ، \quad ل_٠ = ٠,٩$$

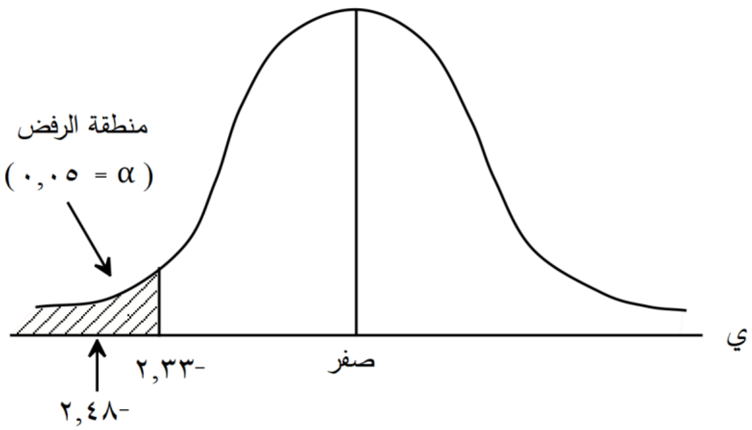
ويمكن صياغة فرضية العدم والفرضية البديلة كما يلي:

$$ف_٠:\ ل = ٠,٩ \quad ، \quad ف_١:\ ل < ٠,٩$$

وقد صيغت الفرضية البديلة بهذا الشكل وذلك لأن نقص نسبة الذي يشفون نتيجة لتناول هذا الدواء عن ٠,٩ هو الذي يؤكد عدم صحة ادعاء الشركة. بينما لو أخذنا الفرضية البديلة على الصورة ($ف_١:\ ل \neq ٠,٩٠$)، فإن هذا يعني في جزء منه أن نسبة الذين يشفيهم هذا الدواء يمكن أن تكون أكبر من ٩٠٪ وهذا في حد ذاته تأكيد لزعم الشركة بزيادة فعالية الدواء المنتَج في مصانعها.

وبافتراض صحة فرضية العدم ($ف_٠$):

$$ي = \frac{\hat{ل} - ل_٠}{\sqrt{\frac{ل_٠(١ - ل_٠)}{ن}}} = \frac{٠,٨٣٣ - ٠,٩}{\sqrt{\frac{٠,٩ \times ٠,١}{١٢٠}}} = \frac{-٠,٠٦٧}{٠,٠٢٧} = -٢,٤٨$$



وحيث أن: α = ٠,٠١ فإن $ي_{\frac{\alpha}{٢}}$ = ٢,٣٣ (الاختبار هنا ذو ذيل واحد). لذلك نجد أن قيمة **ي** المحسوبة (**ي** = - ٢,٤٨) تقع في منطقة الرفض (منطقة الرفض هي **ي** < -٢,٣٣).

أي أننا - ووفقاً لنتائج العينة - نستطيع القول بأنه ليست هناك أدلة كافية على صحة ادعاء الشركة بأن دواءها المنتج في مصانعها يستطيع إشفاء ٩٠٪ من المرضى الذين يتعاطونه، وأن هذه النسبة تقل عن ٩٠٪. وتأتى تلك النتيجة انعكاساً (لرفضنا لفرضية العدم وقبولنا للفرضية البديلة).

مثال (٩):

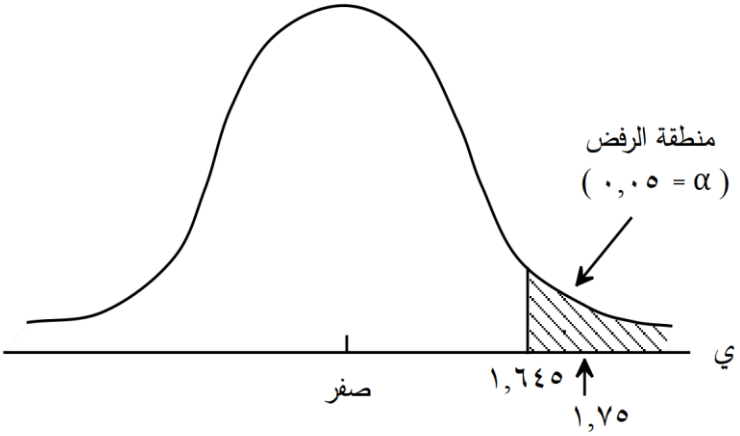
أُختيرت عينة حجمها ٢٠٠ طالب من بين طلاب إحدى كليات التجارة فوُجد أن من بينهم ٢٥ طالباً يدخنون السجائر. هل يمكن القول بأن نسبة المدخنين في هذه الكلية تزيد على ٩٪ وذلك حسب ما توصلت إليه إحدى الدراسات التي أُجريت في هذا الصدد من قبل؟ استخدم مستوى معنوية ٠,٠٥.

الحـــل:

$$\hat{ل} = \frac{٤٥}{٢٠٠} = ٠,١٢٥ \quad ، \quad ن = ٢٠٠ \quad ، \quad ل_{٠} = ٠,٠٩$$

$ف_{٠}$: ل = ٠,٠٩ ، $ف_{١}$: ل > ٠,٠٩

$$ي = \frac{\hat{ل} - ل_{٠}}{\sqrt{\frac{ل_{٠}(١ - ل_{٠})}{ن}}} = \frac{٠,١٢٥ - ٠,٠٩}{\sqrt{\frac{٠,٠٩ \times ٠,٩١}{٢٠٠}}} = \frac{٠,٠٣٥}{٠,٠٢} = ١,٧٥$$



وحيث أن $ي_{\alpha}$ = $ي_{٠,٠٥}$ = ١,٦٤٥، وأن منطقة الرفض تقع عند الذيل الأيمن (**ي** > ١,٦٤٥). فإننا نجد أن قيمة **ي** المحسوبة (١,٧٥) تقع في منطقة الرفض. أي أننا

نرفض فرضية العدم ونقبل الفرضية البديلة القائلة بأن نسبة الطلاب المدخنين في هذه الكلية تزيد على ٩٪ وذلك عند مستوى معنوية ٠,٠٥. وهذا يؤكد صحة النتائج التي توصلت إليها إحدى الدراسة من قبل.

مثال (١٠):

يدّعي أحد المرشحين لعضوية مجلس الشعب بأنه سوف يحصل على تأييد ٧٥٪ من ناخبي دائرته. بينما يرى منافسوه بأن نسبة المؤيدين له لن تبلغ تلك النسبة (٧٥٪). وللتحقق من ذلك تم اختيار عينة مكونة من ٤٥٠ ناخباً من هذه الدائرة فوجد أن من بينهم ٣٢٥ ناخباً يؤيدون هذا المرشح.

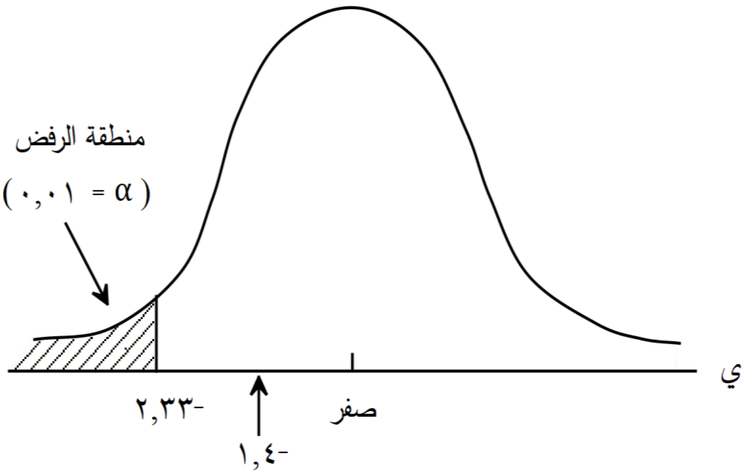
أيهما على حق؟ المرشح أم منافسوه؟ استخدم مستوى معنوية ٠,٠١.

الحــــــــل:

$$\hat{\text{ل}} = \frac{٣٢٥}{٤٥٠} = ٠,٧٢٢ \quad ، \quad \text{ن} = ٤٥٠ \quad ، \quad \text{ل}_{.} = ٠,٧٥$$

$$\text{ف}_{.}: \text{ل} = ٠,٧٥ \quad ، \quad \text{ف}_{١}: \text{ل} < ٠,٧٥$$

$$\text{ي} = \frac{\hat{\text{ل}} - \text{ل}_{.}}{\sqrt{\frac{\text{ل}_{.}(١ - \text{ل}_{.})}{\text{ن}}}} = \frac{٠,٧٢٢ - ٠,٧٥}{\sqrt{\frac{٠,٧٥ \times ٠,٢٥}{٤٥٠}}} = \frac{-٠,٠٢٨}{٠,٠٢} = -١,٤$$



تقع منطقة الرفض عند الذيل الأيسر (ي < - ٢,٣٣). وحيث أن قيمة ي المحسوبة، وهي - ١,٤، تقع في منطقة عدم الرفض حيث أنها أكبر من - ٢,٣٣، فإننا لا نرفض فرضية العدم والتي تفيد بأن المرشح على حق في حين أن منافسيه قد جانبهم الصواب في ادعائهم بأن نسبة المؤيدين له لن تبلغ ٧٥٪ من إجمالي الناخبين.

ثالثاً: اختبار الفرضيات الإحصائية حول الفرق بين متوسطيْ مجتمعيْن

كثيراً ما نجد أنفسنا في حاجة لإجراء مقارنة بين مجتمعيْن. كأَنْ نقارن مثلاً أداء مجموعة من الطالبات في اختبار معين بأداء مجموعة من الطلاب الذكور في نفس الاختبار. أو نريد الحكم على أن قوة تحمل منتَج معين تفوق نظيرتها في منتَج آخر. إلى غير ذلك من الأمثلة التي تندرج تحت الحكم على الفرق بين متوسطي مجتمعيْن من حيث كونه حقيقياً (جوهرياً) أم أن هذا الفرق يرجع إلى عوامل الصدفة.

وفي دراستنا لاختبار الفرضيات الإحصائية المتعلقة بالفرق بين متوسطي مجتمعيْن ($م_١ - م_٢$) تلزم الإشارة إلى نقطتيْن هامتيْن. أولاهما أنه عند سحب عينة من مجتمع، وعينة أخرى من مجتمع ثانٍ، فإنه يفترض أن العينتيْن مستقلتان. وأما الحالات التي لا تكون فيها العينتان مستقلتيْن فإنها تخرج عن نطاق هذا الكتاب. وأما النقطة الثانية فهي أنه في حالة إجراء اختبار ذي ذيليْن على الفرق بين متوسطي مجتمعيْن وكانت النتيجة هي عدم رفض فرضية العدم. فإن ذلك يكون مكافئاً لوقوع قيمة معلمة المجتمع ($م_١ - م_٢$) المفترضة في فرضية العدم ضمن فترة الثقة لهذه المعلمة والتي تعرضنا لكيفية تكوينها في الفصل السابق من هذا الكتاب. وأما وقوع هذه القيمة خارج فترة الثقة فإنه يعني رفض فرضية العدم. وأما في حالة الاختبار ذي الذيل الواحد فإن الأمر ربما يستلزم بعض الحرص عند محاولة الربط بين أسلوبي فترة الثقة واختبار الفرضيات وصولاً إلى نتيجة واحدة فيما يتعلق بالفرق بين متوسطيْ مجتمعيْن.

هذا ووفقاً لاختبار معنوية الفرق بين متوسطيْ مجتمعيْن فإننا نقوم بسحب عينة من المجتمع الأول وعينة من المجتمع الثاني ثم نقوم بحساب متوسط كل عينة. فإذا كان $م_١$، $م_٢$ هما متوسطا المجتمعيْن الأول والثاني على التوالي. كما كانت $\bar{س}_١$، $\bar{س}_٢$ هما متوسطا العينتيْن المسحوبتيْن من هذيْن المجتمعيْن على الترتيب. فإن الفرق بين متوسطي العينتيْن ($\bar{س}_١ - \bar{س}_٢$)، - وكما عرفنا في الفصل السابق - هو متغير عشوائي متوسطه ($م_١ - م_٢$). وإذا كانت العينتان مستقلتيْن فإن الانحراف المعياري للمتغير ($\bar{س}_١ - \bar{س}_٢$) هو: $\sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}}$ حيث $\sigma_١^٢$ ، $\sigma_٢^٢$ هما تباينا المجتمعيْن الأول والثاني على التوالي. كما أن $ن_١$، $ن_٢$ هما حجم كلٍّ من العينتيْن الأولى والثانية على الترتيب. وإذا كان توزيع كل من المجتمعيْن معتاداً فإن توزيع المعاينة للفرق بين متوسطي العينتيْن ($\bar{س}_١ - \bar{س}_٢$) هو توزيع معتاد بمتوسط ($م_١ - م_٢$) وانحراف معياري $\sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}}$ وذلك بغض النظر عن حجم العينة صغيراً كان أم كبيراً.

وأما إذا كان حجم كل من العينتيْن كبيراً كبراً كافياً ($ن_١$، $ن_٢ \geq ٣٠$) فإنه وفقاً لنظرية الحد المركزية نجد أن توزيع المعاينة للمتغير ($\bar{س}_١ - \bar{س}_٢$) هو توزيع قريب من المعتاد له وسط حسابي وانحراف معياري سبق الإشارة إليهما وذلك بغض النظر عن توزيع كل من المجتمعيْن معتاداً كان أم غير ذلك.

ولاختبار معنوية الفرق بين متوسطي مجتمعيْن ($م_١ - م_٢$) فإن المختبِر الإحصائي في هذه الحالة يأخذ الصورة:

$$ي = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}}}$$

وذلك إذا كان تباين كل من المجتمعيْن معروفاً. وهذا المختبر (ي) يتبع توزيعاً معتاداً معيارياً. وأما إذا كان التباينان مجهوليْن وحجم كل من العينة كبير فإنه يمكننا استخدام $ع_١$، $ع_٢$ كتقريب لـ $\sigma_١$، $\sigma_٢$ على الترتيب.

وفي هذه الحالة يأخذ المختبِر الإحصائي الصورة:

$$ي = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{ع_١^٢}{ن_١} + \frac{ع_٢^٢}{ن_٢}}}$$

وهو أيضا له توزيع معتاد معياري.

وأما إذا كان التباينان مجهوليْن وحجم كل من العينتيْن صغير ($ن_١ < ٣٠$، $ن_٢ < ٣٠$) فإن المختبِر الإحصائي في هذه الحالة يأخذ الصورة:

$$ت = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{ع_م^٢\left(\frac{١}{ن_١} + \frac{١}{ن_٢}\right)}}$$

حيث $ع_م^٢ = \frac{(ن_١ - ١)ع_١^٢ + (ن_٢ - ١)ع_٢^٢}{ن_١ + ن_٢ - ٢}$

وذلك بافتراض تساوي التباينيْن في المجتمعيْن اللذين سُحبت منهما العينتان.
والمختبِر الإحصائي في هذه الحالة يتبع توزيع ت بدرجات حرية $ن_١ + ن_٢ - ٢$.

وفيما يلي الحالات المختلفة التي سيتم تتناولها في هذا الشأن والموضحة في الجدوليْن (٣ - ٦) و (٣ - ٧).

الجدول (٣ - ٦)
اختبار الفرضيات حول الفرق بين متوسطي مجتمعيْن باستخدام التوزيع المعتاد (اختبار ي)

<table>
<tr><th>نوع الاختبار / بيان</th><th>اختبار ذو ذيل واحد
(الفرضية البديلة متجهة)</th><th>اختبار ذو ذيليْن
(الفرضية البديلة غير متجهة)</th></tr>
<tr><td>فرضية العدم والفرضية البديلة</td><td>ف$_٠$: (م$_١$ − م$_٢$) = صفر
ف$_١$: (م$_١$ − م$_٢$) $<$ صفر
أو ف$_١$: (م$_١$ − م$_٢$) $>$ صفر</td><td>ف$_٠$: (م$_١$ − م$_٢$) = صفر
ف$_١$: (م$_١$ − م$_٢$) $\neq$ صفر</td></tr>
<tr><td>المختبِـــــــر الإحصائــــــي</td><td colspan="2">$ي = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{\sigma_١^٢}{ن_١} + \frac{\sigma_٢^٢}{ن_٢}}}$
أو باستخدام ع$_١$، ع$_٢$ كتقريب لـ $\sigma_١$، $\sigma_٢$ على التوالي.</td></tr>
<tr><td>منطقة الرفـــض</td><td>في حالة:
ف$_١$: (م$_١$ − م$_٢$) $<$ صفر
ي $<$ -ي$_{\alpha}$
في حالة:
ف$_١$: (م$_١$ − م$_٢$) $>$ صفر
ي $>$ ي$_{\alpha}$
حيث: ϕ(ي$_{\alpha}$) = ١ − α</td><td>ي $<$ -ي$_{\frac{\alpha}{٢}}$
أو:
ي $>$ ي$_{\frac{\alpha}{٢}}$
حيث:
ϕ(ي$_{\frac{\alpha}{٢}}$) = ١ − $\frac{\alpha}{٢}$</td></tr>
<tr><td>شروط التطبيـــق</td><td colspan="2">في حالة العينات الكبيرة (ن$_١$، ن$_٢$ $\geq$ ٣٠):
العينتان عشـوائيتان ومسـتقلتان. حينئذ يمكن اسـتخدام اختبار ي بغض النظر عن توزيع المجتمع معتاداً كان أم غير ذلك. كما يمكن أيضاً استخدام ع$_١$، ع$_٢$ كتقريب جيد لـ $\sigma_١$، $\sigma_٢$.
في حالة العينات الصغيرة (ن$_١$، ن$_٢$ $<$ ٣٠):
العينتان عشـوائيتان ومسـتقلتان - توزيع كل من المجتمعيْن معتاد - تباين كل من المجتمعيْن معلوم.</td></tr>
</table>

الجدول (٣ - ٧)

اختبار الفرضيات حول الفرق بين متوسطي مجتمعيْن باستخدام التوزيع ت (اختبار ت)

نوع الاختبار / بيان	اختبار ذو ذيل واحد (الفرضية البديلة متجهة)	اختبار ذو ذيليْن (الفرضية البديلة غير متجهة)
فرضية العدم والفرضية البديلة	$\text{ف}_{.}: (\text{م}_{١} - \text{م}_{٢}) = \text{صفر}$ $\text{ف}_{١}: (\text{م}_{١} - \text{م}_{٢}) < \text{صفر}$ أو $\text{ف}_{١}: (\text{م}_{١} - \text{م}_{٢}) > \text{صفر}$	$\text{ف}_{.}: (\text{م}_{١} - \text{م}_{٢}) = \text{صفر}$ $\text{ف}_{١}: (\text{م}_{١} - \text{م}_{٢}) \neq \text{صفر}$
المختبِـــــــر الإحصائـــــي	$\text{ت} = \dfrac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\sqrt{\text{ع}^{٢}_{\text{م}}\left(\dfrac{١}{\text{ن}_{١}} + \dfrac{١}{\text{ن}_{٢}}\right)}}$ حيث $\text{ع}^{٢}_{\text{م}} = \dfrac{(\text{ن}_{١} - ١)\text{ع}^{٢}_{١} + (\text{ن}_{٢} - ١)\text{ع}^{٢}_{٢}}{\text{ن}_{١} + \text{ن}_{٢} - ٢}$ وذلك بافتراض تساوي التباينيْن في المجتمعيْن اللذين سُحبت منهما العينتان.	
منطقة الرفـــض	في حالة: $\text{ف}_{١}: (\text{م}_{١} - \text{م}_{٢}) < \text{صفر}$ $\text{ت} < -\text{ت}_{\alpha}$ في حالة: $\text{ف}_{١}: (\text{م}_{١} - \text{م}_{٢}) > \text{صفر}$ $\text{ت} > \text{ت}_{\alpha}$ حيـث تتوقف قيمـة ت_{α} على درجات الحرية $\text{ن}_{١} + \text{ن}_{٢} - ٢$	$\text{ت} < -\text{ت}_{\frac{\alpha}{٢}}$ أو: $\text{ت} > \text{ت}_{\frac{\alpha}{٢}}$ حيـث تتوقف قيمـة $\text{ت}_{\frac{\alpha}{٢}}$ على درجات الحرية $\text{ن}_{١} + \text{ن}_{٢} - ٢$
شروط التطبيـــق	١- العينتان عشوائيتان مستقلتان وصغيرتان ($\text{ن}_{١}$، $\text{ن}_{٢}$ < ٣٠). ٢- كلٌ من المجتمعيْن المســحوب منهما العينتان له توزيع معتاد أو قريب من المعتاد. ٣- تساوى التباينيْن في المجتمعيْن. ٤- التباين في كل من المجتمعيْن غير معلوم.	

وفيما يلي بعض الأمثلة التي تلقى الضوء على كيفية إجراء اختبار الفرضيات حول الفرق بين متوسطي مجتمعيْن آخذين في الاعتبار الحالات المختلفة التي شملها الجدولان السابقان.

مثال (١١):

في دراسة لتقييم أسلوبيْن من أساليب تدريس مادة الإحصاء التطبيقي، أُختيرت إحدى الكليات وطُبق فيها الأسلوب الأول، كما أُختيرت كلية ثانية وطُبق فيها الأسلوب بالثاني. ثم أُعطي جميع الطلاب نفس الاختبار فكانت النتائج التالية:

بيان	الكلية الأولى	الكلية الثانية
عـــدد الطـــلاب	١٠٠٠	١٢٠٠
متوسط الدرجـــات	٧٥	٧٤
الانحراف المعياري	١٢	١٠

بافتراض أن طلاب كل كلية يمكن اعتبارهم عينة عشوائية من مجتمع الطلاب عموماً وأن مستوياتهم في الاستيعاب واحدة، هل هناك ما يدل على تفوق أحد أساليب التدريس على الأسلوب الآخر؟ استخدم مستوى معنوية ٠,٠٥

الحـــــل:

من بيانات العينتيْن نجد أن:

$$\bar{س}_{١} = ٧٥ \quad ، \quad ع_{١} = ١٢ \quad ، \quad ن_{١} = ١٠٠٠$$

$$\bar{س}_{٢} = ٧٤ \quad ، \quad ع_{٢} = ١٠ \quad ، \quad ن_{٢} = ١٢٠٠$$

كما يمكن صياغة فرضية العدم والفرضية البديلة كما يلي:

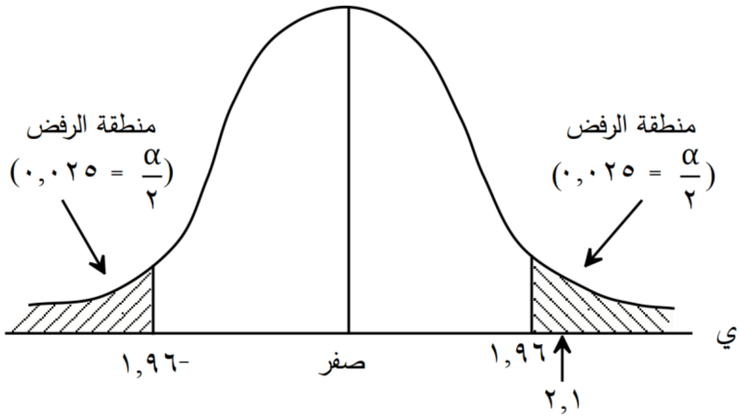
$$ف_{.}: م_{١} - م_{٢} = صفر \quad ، \quad ف_{١}: م_{١} - م_{٢} \neq صفر$$

وبافتراض صحة فرضية العدم (ف.) نجد أن:

$$\text{المختبِر الإحصائي (ي)} = \frac{\bar{س}_{٢} - \bar{س}_{٢}}{\sqrt{\frac{ع_{١}^{٢}}{ن_{١}} + \frac{ع_{٢}^{٢}}{ن_{٢}}}}$$

$$= \frac{٧٥ - ٧٤}{\sqrt{\frac{(١٢)^{٢}}{١٠٠٠} + \frac{(١٠)^{٢}}{١٢٠٠}}} = \frac{١}{٠,٤٧٧} = ٢,١$$

الاختبار هنا ذو ذيليْنا حيث تقع منطقة الرفض عند الذيليْن ($\text{ي} < -\text{ي}_{\frac{\alpha}{\text{٢}}}$ و $\text{ي} > \text{ي}_{\frac{\alpha}{\text{٢}}}$). وحيث أن α = ٠,٠٥ فإن $\text{ي}_{\frac{\alpha}{\text{٢}}}$ = ١,٩٦.



وحيث أن قيمة **ي** المحسوبة (٢,١) هي أكبر من ١,٩٦ فإنه يمكننا رفض فرضية العدم وبالتالي قبول الفرضية البديلة. وهذا يعني أن أحد الأسلوبيْن يتفوق على الآخر من حيث استيعاب الطلاب له.

مثـــال (١٢):

باستخدام بيانات المثال السابق.

المطلوب:

أولاً: إجراء الاختبار المطلوب في مثال (١٠) مستخدماً في ذلك فترة الثقة.

ثانياً: هل يمكن القول بأن الأسلوب الأول في التدريس هو أفضل من الأسلوب الثاني؟

الحـــل:

أولاً: يمكن تكوين فترة الثقة للفرق بين متوسطي المجتمعيْن وذلك على النحو الموضح في الفصل السابق من هذا الكتاب حيث نجد أن:

$$\text{الحد الأدنى لفترة الثقة} = (\bar{\text{س}}_{\text{١}} - \bar{\text{س}}_{\text{٢}}) - \text{ي}_{\frac{\alpha}{\text{٢}}} \sqrt{\frac{\text{ع}_{\text{١}}^{\text{٢}}}{\text{ن}_{\text{١}}} + \frac{\text{ع}_{\text{٢}}^{\text{٢}}}{\text{ن}_{\text{٢}}}}$$

$$= (\text{٧٥} - \text{٧٤}) - \text{١,٩٦} \sqrt{\frac{(\text{١٢})^{\text{٢}}}{\text{١٠٠٠}} + \frac{(\text{١٠})^{\text{٢}}}{\text{١٢٠٠}}}$$

$$= \text{١} - \text{١,٩٦} \times \text{٤,٤٧٧} = \text{٠,٠٦٥}$$

$$\text{الحد الأعلى لفترة الثقة} = (\bar{\text{س}}_{\text{١}} - \bar{\text{س}}_{\text{٢}}) + \text{ي}_{\frac{\alpha}{\text{٢}}} \sqrt{\frac{\text{ع}_{\text{١}}^{\text{٢}}}{\text{ن}_{\text{١}}} + \frac{\text{ع}_{\text{٢}}^{\text{٢}}}{\text{ن}_{\text{٢}}}}$$

$$= (٧٥ - ٧٤) + ١,٩٦ \sqrt{\frac{(١٢)^٢}{١٠٠٠} + \frac{(١٠)^٢}{١٢٠٠}}$$

$$= ١ + ١,٩٦ \times ٤,٤٧٧ = ١,٩٣٥$$

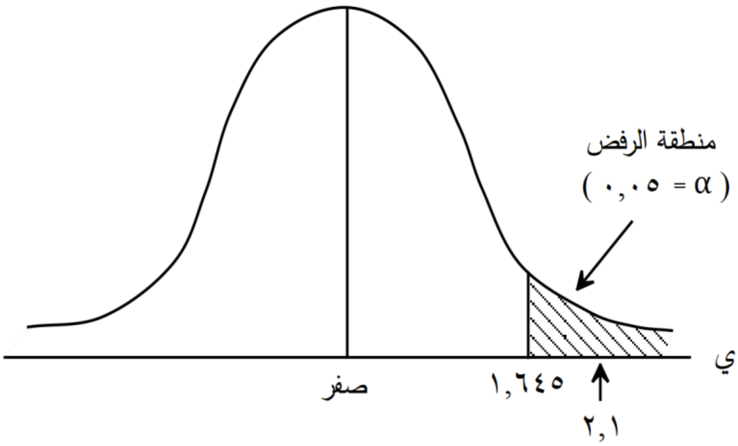
أي أن الفرق بين متوسطي المجتمعيْن يتراوح ما بين (٠,٠٦٥، ١,٩٣٥). وحيث أن قيمة معلمة المجتمع ($م_١ - م_٢$) في فرضية العدم هي الصفر وهي قيمة لا تدخل ضمن فترة الثـقة.

فإن ذلك يعني رفضنا لفرضية العدم وقبولنا للفرضية البديلة والتي تفيد بأن أحد الأسلوبيْن أفضل من الآخر.

ثانياً: حيث أن المطلوب هنا هو اختبار ما إذا كان الأسلوب الأول أفضل من الثاني فإن فرضية العدم والفرضية البديلة يمكن صياغتهما على النحو التالي:

ف.: $م_١ - م_٢$ = صفر ، **ف**$_١$: $م_١ - م_٢$ > صفر

وقد سبق إيجاد قيمة المختبِر الإحصائي (**ي**) ووجد أنها تساوي ٢,١، أي أن: **ي** = ٢,١.



الاختبار هنا ذو ذيل واحد حيث تقع منطقة الرفض عند الذيل الأيمن (**ي** > **ي**$_{\frac{\alpha}{٢}}$) وحيث

ي$_{\frac{\alpha}{٢}}$ = ١,٦٤٥. وبذلك يمكننا استنتاج أن القيمة المحسوبة للمختبر الإحصائي (**ي** = ٢,١)

تقع في منطقة الرفض. ومن ثم نرفض فرضية العدم ونقبل الفرضية البديلة والتي تقول بأن الأسلوب الأول في التدريس هو أفضل من الأسلوب الثاني.

مثـــال (١٣):

في مصنع لإنتاج اللمبات الكهربائية استخدمـــت طريقتان مختلفتان للإنتاج (أ و ب).
فإذا اخترنا عينة من اللمبات المنتَجة باستخدام كل من الطريقتيْن فكانت النتائج كما يلي:

البيان / الطريقة	حجم العينة	متوسط أعمال المصابيح في العينة (ساعة)	الانحراف المعياري لعمر المصابيح
أ	١٢	١٥٠٠	٦٠
ب	١٥	١٥٥٠	٥٠

بافتراض أن عمر المصابيح يتبع توزيعاً معتاداً، اختبر صحة الفرضية القائلة بأن الطريقة (ب) تنتِج مصابيح أطول عمراً من تلك المنتَجة بواسطة الطريقة (أ) وذلك عند مستوى معنوية ٠,٠١.

الحـــــل:

من بيانات العينتيْن نجد أن:

$$\bar{\text{س}}_{١} = ١٥٠٠ \quad ، \quad \text{ع}_{١} = ٦٠ \quad ، \quad \text{ن}_{١} = ١٢$$

$$\bar{\text{س}}_{٢} = ١٥٥٠ \quad ، \quad \text{ع}_{٢} = ٥٠ \quad ، \quad \text{ن}_{٢} = ١٥$$

حيث أن المطلوب هو اختبار ما إذا كانت الطريقة (ب) تنتج مصابيح أطول عمراً من الطريقة (أ) فإنه يمكن صياغة الفرضيات على النحو التالي:

$$\text{ف}_{٠}: \text{م}_{١} - \text{م}_{٢} = \text{صفر} \quad ، \quad \text{ف}_{١}: \text{م}_{١} - \text{م}_{٢} < \text{صفر}$$

وبافتراض تساوي التباينيْن لأعمار المصابيح المنتَجة باستخدام الطريقتيْن فإن التباين المشترك هو:

$$\text{ع}_{\text{م}}^{٢} = \frac{(\text{ن}_{١} - ١)\text{ع}_{١}^{٢} + (\text{ن}_{٢} - ١)\text{ع}_{٢}^{٢}}{\text{ن}_{١} + \text{ن}_{٢} - ٢}$$

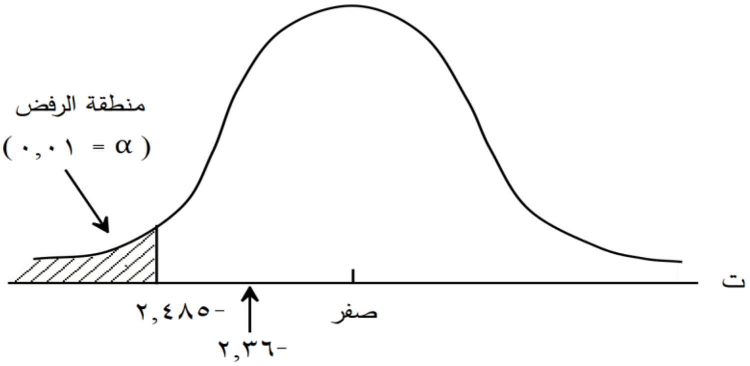
$$= \frac{(١٢ - ١)(٦٠)^{٢} + (١٥ - ١)(٥٠)^{٢}}{١٢ + ١٥ - ٢} = \frac{٧٤٦٠٠}{٢٥} = ٢٩٨٤$$

وحيث أن حجم كل من العينتيْن صغير ($\text{ن}_{١}$، $\text{ن}_{٢}$ < ٣٠)، كما أن التباين مجهول لكل من المجتمعيْن. هذا بالإضافة إلى أن توزيع أعمار المصابيح في كل من الطريقتيْن يتبع التوزيع المعتاد، فإن المختبِر الإحصائي الملائم هو:

$$\text{ت} = \frac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\sqrt{\text{ع}_{\text{م}}^{٢}\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}}$$

$$= \frac{١٥٠٠ - ١٥٥٠}{\sqrt{٢٩٨٤\left(\frac{١}{١٢} + \frac{١}{١٥}\right)}} = \frac{-٥٠}{٢١,١٦} = -٢,٣٦$$

الاختبار في هذه الحالة ذو ذيل واحد حيث تقع منطقة الرفض عند الذيل الأيسر ($ت < -ت_{\frac{\alpha}{٢}}$) وذلك عند درجات حرية ($ن_١ + ن_٢ - ٢$).



ومن جدول المساحة تحت منحنى توزيع ت نجد أن قيمة $ت_{٠,٠١}$ عند درجات حرية ٢٥ هي ٢,٤٨٥.

وبذلك يمكننا استنتاج أن قيمة المختبِر الإحصائي المسحوبة (-٢,٣٦) تقع داخل منطقة عدم الرفض. أي أننا لا نرفض فرضية العدم القائلة بأن متوسط أعمار المصابيح واحد في كل من الطريقتيْن وأن القول بأن الطريقة (ب) تنتِج مصابيح أطول عمراً هو قول ليس هناك ما يؤكده.

مثـــال (١٤):

باستخدام بيانات المثال السابق، أجْرِ الاختبار التالي:

$ف_٠: م_١ - م_٢ = صفر$ ، $ف_١: م_١ - م_٢ \neq صفر$

مستخدماً في ذلك فترة الثقة ٩٩٪.

الحـــل:

فترة الثقة للفرق بين متوسطي المجتمعيْن ($م_١ - م_٢$) هي:

$$(\bar{س}_١ - \bar{س}_٢) \pm ت_{(ن_١ + ن_٢ - ٢)، \frac{\alpha}{٢}} \sqrt{ع_م^٢ \left(\frac{١}{ن_١} + \frac{١}{ن_٢}\right)}$$

أي أن:

$$\text{فترة الثقة} = (١٥٠٠ - ١٥٥٠) \pm ت_{٢٥، ٠,٠٠٥} \sqrt{٢٩٨٤ \left(\frac{١}{١٢} + \frac{١}{١٥}\right)}$$

$$= -٥٠ \pm ٢,٧٨٧ \times ٢١,١٦ = -٥٠ \pm ٥٨,٩٧$$

$$= [-١٠٨,٩٧ ، + ٨,٩٧]$$

وحيث أن قيمة معلمة المجتمع ($م_١ - م_٢$) في فرضية العدم وهي الصفر تقع داخل فترة الثقة فإن ذلك يعني قبولنا لفرضية العدم والتي تفيد بأن متوسط أعمار المصابيح المنتَّجة باستخدام الطريقة (أ) لا يختلف عن نظيره في حالة المصابيح المنتَّجة باستخدام الطريقة (ب).

مثال (١٥):

أُختيرت عينة مكوَّنة من ٢٥٠ شخصاً من أحد المجتمعات فكان الوسط الحسابي والانحراف المعياري لأطوالهم ١٧٥ سم و ١٨ سم على الترتيب. كما أُخذت عينة أخرى مكوَّنة من ١٨٠ شخصاً من مجتمع آخر فوُجد أن الوسط الحسابي والانحراف المعياري لأطوال الأشخاص فيها ١٧٣ سم و ٢٠سم على التوالي. بافتراض أن الأشخاص في العينتيْن لهم نفس الفئات العمرية، هل يمكن القول بأن المجتمع الأول يتمتع أفراده بطول أكبر؟ استخدم في ذلك مستوى معنوية ٠,١٠.

الحـــل:

من بيانات العينتيْن نجد أن:

$$\bar{س}_١ = ١٧٥ \quad ، \quad ع_١ = ١٨ \quad ، \quad ن_١ = ٢٥٠$$

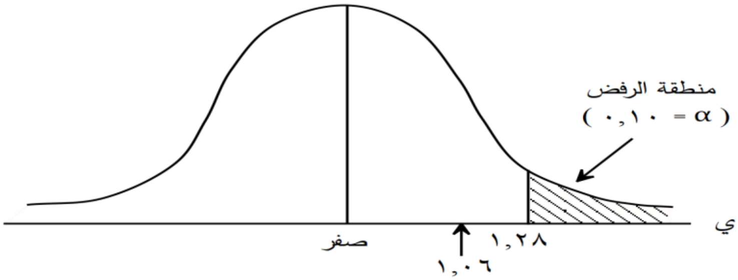
$$\bar{س}_٢ = ١٧٣ \quad ، \quad ع_٢ = ٢٠ \quad ، \quad ن_٢ = ١٨٠$$

$$ف_٠ : م_١ - م_٢ = \text{صفر} \quad ، \quad ف_١ : م_١ - م_٢ > \text{صفر}$$

$$\text{المختبِر الإحصائي (ي)} = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\frac{ع_١^٢}{ن_١} + \frac{ع_٢^٢}{ن_٢}}}$$

$$= \frac{١٧٥ - ١٧٣}{\sqrt{\frac{(١٨)^٢}{٢٥٠} + \frac{(٢٠)^٢}{١٨٠}}} = \frac{٢}{١,٨٨} = ١,٠٦$$

الاختبار هنا ذو ذيل واحد حيث تقع منطقة الرفض عند الذيل الأيمن ($ي > ي_\alpha$) وحيث أن مستوى المعنوية المطلوب هو α= ٠,١٠. فإن $ي_\alpha$ = ١,٢٨.



وبذلك يمكننا استنتاج أن قيمة **ي** المحسوبة (١,٠٦) تقع داخل منطقة عدم الرفض. أي أننا نقبل فرضية العدم والتي مؤداها أن متوسط أطوال الأفراد في المجتمع الأول لا يختلف عن متوسط أطوال الأفراد في المجتمع الثاني. وأن القول بأن أفراد المجتمع الأول أكثر طولاً هو قول تنقصه الأدلة وذلك بمستوى معنوية ٠,٠١.

مثال (١٦):

وضِّح ماذا يحدث بشأن القرار المتَّخذ في المثال السابق لو أن مستوى المعنوية أصبح ٠,٠٥ بدلاً من ٠,١٠؟

الحـــــل:

حيث أن قيمة **ي** المحسوبة في المثال السابق كانت أقل من **ي**$_{٠,١٠}$(١,٢٨)، وبما أن منطقة الرفض تتحدد بـ (**ي** > **ي**$_{\alpha}$). فإن استخدام أي مستوى معنوية أقل من ٠,١٠ من شأنه أن يحتفظ بقرار فرضية العدم كما هو. إذ أن القيمة (١,٠٦) سوف تقع دائماً في منطقة عدم الرفض وذلك لأن قيمة **ي**$_{\alpha}$ في تلك الحالات سوف تزيد على ١,٢٨ لأي مستوى معنوية أقل من ٠,١٠.

رابعاً: اختبارات الفرضيات الإحصائية حول الفرق بين نسبتيْ مجتمعيْن

إن اختبار الفرضيات الإحصائية فيما يتعلق بالفرق بين نسبتيْ مجتمعيْن على قدر كبير من الأهمية في التطبيقات العملية. إذا أننا وفي كثير من البحوث والدراسات نحتاج إلى معرفة ما إذا كانت نسبة الذين يتمتعون بخاصية معينة متساوية في مجتمعيْن أم لا. فَعلى سبيل المثال، قد تواجهنا مثل التساؤلات التالية:

- هل يمكننا القول بأن الذكور أكثر تغيباً عن المدارس من الإناث؟
- إلى أي حد يمكننا التسليم بصحة الرأي الطبي القائل بأن التدخين سبب في الإصابة بأمراض الصدر؟
- هل يمكن القول بأن نسبة الأمية واحدة في كل من مدينتي القاهرة والاسكندرية؟ أم أنها تميل إلى زيادة في إحدى المدينتيْن عن الأخرى؟
- ما هو مدى صحة ادعاء صاحب مصنع ما بأن نسبة الوحدات التالفة في إنتاج مصنعه هي أقل من نظيرتها في المصانع المشابهة الأخرى؟
- هل يمكن القول بأن ظاهرة التدخين هي أكثر تفشياً بين شباب المدن عنها بين شباب القرى؟

إلى غير ذلك من التساؤلات والتي نستطيع الإجابة عليها من خلال تطبيق الاختبارات الإحصائية فيما يتعلق بالفرق بين نسبتي مجتمعيْن.

هذا وتقوم نظرية اختبار الفرق بين نسبتيْن على ما يلي:
إذا سُحبت عينتان مستقلتان من مجتمعيْن، وكان حجم العينة المسحوبة من المجتمع الأول $ن_١$ وحجم العينة المسحوبة من المجتمع الثاني $ن_٢$. فإذا كنا معنيين بخاصية معينة في كل من المجتمعيْن، وكانت نسبة الذين يتمتعون بهذه الخاصية في المجتمع الأول $ل_١$ ونسبة الذين يتمتعون بها في المجتمع الثاني $ل_٢$. وإذا أردنا اختبار أن $ل_١ = ل_٢$، أي اختبار أن النسبة في المجتمعيْن واحدة. فإننا نحسب نسبة الذين يمتتعون بهذه الخاصية في كل من العينتيْن ولنفترض أنهما $\hat{ل}_١$ ، $\hat{ل}_٢$، على الترتيب.
وباستخدام بيانات العينتيْن وبافتراض صحة فرضية العدم، أي بافتراض أن: $ل_١ = ل_٢ = ل$، فإن $\hat{ل}$ وهي مقدِّر للنسبة ل في المجتمع تتحدد كما يلي:

$$\hat{ل} = \frac{ن_١\hat{ل}_١ + ن_٢\hat{ل}_٢}{ن_١ + ن_٢} = \frac{س_١ + س_٢}{ن_١ + ن_٢}$$

حيث: $س_١$ عدد الذين يتمتعون بهذه الخاصية في المجتمع الأول،
$س_٢$ عدد الذين يتمتعون بهذه الخاصية في المجتمع الثاني.
كما أن المقدِّر:

$$ع_{(\hat{ل}_١ - \hat{ل}_٢)} = \sqrt{\hat{ل}\left(١ - \hat{ل}\right)\left(\frac{١}{ن_١} + \frac{١}{ن_٢}\right)}$$

هو مقدِّر للخطأ المعياري للفرق بين النسبتيْن وذلك بافتراض صحة فرضية العدم.
وإذا كان حجم كل من العينتيْن كبير كبراً كافياً فإن المختبِر الإحصائي:

$$ي = \frac{\hat{ل}_١ - \hat{ل}_٢}{ع_{(\hat{ل}_١ - \hat{ل}_٢)}}$$

تمثل متغيراً يتبع التوزيع المعتاد المعياري.

وقبل أن نتناول بعض الأمثلة في هذا الشأن، فإنه من الملائم أن نوجز الحالات المختلفة التي يمكن أن تواجهنا وذلك على النحو الموضح في الجدول (٣ - ٨). وتقتصر الحالات هنا على العينات كبيرة الحجم وذلك للأسباب التي ورد ذكرها من قبل عند تناولنا لاختبار الفرضيات حول النسبة في المجتمع.

الجدول (٣ - ٨)

اختبار الفرضيات حول الفرق بين نسبتي مجتمعيْن باستخدام التوزيع المعتاد (اختباري)

نوع الاختبار / بيان	اختبار ذو ذيل واحد (الفرضية البديلة متجهة)	اختبار ذو ذيليْن (الفرضية البديلة غير متجهة)
فرضية العدم والفرضية البديلة	$\text{ف}_{.}: (\text{ل}_{١} - \text{ل}_{٢}) = \text{صفر}$ $\text{ف}_{١}: (\text{ل}_{١} - \text{ل}_{٢}) < \text{صفر}$ أو $\text{ف}_{١}: (\text{ل}_{١} - \text{ل}_{٢}) > \text{صفر}$	$\text{ف}_{.}: (\text{ل}_{١} - \text{ل}_{٢}) = \text{صفر}$ $\text{ف}_{١}: (\text{ل}_{١} - \text{ل}_{٢}) \neq \text{صفر}$
المختبِر الإحصائي	$\text{ي} = \dfrac{\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢}}{\sigma_{(\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢})}}$ حيث: $\sigma_{(\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢})} = \sqrt{\dfrac{\hat{\text{ل}}_{١}(١ - \hat{\text{ل}}_{١})}{\text{ن}_{١}} + \dfrac{\hat{\text{ل}}_{٢}(١ - \hat{\text{ل}}_{٢})}{\text{ن}_{٢}}}$ $= \sqrt{\hat{\text{ل}}(١ - \hat{\text{ل}})\left(\dfrac{١}{\text{ن}_{١}} + \dfrac{١}{\text{ن}_{٢}}\right)}$ وذلك بافتراض صحة فرضة العدم. حيث: $\hat{\text{ل}} = \dfrac{\text{ن}_{١}\hat{\text{ل}}_{١} + \text{ن}_{٢}\hat{\text{ل}}_{٢}}{\text{ن}_{١} + \text{ن}_{٢}}$	
منطقة الرفـــض	في حالة: $\text{ف}_{١}: (\text{ل}_{١} - \text{ل}_{٢}) < \text{صفر}$ $\text{ي} < -\text{ي}_{\alpha}$ في حالة: $\text{ف}_{١}: (\text{ل}_{١} - \text{ل}_{٢}) > \text{صفر}$ $\text{ي} > \text{ي}_{\alpha}$	$\text{ي} < -\text{ي}_{\frac{\alpha}{٢}}$ أو: $\text{ي} > \text{ي}_{\frac{\alpha}{٢}}$
شروط التطبيـــق	• العينات مستقلة وعشوائية. • حجم كل من العينتيْن كبير، وذلك حتى يمكن استخدام التوزيع المعتاد في عملية الاختبار.	

والآن نتعرض بشيء من التفصيل لبعض الأمثلة التي تغطى الحالات المختفة التي أوجزناها في الجدول (٣ - ٨).

مثال (١٧):

في دراسة لظاهرة التغيب عن المدرسة ومدى ارتباطها بجنس دون آخر (ذكور وإناث)، راقبت إحدى الإدارات التعليمية الغياب في مدرسة للذكور عدد طلابها ٦٠٠ وفي مدرسة للطالبات عدد طالباتها ٥٠٠ طالبة. وفي نهاية العام الدراسي كان عدد الذين تجاوز غيابهم عشرة أيام في مدرسة الذكور ١٢٠ طالباً وفي مدرسة الإناث ٨٥ طالبة.

هل تؤكد هذه البيانات صحة الادعاء القائل بأن الذكور أكثر غياباً من الإناث؟ استخدم مستوى معنوية ٠,٠٥.

الحـــل:

بافتراض أن طلاب مدرسة الذكور يمثلون عينة عشوائية من مجتمع طلاب المدارس عموماً، وأن طالبات مدرسة الإناث يمثلن عينة عشوائية من مجتمع طالبات المدارس ككل. فإنه من بيانات العينتيْن نجد أن:

$$\hat{ل}_{١} = \frac{١٢٠}{٦٠٠} = ٠,٢ \quad ، \quad ن_{١} = ٦٠٠،$$

$$\hat{ل}_{٢} = \frac{٨٥}{٥٠٠} = ٠,١٧ \quad ، \quad ن_{٢} = ٥٠٠$$

وحيث أننا معنيِّين بمعرفة ما إذا كانت نسبة الغياب عن المدرسة هي أكبر عند الذكور منها عند الإناث فإن فرضية العدم والفرضية البديلة يمكن أن تصاغا على النحو التالي:

$$ف_{٠}: ل_{١} - ل_{٢} = صفر \quad ، \quad ف_{١}: ل_{١} - ل_{٢} > صفر$$

وهنا يلاحَظ أن $ل_{١}$ هي النسبة المتعلقة بالذكور وأما $ل_{٢}$ فهي النسبة المتعلقة بالإناث. ومن ثم فإن زيادة نسبة تغيب الذكور عن نظيرتها عند الإناث تعني $ل_{١} > ل_{٢}$، أي أن $(ل_{١} - ل_{٢}) > صفر$.

وبافتراض صحة فرضية العدم (أي $ل_{١} - ل_{٢} = صفر$، بمعنى $ل_{١} = ل_{٢}$) فإننا نجد أن:

$$\hat{ل} = \frac{س_{١} + س_{٢}}{ن_{١} + ن_{٢}} = \frac{١٢٠ + ٨٥}{٦٠٠ + ٥٠٠} = ٠,١٩$$

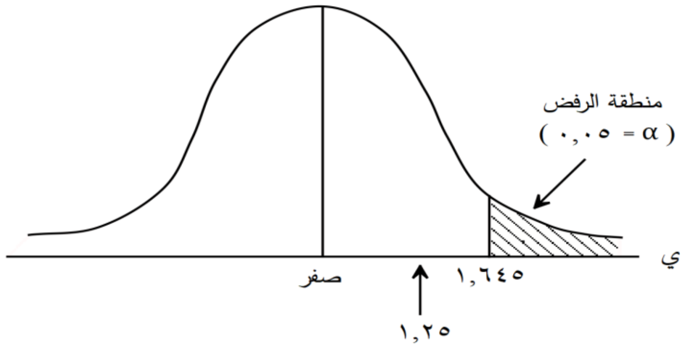
$$أو \quad \hat{ل} = \frac{ن_{١}\hat{ل}_{١} + ن_{٢}\hat{ل}_{٢}}{ن_{١} + ن_{٢}}$$

$$= \frac{٦٠٠ \times ٠,٢ + ٥٠٠ \times ٠,١٧}{٦٠٠ + ٥٠٠} = \frac{١٢٠ + ٨٥}{٦٠٠ + ٥٠٠} = ٠,١٩$$

$$\therefore \; ع_{(\hat{ل}_{١} - \hat{ل}_{٢})} = \sqrt{\hat{ل}\left(١ - \hat{ل}\right)\left(\frac{١}{ن_{١}} + \frac{١}{ن_{٢}}\right)}$$

$$= \sqrt{٠,١٩ \times ٠,٨١ \left(\frac{١}{٦٠٠} + \frac{١}{٥٠٠}\right)} = ٠,٠٢٤$$

وكذلك **ي** = $\frac{\hat{ل}_١ - \hat{ل}_٢}{ع_{(\hat{ل}_١ - \hat{ل}_٢)}} = \frac{٠,٢٠ - ٠,١٧}{٠,٠٢٤} = \frac{٠,٠٣}{٠,٠٢٤} = ١,٢٥$



وحيث أن الاختبار هنا ذو ذيل واحد وأن منطقة الرفض تقع عند الذيل الأيمن (**ي** > ي$_{٠,٠٥}$)، أي (**ي** > ١,٦٤٥) فإننا نجد أن قيمة **ي** المحسوبة (١,٢٥) تقع داخل منطقة عدم الرفض. وهذا يعني عدم رفضنا لفرضية العدم القائلة بأن نسبة التغيب عن المدرسة هي واحدة عند كل من الذكور والإناث، وأن الافتراض القائل بأن الذكور أكثر تغيباً من الإناث هو محض ادعاء ليس هناك ما يؤيده أو يبرره.

مثال (١٨):

أُخذت عينة حجمها ٤٠٠ شخصاً من مجتمع غير المدخنين حيث وُجد أن عدد المصابين بأحد أمراض الصدر هو ١٠٠ شخص. كما أُخذت عينة حجمها ٢٠٠ شخصاً من مجتمع المدخنين فوجد أن عدد المصابين بأحد أمراض الصدر هو ٦٤ شخصاً. بافتراض أن المجتمعيْن المسحوب منهما العينتان جميع ظروفهما - ماعدا التدخين - متشابهة. هل يمكن القول بأن نسبة الإصابة بأحد أمراض الصدر هي واحدة بين الأشخاص المدخنين وغير المدخنين؟ أم أن الشخص المدخن يكون أكثر عرضه للإصابة بأحد هذه الأمراض؟ استخدام مستوى معنوية ٠,٠٥.

الحـــل:

$\hat{ل}_١ = \frac{١٠٠}{٤٠٠} = ٠,٢٥$ ، $ن_١ = ٤٠٠$،

$\hat{ل}_٢ = \frac{٦٤}{٢٠٠} = ٠,٣٢$ ، $ن_٢ = ٢٠٠$

وفقاً للمطلوب في هذا المثال فإن فرضية العدم والفرضية البديلة هما:

$\text{ف}_{٠}$: $\text{ل}_{١} - \text{ل}_{٢}$ = صفر ، $\text{ف}_{١}$: $\text{ل}_{١} - \text{ل}_{٢}$ < صفر

وبافتراض صحة فرضية العدم (أي: $\text{ل}_{١} = \text{ل}_{٢}$) فإننا نجد أن:

$$\hat{\text{ل}} = \frac{\text{١٠٠} + \text{٦٤}}{\text{٤٠٠} + \text{٢٠٠}} = \frac{\text{١٦٤}}{\text{٦٠٠}} = \text{٠,٢٧٣}$$

أو

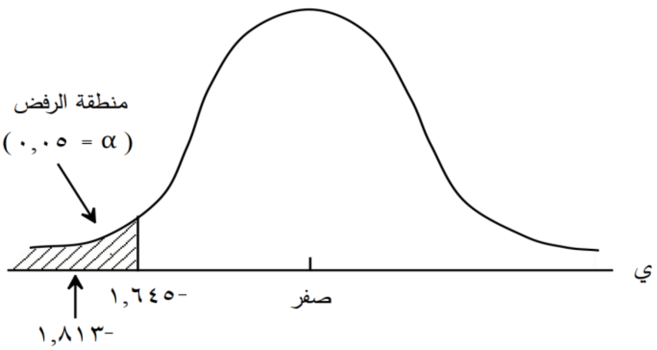
$$\hat{\text{ل}} = \frac{\text{٤٠٠} \times \text{٠,٢٥} + \text{٢٠٠} \times \text{٠,٣٢}}{\text{٤٠٠} + \text{٢٠٠}} = \frac{\text{١٠٠} + \text{٦٤}}{\text{٦٠٠}} = \text{٠,٢٧٣}$$

$$\therefore \ \text{ع}_{(\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢})} = \sqrt{\hat{\text{ل}}\left(\text{١} - \hat{\text{ل}}\right)\left(\frac{\text{١}}{\text{ن}_{١}} + \frac{\text{١}}{\text{ن}_{٢}}\right)}$$

$$= \sqrt{\text{٠,٢٧٣} \times \text{٠,٧٢٧}\left(\frac{\text{١}}{\text{٤٠٠}} + \frac{\text{١}}{\text{٢٠٠}}\right)} = \text{٠,٠٣٨٦}$$

أي أن

$$\text{ي} = \frac{\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢}}{\text{ع}_{(\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢})}} = \frac{\text{٠,٢٥} - \text{٠,٣٢}}{\text{٠,٠٣٨٦}} = \text{-١,٨١٣}$$



وحيث أن الاختبار هنا ذو ذيل واحد، وأن منطقة الرفض تقع عند الذيل الأيسر (**ي** < -$\text{ي}_{٠,٠٥}$)، أي (**ي** > -١,٦٤٥) فإن قيمة **ي** المحسوبة (- ١,٨١٣) تقع داخل منطقة الرفض.

أي أننا نرفض فرضية العدم ونقبل الفرضية البديلة القائلة بأن نسبة تعرض المدخنين بأحد أمراض الصدر هي أكبر من نسبة تعرض غير المدخنين للإصابة بأحد هذه الأمراض وذلك عند مستوى مئوية ٠,٠٥.

مثال (١٩):

سحبت عينة عشوائية مكوَّنة من ٨٠ وحدة من إنتاج أحد المصانع فوُجِد من بينها ٢٤ وحدة معيبة. كما سحبت عينة عشوائية من ١٢٠ وحدة من إنتاج مصنع آخر له نفس المنتَج

فوُجد من بينها ٤٠ وحدة معيبة. باستخدام مستوى معنوية ٠,٠١، هل يمكن القول بأن نسبة الإنتاج المعيب واحدة في المصنعيْن؟ أم أن نسبة الوحدات المعيبة في إنتاج أحد المصنعيْن تزيد عن نظيرتها في المصنع الآخر؟

الحـــــل:

$$\hat{\text{ل}}_{١} = \frac{٢٤}{٨٠} = ٠,٣٠ \quad ، \quad \hat{\text{ل}}_{٢} = \frac{٤٠}{١٢٠} = ٠,٣٣$$

فرضية العدم والفرضية البديلة في هذه الحالة تأخذان الصيغ التالية:

$$\text{ف}_{.}: \text{ل}_{١} - \text{ل}_{٢} = \text{صفر} \quad ، \quad \text{ف}_{١}: \text{ل}_{١} - \text{ل}_{٢} \neq \text{صفر}$$

وبافتراض صحة فرضية العدم نجد أن:

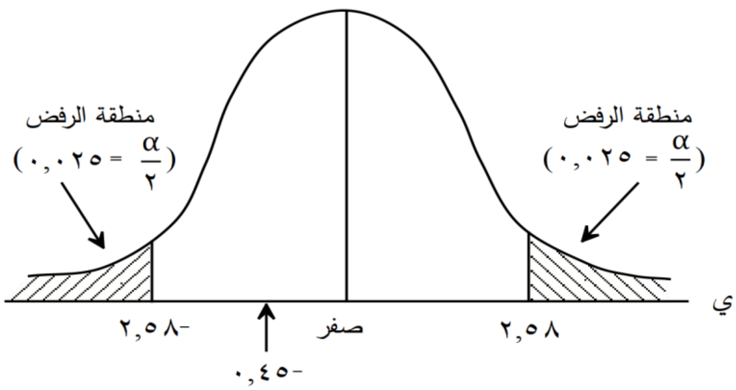
$$\hat{\text{ل}} = \frac{٢٤ + ٤٠}{٨٠ + ١٢٠} = \frac{٦٤}{٢٠٠} = ٠,٣٢$$

$$\text{أو} \quad \hat{\text{ل}} = \frac{٨٠ \times ٠,٣ + ١٢٠ \times ٠,٣٣}{٨٠ + ١٢٠} = \frac{٢٤ + ٤٠}{٢٠٠} = ٠,٣٢$$

$$\therefore \quad \text{ع}_{(\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢})} = \sqrt{\hat{\text{ل}}\left(١ - \hat{\text{ل}}\right)\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}$$

$$= \sqrt{٠,٣٢ \times ٠,٦٨\left(\frac{١}{٨٠} + \frac{١}{١٢٠}\right)} = ٠,٠٦٧٣$$

$$\text{أي أن} \quad \text{ي} = \frac{\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢}}{\text{ع}_{(\hat{\text{ل}}_{١} - \hat{\text{ل}}_{٢})}} = \frac{٠,٠٣ - ٠,٣٣}{٠,٠٦٧٣} = -٠,٤٥$$



وبما أن الاختبار هنا ذو ذيلين، وأن منطقتي الرفض تقعان عند كل من الذيل الأيمن والذيل الأيسر ($\text{ي} < -\text{ي}_{٠,٠٢٥}$ ، $\text{ي} > \text{ي}_{٠,٠٢٥}$) حيث $\text{ي}_{٠,٠٢٥} = ٢,٥٨$. فإننا نجد أن قيمة ي المحسوبة (-٠,٤٥) تقع داخل منطقة عدم الرفض. أي أننا لا نرفض فرضية العدم القائلة

بأن نسـبة الوحدات المعيبة في إنتاج كل من المصـنعيْن متساوية وذلك عند مستوى معنوية ٠,٠١.

وفي النهاية، وبنظرة سريعة على الأوجه المختلفة التي طبقنا فيها اختبار الفرضيات وهي:

١- متوسط مجتمع ٢- النسبة في المجتمع

٣- الفرق بين متوسطيْ مجتمعيْن ٤- الفرق بين نسبتيْ مجتمعين

فإنه يمكننا ملاحظة أن مناطق الرفض تأخذ شكلاً واحداً في جميع الحالات وهو:

في حالة الاختبار ذي الذيل الواحد:

- **ي** < -**ي**$_{\alpha}$ إذا وقعت منطقة الرفض عند الذيل الأيسر (أي إذا ظهرت علامة < في الفرضية البديلة).
- **ي** > **ي**$_{\alpha}$ إذا وقعت منطقة الرفض عند الذيل الأيمن (أي إذا ظهرت علامة > في الفرضية البديلة).

في حالة الاختبار ذو الذيليْن:

- **ي** < -**ي**$_{\frac{\alpha}{٢}}$ هي منطقة الرفض عند الذيل الأيسر.
- **ي** > **ي**$_{\frac{\alpha}{٢}}$ هي منطقة الرفض عند الذيل الأيمن.

وأن الاختلاف في الحالات الأربع فيما يتعلق بإجراءات اختبار الفرضيات يتمثل في:

أ- صياغة فرضية العدم والفرضية البديلة.

ب- تعريف المختبِر الإحصائي (ي) أو (ت) حيث يختلف في صياغته من حالة لأخرى.

وهنا، لعله من المفيد أن نشير إلى أن رفض فرضية العدم في الحالات الأربعة المُشار إليها يمكن اتخاذ قرار بشأنه وذلك بشكل أبسط مما سبق وذلك على النحو التالي:

- **في حالة الاختبار ذي الذيل الواحد:**

تُرفض فرضية العدم إذا كانت |**ي**| > **ي**$_{\alpha}$ وذلك بغض النظر عن الصورة التي تأخذها الفرضية البديلة (يلاحظ أن |**ي**| هي القيمة المطلقة للمختبر الإحصائي).

هذا مع ملاحظة أن: |**ي**| > **ي**$_{\alpha}$ تكافئ تماماً أياً من الحالتيْن التاليتيْن أو كليهما معاً:

ي < -**ي**$_{\alpha}$ أو **ي** > **ي**$_{\alpha}$

- **في حالة الاختبار ذي الذيليْن:**

تُرفض فرضية العدم إذا كانت |**ي**| > **ي**$_{\frac{\alpha}{٢}}$.

كما أن: $|\text{ي}| > \text{ي}_{\frac{\alpha}{٢}}$ تكافئ تماماً أياً من الحالتيْن التاليتيْن أو كليهما معاً:

$$\text{ي} < -\text{ي}_{\frac{\alpha}{٢}} \qquad \text{أو} \qquad \text{ي} > \text{ي}_{\frac{\alpha}{٢}}$$

وأما في حالة استخدام توزيع ت بدلاً من التوزيع المعتاد المعياري فإن ما سبق وأن أشرنا إليه ينطبق أيضاً على توزيع ت وذلك باستخدام المختبِر الإحصائي ت بدلاً من المختبِر الإحصائي **ي**. هذا مع ملاحظة أن قيمة كل من ت_{α}، $\text{ت}_{\frac{\alpha}{٢}}$ - لمستوى المعنوية الواحد - تختلف من حالة لأخرى وفقاً لدرجات الحرية. وأما قيمة كل من ي_{α}، $\text{ي}_{\frac{\alpha}{٢}}$ فهي ثابتة في جميع الحالات وذلك لمستوى معنوية محدد.

وفيما يلي نوجز حالات اختبار الفرضيات الخاصة بالوسط الحسابي والنسبة سواء تعلق الأمر بمجتمع واحد وعينه واحدة أو بمجتمعيْن وعينتيْن وذلك كما يوضحه الجدولان (٣ - ٩)، (٣ - ١٠).

وفي نهاية هذا الفصل نأمل أن نكون قد قدمنا للقارئ قدراً ملائماً ومفيداً من موضوع اختبار الفرضيات، مؤكدين أنه لازال هناك الكثير مما يمكن أن يقال في هذا الشأن. ولكننا أكتفينا بهذا القدر تاركين سواه إلى مراحل أخرى بإذن الله.

الجدول (٣ - ٩)

اختيار الفرضيات المتعلقة بالوسط الحسابي والنسبة (باستخدام التوزيع المعتاد: اختبار ي)

م	الحالة	فرضة العدم (ف٠)	الفرضية البديلة (ف١)	المختبِر الإحصائي (إحصاء الاختبار)	شروط رفض فرضية العدم (ف٠)
١	متوسط مجتمع	$م = م_٠$	(١) $م < م_٠$ (٢) $م > م_٠$ (٣) $م \neq م_٠$	$ي = \dfrac{\bar{س} - م_٠}{\dfrac{\sigma}{\sqrt{ن}}}$ أو باستخدام ع كتقريب لـ α	شـــروط الرفض واحـدة في جميع الحـالات الأربع وهي: • في حالة الفرضية البديلة (١): $ي < -ي_{\alpha}$
٢	النسبة في مجتمع	$ل = ل_٠$	(١) $ل < ل_٠$ (٢) $ل > ل_٠$ (٣) $ل \neq ل_٠$	$ي = \dfrac{\hat{ل} - ل_٠}{\sqrt{\dfrac{ل_٠(١ - ل_٠)}{ن}}}$	• في حالة الفرضية البديلة (٢): $ي > ي_{\alpha}$
٣	الفرق بين متوسطيْ مجتمعيْن	$(م_١ - م_٢) = صفر$	(١) $(م_١ - م_٢) < صفر$ (٢) $(م_١ - م_٢) > صفر$ (٣) $(م_١ - م_٢) \neq صفر$	$ي = \dfrac{\bar{س}_١ - \bar{س}_٢}{\sqrt{\dfrac{\sigma_١^٢}{ن_١} + \dfrac{\sigma_٢^٢}{ن_٢}}}$ أو باستخدام ع١، ع٢ كتقريب لـ $\sigma_١$، $\sigma_٢$ على التوالي.	• في حالة الفرضية البديلة (٣): $ي < -ي_{\frac{\alpha}{٢}}$ أو $ي > ي_{\frac{\alpha}{٢}}$ • في حالة الفرضية البديلة (١)، (٢): $\lvert ي \rvert < ي_{\alpha}$
٤	الفرق بين نسبتيْ مجتمعيْن	$(ل_١ - ل_٢) = صفر$	(١) $(ل_١ - ل_٢) < صفر$ (٢) $(ل_١ - ل_٢) > صفر$ (٣) $(ل_١ - ل_٢) \neq صفر$	$ي = \dfrac{\hat{ل}_١ - \hat{ل}_٢}{\sigma_{(\hat{ل}_١ - \hat{ل}_٢)}}$ حيث: $\sigma_{(\hat{ل}_١ - \hat{ل}_٢)} = \sqrt{\hat{ل}(١ - \hat{ل})\left(\dfrac{١}{ن_١} + \dfrac{١}{ن_٢}\right)}$ $\hat{ل} = \dfrac{ن_١\hat{ل}_١ + ن_٢\hat{ل}_٢}{ن_١ + ن_٢}$	• في حالة الفرضية البديلة (٣): $\lvert ي \rvert < ي_{\alpha}$ حيث: $\phi(ي_{\alpha}) = ١ - \alpha$ $\phi(ي_{\frac{\alpha}{٢}}) = ١ - \dfrac{\alpha}{٢}$

الجدول (٣ - ١٠)

اختبار الفرضيات المتعلقة بالوسط الحسابي باستخدام توزيع ت (اختبار ت)

م	الحالة	فرضة العدم (ف.)	الفرضية البديلة (ف١)	المختبِر الإحصائي (إحصاء الاختبار)	شروط رفض فرضية العدم (ف.)
١	متوسط مجتمع	م = م.	(١) م < م. (٢) م > م. (٣) م ≠ م.	$ت = \frac{\bar{س} - م.}{\frac{ع}{\sqrt{ن}}}$	شروط الرفض واحدة الحالتين وهي: • في حالة الفرضية البديلة (١): $ت < -ت_{\alpha}$ • في حالة الفرضية البديلة (٢): $ت > ت_{\alpha}$ • في حالة الفرضية البديلة (٣): $ت < -ت_{\frac{\alpha}{2}}$ أو $ت > ت_{\frac{\alpha}{2}}$ • في حالة الفرضية البديلة (١)، (٢): $\|ت\| > ت_{\alpha}$ • في حالة الفرضية البديلة (٣): $\|ت\| > ت_{\frac{\alpha}{2}}$ حيث تتوقف قيمة $ت_{\alpha}$ ، $ت_{\frac{\alpha}{2}}$ على درجات الحرية (ن – ١) في حالة عينة واحدة وعلى درجات الحرية ($ن_١ + ن_٢ - ٢$) في حالة عينتيْن.
٢	الفرق بين متوسطيْ مجتمعيْن	($م_١ - م_٢$) = صفر	(١) ($م_١ - م_٢$) < صفر (٢) ($م_١ - م_٢$) > صفر (٣) ($م_١ - م_٢$) ≠ صفر	$ت = \frac{\bar{س}_١ - \bar{س}_٢}{\sqrt{ع_م^٢\left(\frac{١}{ن_١} + \frac{١}{ن_٢}\right)}}$ حيث: $ع_م^٢ = \frac{(ن_١ - ١)ع_١^٢ + (ن_٢ - ١)ع_٢^٢}{ن_١ + ن_٢ - ٢}$ وذلك بافتراض تساوي التباينيْن في المجتمعيْن اللذين سُحبت منهما العينتان.	

تمارين

الفصل الثالث

١- ما هو المقصود بكل مما يلي:

أ- منطقة الرفض　　　　**ب-** اختبار ذو ذيل واحد

ج- اختبار ذو ذيلَيْن　　　　**د-** مستوى المعنوية

ه- فرق معنوي (جوهري)

٢- ماهي العوامل المؤثرة على تحديد منطقة الرفض؟

٣- في دراسة إحصائية لمعرفة معدل استهلاك نوع معين من السيارات للوقود، تم اختيار عينة مكونة من ٦٠ سيارة من هذا النوع من السيارات وسُجِّلت المسافة التي تقطعها كل سيارة بالكيلومترات لكل لتر من البنزين. فإذا كانت النتائج كالتالي:
متوسط المسافة التي تقطعها السيارة لكل لتر من البنزين = ١٢كم.
الانحراف المعياري = ٢,٥ كم.
باستخدام مستوى معنوية ٠,٠٥، هل يمكن القول بأنه معدل استهلاك هذا النوع من السيارات يزيد على ١١,٥كم / لتر؟

٤- إذا كان متوسط عدد البيض لسلالة معينة من الدجاج في الشهر هو ٢٠ بيضة للدجاجة الواحدة. فإذا تم اختيار عينة مكونة من ١٠٠ دجاجة حيث وُجِد أن متوسط عدد البيض للعينة ١٩ بيضة والانحراف المعياري ٢,٣ بيضة.
هل يمكن القول بأن هذه العينة تنتمي إلى هذه السلالة من الدجاج؟
استخدم مستوى معنوية ٠,٠١.

٥- لنفترض أن مفتشاً يرغب في إجراء مراجعة سريعة على وزن رغيف الخبز الذي ينتجه أحد المخابز. فأخذ عينة عشوائية من إنتاج هذا المخبز حجمها ٦٤ رغيفاً حيث وجد أن متوسط وزن الرغيف في هذا المخبز هو ١٢٠ جم.
هل تؤكد هذه البيانات صحة ادعاء البعض بأن وزن الرغيف في هذا المخبز يقل عن ١٢٥ جم مستخدماً في ذلك مستوى معنوية ٠,٠٥؟

٦- تتلقى جهة حكومية شكاوى كثيرة من المستهلكين فحواها أن عبوات مسحوق الصابون التي تبيعها إحدى الشركات تحتوي على كمية أقل من ٢٥٠ جم من المسحوق وهي الكمية المعلَن عنها. وللتحقق من صحة شكاوى المستهلكين اشترت الجهة الحكومية ٩ عبوات من المسحوق حيث وُجد أن:

- متوسط وزن المسحوق في العبوة = ٢٤٥ جم.
- الانحراف المعياري = ١٠ جم.

كيف يمكن للجهة الحكومية التحقق من صحة شكاوى المستهلكين عند مستوى معنوية ٠,٠١ وذلك إذا عُلم أن وزن المسحوق في العبوات يتبع توزيعاً معتاداً؟

٧- بافتراض أن نسبة المدخنين بين طلاب إحدى الكليات هي ١٨٪. فإذا أراد باحث أن يختار عينة ممثلة لطلاب هذه الكلية بهدف دراسة ظاهرة التدخين فيها حيث وُجِد أنه من بين ٤٠٠ طالباً شملتهم العينة كان هناك ٨٠ طالباً مدخناً.
هل يمكن - إحصائياً - اعتبار هذه العينة ممثلة لطلاب هذه الكلية فيما يتعلق بهذه الظاهرة؟ استخدم مستوى معنوية ٠,٠٥.

٨- في استطلاع لآراء أولياء الأمور حول الفصل بين الذكور والإناث في المرحلة الابتدائية، أجاب ٢٥٥ من بين ٥٠٠ بالموافقة. هل يمكن القول - وبمستوى معنوية ٠,٠٥ - بأن أغلبية أولياء الأمور يؤيدون هذا الأمر؟

٩- سحبت عينة عشوائية من ٨٠ مُفردة من إنتاج إحدى الآلات فوجد بها ٥ وحدات معيبة. المطلوب اختبار الفرضية القائلة بأن نسبة المعيب في إنتاج هذه الآلة هو ٦,٥٪ في مقابل الفرضية البديلة بأن هذه النسبة لا تساوي ٦,٥٪ وذلك عند مستوى معنوية ٠,٠٥.

١٠- أُلقيت قطعة نقود ٥٠ مرة وذلك لتحديد ما إذا كانت هذه القطعة متزنة أم لا، فظهرت الكتابة ٢٢ مرة. هل يمكن القول بأن قطعة النقود متزنة أم أنها تميل للانحياز للوجه الذي عليه الصورة (استخدم مستوى المعنوية ٠,٠١)؟

١١- فيما يلي أطوال ٧ طلاب ذكور و٩ طالبات إناث بالسنتيمتر تم اختيارهم عشوائياً من بين طلاب إحدى الكليات:

الذكـــور:	١٧٣	١٦٤	١٦٩	١٦٧	١٦٨	١٨٥	١٦٥
الإنـــاث:	١٦٣	١٦٤	١٦٥	١٦٧	١٦٨	١٦٤	١٧١
	١٧٣	١٦٤					

باستخدام مستوى المعنوية ٠,٠٥، هل تؤيد هذه البيانات صحة القول بأن الذكور أكثر طولاً من الإناث؟ وذلك بافتراض:

- أن طلاب هذه الكلية (ذكوراً، إناثاً) يمثلون عينات عشوائية من المجتمع عموماً وذلك فيما يتعلق بظاهرة الطول.
- أن أطوال الذكور والإناث في المجتمع تتبع توزيعاً معتاداً.

١٢- قيس الزمن الذي يستغرقه عاملان فنيان في إصلاح عطل ما في السيارة في ٦٤ حالة. فإذا افترضنا أن قياساتهما تخضع لتوزيع معتاد وأن الوسط الحسابي لقياسات العامل الأول كان ١٥٠ دقيقة وللعامل الثاني ١٤٥ دقيقة وكان التباين في قياساتهما هو ٣٦، ٢٥ على التوالي.
هل يمكن القول بأن كفاءة العامليْن متعادلة في إصلاح ذلك العطل؟
استخدم مستوى معنوية ٠,٠٥.

١٣- أخذت عينتان من مصنعيْن ينتجان المصابيح الكهربائية وذلك لقياس العمر الاستهلاكي لهذه المصابيح (بالساعة) فكانت البيانات كما يلي:

العينة من المصنع الاول (أ) ن١ = ٢٠				**العينة من المصنع الثاني (ب)** ن٢ = ١٥		
٥٠	٩٥	٢١٠	١٢٠	٢١٠	٨٨٠	١٠
٧٥	٩٠	٢٠٠	١٦٠	١٣٠	١٢٥	١٤٠
١٨٥	١١٠	١٠٠	١٣٠	٢٠٥	١٩٠	١٤٥
١٩٥	١٢٠	١١٠	١٥٠	١٦٠	١٨٥	١٢٠
١١٢	٨٠	٩٥	٧٠	١٠٠	٩٥	١١٠

بافتراض أن أعمار المصابيح الكهربائية في كل من المصنعيْن تتبع التوزيع المعتاد.
المطلوب:
أولاً: هل يمكن القول بأن متوسط عمر المصابيح في المصنع (ب) يبلغ ١٢٥ ساعة؟
استخدم α = ٠,٠١
ثانياً: ما مدى صحة ادعاء مدير المصنع (ب) والقائل بأن إنتاج مصنعه يفوق إنتاج المصنع (أ) من حيث كون مصابيحه أطول عمراً؟ استخدام α = ٠,٠٥
ثالثاً: هل يمكن القول بأن أعمار المصابيح في المصنعيْن واحدة؟ استخدام مستوى معنوية ٠,٠١.
رابعاً: ماهي الشروط التي يجب توافرها حتى يمكن استخدام الاختيار الإحصائي الملائم في ثانياً، ثالثاً؟.

١٤- أُعطى اختبار في الإحصاء لعينتيْن من الطلاب في كليتيْن مختلفتيْن، تتكون العينة الأولى من ٦٠ طالباً والثانية من ٩٠ طالباً وكانت النتائج كالتالي:
س̅١ = ٧٥ ، ع١ = ٦ ، س̅٢ = ٧٣ ، ع٢ = ٤

المطلوب:

أولاً: هل يمكن القول بأن مستوى طلاب الكلية الأولى هو أفضل من نظيره في الكلية الثانية؟ استخدم α = ٠,٠٥.

ثانياً: هل هناك اختلاف معنوي بين أداء الطلاب في الكلتيْن؟ استخدم مستوى معنوية: α = ٠,٠١.

١٥- في عينة تحتوي على ١٠٠ قارئ للصحف اليومية في مدينة القاهرة وُجد أن منهم ٦٥ قارئاً يحرصون على قراءة صحيفة الأخبار. وفي عينة أخرى من مدينة الزقازيق حجمها ٨٠ قارئاً وجد أن منهم ٥٠ قارئاً يحرصون على قراءة نفس الصحيفة.
هل يمكن القول بأن صحيفة الأخبار هي أكثر تفضيلاً لدى قراء الصحف في مدينة القاهرة عنها في مدينة الزقازيق؟ استخدم مستوى معنوية ٠,٠٥.

١٦- في دراسة لمعرفة مدى تفضيل الأشخاص البالغين والأشخاص في سن المراهقة لمشاهدة برنامج تليفزيوني معين، أُخذت عينة حجمها ٢٥٠ شخصاً من البالغين، كما أخذت عينة أخرى مكوَّنة من ٤٠٠ شخص في سن المراهقة. فإذا وجدنا أن ٨٠ شخصاً من البالغين و١٢٥ شخصا في سن المراهقة يفضلون مشاهدة هذا البرنامج.
هل يمكن القول بأن درجة تفضيل مشاهدة هذا البرنامج هي واحدة عند كل من الأشخاص البالغين والأشخاص في سن المراهقة؟ استخدم مستوى معنوية ٠,٠٥.

١٧- لنفترض أن لدينا عينتيْن الأولى حجمها ١٢٠ شخصاً والثانية حجمها ١٨٠ شخصاً. وأن عدد المدخنين في العينة الأولى هو ١٢ شخصاً وفي العينة الثانية ٢٥ شخصاً.

المطلوب:

أولاً: اختبر صحة الفرضية القائلة بأن نسبة المدخنين في كل من المجتمعيْن الذيْن سحبت منهما العينتان واحدة. استخدم مستوى معنوية ٠,٠١.

ثانياً: هل يمكن القول بأن ظاهرة التدخين هي أكثر انتشاراً في المجتمع الثاني عنها في المجتمع الأول؟ استخدم معنوية ٠,٠٥.

١٨- أُخذت عينتان الأولى حجمها ٣٠٠ شخص من مدينة القاهرة والثانية حجمها ٢٠٠ شخص من مدينة الإسكندرية. فإذا كان عدد الذين يؤيدون عمل المرأة في عينه القاهرة هو ١٨٠ شخص وفي عينة الاسكندرية ١١٥ شخصاً.
هل يمكن القول بأن نسبة الذين يؤيدون عمل المرأة في المدينتيْن واحدة؟ أم أنها تميل للزيادة في مدينة القاهرة عنها في مدينة الإسكندرية.

١٩- في دراسة لتقييم فعالية نوع جديد من الأسبرين، تم اختيار عينة عشوائية مكونة من ٢٠٠ شخص يعانون من الصداع. فوجد أن ١٦٠ شخصاً منهم قد زالت عنهم آثار الصداع.

ناقش ادعاء المصنع المنتِج لهذا النوع من الأسبرين بأنه يستطيع أن يشفي ٩٠٪ ممن يتعاطونه من الصداع. استخدم مستوى معنوية ٠,٠٥.

٢٠- تنتج آلة منتَجاً معيناً طوله يتبع توزيعاً معتاداً بمتوسط ١٦ سم وانحراف معياري ١,٢ سم. أُختيرت عينة حجمها ١٠٠ من إنتاج هذه الآلة فوجد أن متوسط طول المنتَج فيها هو ١٧,٥ سم.

هل هناك دليل - بمستوى معنوية ٠,٠١ - على أن متوسط هذا المنتَج لم يتغير؟

مقدمة في
الإحصاء التطبيقي

إعداد:
دكتور/ أبوبكر عبد الرحمن
عبد المتعال

الجزء الثاني

2022 - 2023

كلية التجارة
جامعة جنوب الوادي
قسم الأساليب الكمية

مقدمة في

الإحصاء التطبيقي

(الجزء الثاني)

إعداد:

دكتور/ أبوبكر عبد الرحمن عبد المتعال

كلية التجارة – جامعة جنوب الوادي

قسم الأساليب الكمية

العام الجامعي

٢٠٢٢ / ٢٠٢٣

الفصل الرابع
اختبارات مربع كاي
(χ^2 - Tests)

مقدمة:

رأينا في الفصليْن السابقيْن من هذا الكتاب أن استعمال التوزيع المعتاد أو توزيع ت يقتصر على اختبار الفرضيات التي تتعلق بمتغيرات عشوائية متصلة ذات توزيع معتاد أو على الأقل قريب من المعتاد. وفي هذا الفصل نتناول نوعاً آخر من الاختبارات وهو اختبارات مربع كاي Chi – Square Tests. وهذه الاختبارات شائعة الاستخدام في الأبحاث التي تكون نتائجها في شكل تكرارات. حيث تهتم هذه الاختبارات بمدى مطابقة تكرارات مشاهَدة (ملاحَظة) لتكرارات متوقَّعة بناء على نظرية معينة معبَّراً عنها في فرضية العدم.

هذا وفي الوقت الذي ينحصر فيه استخدام التوزيع المعتاد وتوزيع ت على متغيرات كمية متصلة، نجد أن توزيع مربع كاي يتعلق باختبار المتغيرات الوصفية (الصفات) مثل آراء الأشخاص حول قضية من القضايا وعادات التدخين عند الأفراد، إلى غير ذلك من الصفات. هذا بالإضافة إلى أن اختبارات يمكن استخدامها أيضاً في دراسة معنوية العلاقة بين المتغيرات التي يمكن قياسها كمياً وبين بيانات أخرى وصفية كالعلاقة مثلاً بين نسبة المعيب في إنتاج إحدى الآلات وبين وقت العمل (الوردية) من حيث كونه صباحاً أو مساءً. ويتأتَّى ذلك بتقسيم المتغير الكمي أو البيانات الوصفية إلى مجموعات Categories مختلفة وغير متداخلة وذلك بناءً على معايير يراها الباحث. ولعل الأمثلة التالية توضح مجالات استخدام اختبارات مربع كاي.

- معنوية العلاقة بين متغير كمي ومتغير كمي آخر: مثل العلاقة بين درجة الطالب في الاختبار وحجم الأسرة التي يعيش معها. ومرة أخرى نؤكد هنا أن إجراء اختبار مربع كاي على مثل هذا النوع من البيانات إنما يتطلب تقسيم البيانات الكمية إلى مجموعات Categories أو فئات Classes. وهذا ما سوف نراه تفصيلاً فيما بعد.
- معنوية العلاقة ما بين متغير كمي وبيانات وصفية: مثل نسبة المعيب في إنتاج إحدى الآلات ووقت الإنتاج (الوردية).
- معنوية العلاقة ما بين بيانات وصفية وأخرى وصفية أيضاً: مثل الآراء المختلفة تجاه قضية من القضايا وجنس الشخص من حيث كونه ذكراً أم أنثى، وكذلك العلاقة ما بين الإصابة بمرض معين من عدمها وبين التطعيم ضد هذا المرض. أي قياس فعالية حملة التطعيم ضد مرض من الأمراض.

ويعتبر علم الوراثة من أهم المجالات التي يُستخدم فيها اختبار مربع كاي. حيث يتم اختبار صحة بعض قوانين ونظريات الوراثة وذلك من خلال مطابقة التكرارات المشاهَدة عن ظاهرة معينة بالتكرارات المتوقَّعة لهذه الظاهرة والتي تخضع في قياسها لقانون وراثي معين. هذا بالإضافة إلى أن توزيع مربع كاي يستخدم في اختبار ما إذا كانت مجموعة من البيانات تتبع توزيعاً إحصائياً معيناً (التوزيع المعتاد، ثنائي الحدين، بواسون، .... إلخ). وأخيراً فإن توزيع مربع كاي يستخدم أيضاً في مجالات اختبار الفرضيات المتعلقة بالتباين والانحراف المعياري للمجتمع.

هذا وقبل الخوص في مزيد من التفصيلات حول استخدامات اختبارات مربع كاي سوف نتناول باختصار شديد توزيع مربع كاي وأهم خصائصه:

توزيع مربع كاي: Chi – Square Distribution

توزيع مربع كاي هو من التوزيعات المتصلة والتي تُستخدم على نطاق واسع في عملية الاستدلال الإحصائي ويُكتب أحياناً χ^2 - distribution والرمز χ^2 هو حرف يوناني ويُقرأ (كاي). ويعتمد شكل هذا التوزيع على درجات الحرية شأنه في ذلك شأن توزيع ت. أي أن توزيع مربع كاي له معلمة Parameter واحدة هي درجات الحرية Degrees of Freedom.

هذا ولتوزيع مربع كاي الخصائص التالية:

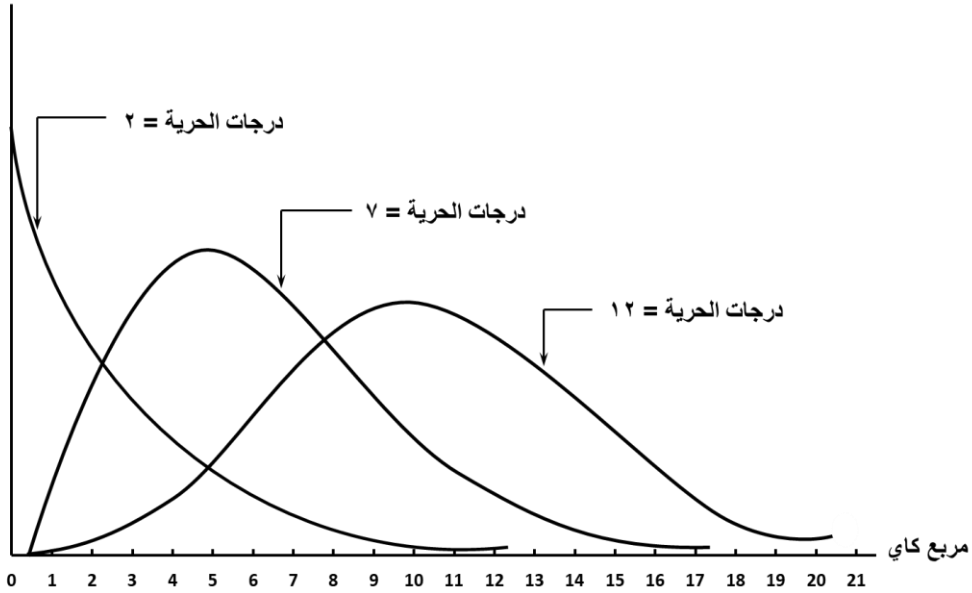
١- أنه يتكون من مربعات قيم المشاهَدات، وبالتالي فإن قيمة مربع كاي دائماً موجبة. أي أن قيمة مربع كاي تتراوح ما بين صفر، ∞.

٢- توزيع مربع كاي هو توزيع ملتو جهة اليمين، أي أنه موجب الالتواء دائماً ومع ذلك فإنه كلما زاد عدد درجات الحرية كلما اقترب التوزيع من التماثل.

٣- أن توزيع مربع كاي له أكثر من شكل حيث يعتمد في ذلك على درجات الحرية (الشكل ٤ - ١). ومن الشكل يتضح أن هذا التوزيع يكون أكثر التواءً جهة اليمين كلما قلت درجات الحرية. أي أنه أكثر قرباً من التماثل بزيادة درجات الحرية.

٤- يعتمد حساب درجات الحرية في اختبارات مربع كاي على نوع الاختبار المُستخدم (سوف نتناول ذلك تفصيلاً فيما بعد) وذلك على عكس اختبار ت والذي تساوي فيه درجات الحرية حجم العينة ناقصاً واحد (ں – ١) وذلك في حالة العينة الواحدة.

٥- أن توزيع مربع كاي هو توزيع متصل وتباينه يتحددان كما يلي:
ت(مربع كاي) = درجات الحرية ،
تباين (مربع كاي) = ٢ × درجات الحرية



الشكل (٤ - ١)
درجات الحرية وشكل توزيع مربع كاي

٦- يعتمد حساب درجات الحرية في اختبارات مربع كاي على نوع الاختبار المستخدَم (سوف نتناول ذلك تفصيلاً فيما بعد) وذلك على عكس اختبار ت والذي تساوي فيه درجات الحرية حجم العينة ناقصاً واحد (ں – ١) وذلك في حالة العينة الواحدة.

مثال (١):

أوجد قيمة مربع كاي عند درجات الحرية ٧ والمساحة على الطرف الأيمن لمنحنى توزيع مربع كاي تساوي ٠,١.

الحل:

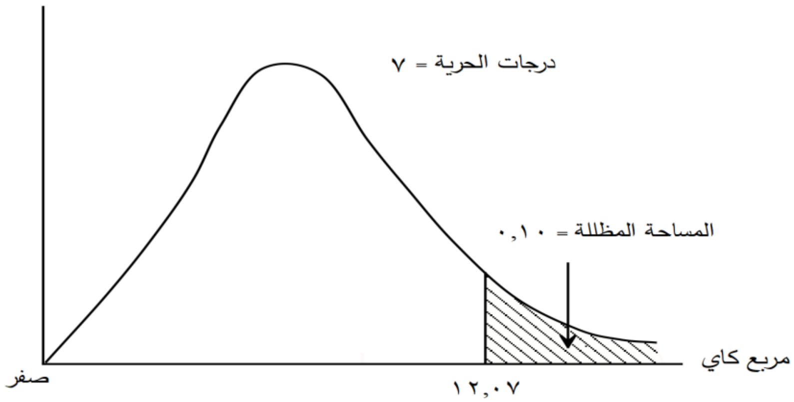
لإيجاد قيمة مربع كاي المطلوبة نبدأ بتحديد الرقم ٧ في العمود المخصص لدرجات الحرية ثم تحديد ٠,١ في الصف العلوي للجدول والمخصص للمساحة على الذيل الأيمن. وقيمة مربع كاي المطلوبة هي القيمة في الجدول عند تقاطع الصف السابع والعمود الممثِّل للقيمة ٠,١ وهذه القيمة تساوي ١٢,٠١٧. والجدول التالي يوضح ذلك، وهو عبارة عن جزء من جدول توزيع مربع كاي المعطى في المحلق (٣) في نهاية الكتاب.

قيمة مربع كاي لدرجات حرية = ٧ ومساحة على الطرف الأيمن = ٠,١

df	Area in the Right Tail Under the Chi-Square Distribution Curve				
	.995	. . .	$\boxed{.100}$	. . .	.005
1	0.000	. . .	2.706	. . .	7.879
2	0.010	. . .	4.605	. . .	10.597
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
$\boxed{7}$	0.989	. . .	$\boxed{12.017}$	. . .	20.278
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
100	67.328	. . .	118.498	. . .	140.169

قيمة مربع كاي المطلوبة

ومن الجدول يمكننا معرفة أن هذه القيمة هي ١٢,٠١٧، وهي القيمة التي يوجد على يسارها مساحة مقدارها ٠,١ تحت منحنى توزيع مربع كاي.



مثال (٢):

أوجد قيمة مربع كاي عندما تكون درجات الحرية ١٢ والمساحة على الطرف الأيسر لمنحنى مربع كاي تساوي ٠,٠٥.

الحل:

بدايةً تلزم الإشارة إلى أن جدول مربع كاي مصمَّم على أساس المساحة على الطرف الأيمن وليس الأيسر. وعليه فإنه إذا كانت المساحة على الطرف الأيسر هي ٠,٠٥ فإن ذلك

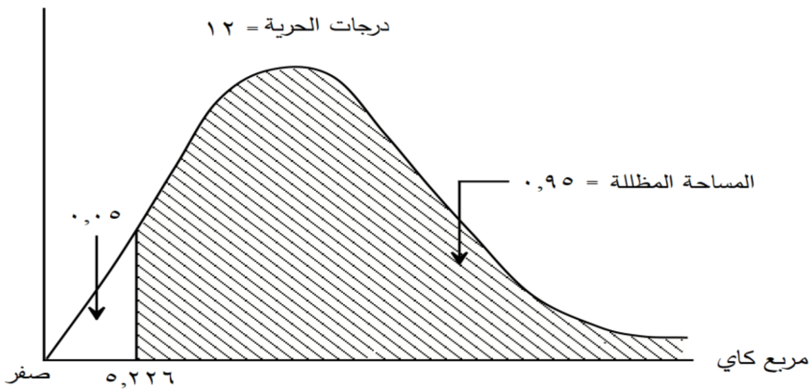
يعني أن المساحة على الطرف الأيمن تساوي ٠,٩٥ وبذلك يمكن إيجاد قيمة مربع كاي المطلوبة بتحديد الصف الذي يحتوي على درجة الحرية ١٢. وفي نقطة التقائه مع العمود الممثل للمساحة ٠,٩٥ تكون هناك القيمة المطلوبة لـ مربع كاي وهي ٥,٢٢٦. والجدول التالي يوضح ذلك، وهو أيضاً يمثل جزءاً من جدول توزيع مربع كاي الموجود في المحلق (٣) في نهاية هذا الكتاب.

قيمة مربع كاي لدرجات حرية = ١٢ ومساحة على الطرف الأيمن = ٠,٩٥

df	Area in the Right Tail Under the Chi-Square Distribution Curve				
	.995	. . .	$\boxed{.950}$	. . .	.005
1	0.000	. . .	0.004	. . .	7.879
2	0.010	. . .	0.103	. . .	10.597
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
$\boxed{12}$	3.074	. . .	$\boxed{5.226}$	. . .	28.300
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
.	. . .	. . .	. . .	. . .	. . .
100	67.328	. . .	77.929	. . .	140.169

قيمة مربع كاي المطلوبة

ويتضح من الشكل التالي أن القيمة ٥,٢٢٦ لـ مربع كاي هي تلك القيمة التي يوجد على يمينها مساحة مقدارها ٠,٩٥ وذلك تحت منحنى توزيع مربع كاي.



اختبارات مربع كاي: Chi – Square Tests

عادة ما يستخدم توزيع مربع كاي في إجراء ثلاثة أنواع مختلفة من الاختبارات نوجزها فيما يلي:

١- اختبار الفرضيات المتعلقة بجداول التوافق Contingency Tables، ويُطلق على مثل هذه الاختبارات اختبارات الاستقلال Independence أو التجانس Homogeneity.

٢- اختبار الفرضيات حول التجارب التي تتكون من أكثر من صفتيْن، ويطلق عليها اختبارات جودة التوفيق Goodness – of – Fit Tests.

٣- اختبار الفرضيات المتعلقة بالتباين والانحراف المعياري لمجتمع واحد.

وفي هذا الباب، سوف تقتصر دراستنا لاختبارات مربع كاي على النوعيْن الأول والثاني من هذه الاختبارات، مُرجِئين النوع الثالث إلى مراحل أخرى في دراستنا لعلم الإحصاء إن شاء الله.

وقبل تناول كلٍّ من هذيْن النوعيْن من الاختبارات بشيء من التفصيل تجدر الإشارة إلى أن اختبارات مربع كاي فيما يتعلق بهذيْن النوعيْن من الاختبارات إنما تتمثل في استخدام توزيع مربع كاي لمعرفة معنوية الفرق بين تكرارات مشاهَدة (ملاحَظة) Observed وبين تكرارات متوقَّعة Expected تعتمد في حساباتها على افتراض صحة نظرية معينة عادة ما تعبِّر عنها فرضية العدم. وبعبارة أخرى، فإن اختبارات مربع كاي لهذيْن النوعيْن من الاختبارات إنما تكشف لنا مدى تطابق التكرارات المشاهَدة مع التكرارات المتوقّعة. هذا وسوف نرمز للتكرار المشاهّد بالرمز ش وللتكرار المتوقَع بالرمز ت.

وقد أشرنا من قبل إلى أن قيمة مربع كاي دائماً موجبة وذلك لأن قيمتها عبارة عن مجموعة قيم مربعة يدخل في حسابها التكرارات المشاهَدة والتكرارات المتوقَعة. وفي واقع الأمر، فإن قيمة مربع كاي تتحدد وفقاً للصيغة التالية:

$$\text{مربع كاي} = \sum \frac{(\text{ش} - \text{ت})^{٢}}{\text{ت}} \qquad (٤ - ١)$$

حيث: ش التكرار المشاهَد ،

ت التكرار المتوقَّع المناظر والذي يعتمد في حسابه على افتراض صحة فرضية العدم.

وهنا نشير إلى كيفية حساب التكرار المتوقَّع تعتمد على نوع الاختبار المستخدم، وهذا ما سوف نراه تفصيلاً فيما بعد.

وبشكل عام، فإن إجراء اختبار مربع كاي لأي من النوعيْن السابق الإشارة إليهما يتطلب الخطوات التالية:

١- صياغة فرضية العدم والفرضية البديلة.

٢- اختيار التوزيع المستخدم في الاختبار، وهو هنا توزيع مربع كاي بطبيعة الحال.

٣- تحديد منطقتي الرفض وعدم الرفض واللتان يفصل بينهما القيمة الحرجة Critical Value لـ مرببع كاي. ويمكن معرفة القيمة الحرجة من جدول المساحة تحت منحنى توزيع مربع كاي (الملحق ٣ في نهاية الكتاب) وذلك بناءً على درجات الحرية ومستوى المعنوية (α) المطلوب. هذا ومستوى المعنوية في اختبارات مربع كاي هنا تعني المساحة على الطرف الأيمن تحت منحنى توزيع مربع كاي. وتسمى قيمة مربع كاي الحرجة بالقيمة الجدولية أيضاً. هذا مع الأخذ في الاعتبار أن تحديد درجات الحرية يختلف في أسلوبه باختلاف نوع الاختبار المستخدَم (كما سنرى فيما بعد).

٤- إيجاد قيمة إحصاء الاختبار كما تحددها الصيغة (٤ - ١) أو أي صيغة أخرى يُشار إليها لاحقاً. وتسمى قيمة إحصاء الاختبار بقيمة مربع كاي المحسوبة.

٥- اتخاذ القرار بشأن رفض أو عدم رفض فرضية العدم. وحيث أن اختبارات مربع كاي هنا هي اختبارات ذات ذيل واحد جهة اليمين، فإن اتخاذ القرار يتمثل في الآتي:
   - رفض فرضية العدم ف. وقبول الفرضية البديلة ف١ إذا كانت:

     قيمة مربع كاي المحسوبة > قيمة مربع كاي الجدولية ،
     أي: قيمة إحصاء الاختبار > القيمة الحرجة لـ مربع كاي.
   - عدم رفض فرضية العدم إذا كانت:
     قيمة مربع كاي المحسوبة < قيمة مربع كاي الجدولية ،
     أي: قيمة إحصاء الاختبار < القيمة الحرجة لـ مربع كاي.

٦- تفسير القرار المتَخذ بما يتلاءم والقضية موضع الاهتمام.

والآن سوف نتناول كل نوع من الاختبارات بشيء من التفصيل على النحو الموضَّح فيما يلي:

أولاً: اختبار الاستقلال أو التجانس

A Test of Independence or Homogeneity

لإجراء مثل هذا النوع من الاختبارات، فإن ذلك يتم من خلال ما يُعرف بـ "جدول التوافق". ولعله من الملائم في البداية أن نعرِّف ما هو المقصود بجدول التوافق.

جدول التوافق: Contingency Table

لكي نوضح ما هو المقصود بـ "جدول التوافق" دعنا نأخذ المثال التالي:
لو افترضنا أننا بصدد معرفة رأي الرجال والنساء تجاه قضية عمل المرأة، وهل تختلف وجهة النظر عند الرجال عنها عند النساء؟ فلو أننا قمنا بسؤال أفراد عينة قوامها ١٠٠ شخص من بينهم ٦٠ رجلاً عن رأيهم تجاه عمل المرأة فكانت النتائج الموضَّحة في الجدول (٤ - ١).

الجدول (٤ - ١)
نموذج لـ "جدول التوافق"

الجنس \ وجهة النظر	مؤيد	معارض	لا رأي له
رجـــل	٣٠	٢٤	٦
امـــرأة	٢٢	١٠	٨

الجدول (٤ - ١) يسمى بـ "جدول التوافق". وهذا الجدول يتكون من صفيْن وثلاثة أعمدة، لذلك يُقال إن حجمه أو رتبته ٢ × ٣. حيث يشير الرقم الأول إلى عدد الصفوف بينما يشير الرقم الثاني إلى عدد الأعمدة. وبذلك يحتوي الجدول على ٦ خلايا، وعددها عبارة عن حاصل ضرب عدد الصفوف في عدد الأعمدة. هذا وعدد صفوف جدول التوافق وكذلك عدد أعمدته إنما يمثلان عدد أقسام أو طبقات Categories كل صفة من الصفتيْن التي يشملهما الجدول. وبصفة عامة، سوف نرمز إلى عدد الصفوف بالرمز ص وإلى عدد الأعمدة بالرمز د. أي أن جدول التوافق يكون بصفة عامة من الرتبة ص × د.

وفي حالتنا هذه، فإن ص = ٢ و د = ٣. وعلى سبيل المثال، فإن الخلية التي تحتوي على الرقم ٦ تعني أن هناك ستة رجال لا رأي لهم تجاه عمل المرأة. وأما الخلية التي تحتوي على العدد ٢٢ فغنها تعني أن هناك ٢٢ امرأة يؤيدن عمل المرأة. وهكذا يمكن تفسير بقية محتويات خلايا جدول التوافق.

والآن عزيزي القارئ، لعله من الملائم أن نتناول كلاً من اختبار الاستقلال واختبار التجانس بشيء من التفصيل.

اختبار الاستقلال: A Test of Independence

اختبار الاستقلال لجدول توافق يعني اختبار فرضية العدم القائلة بأن صفتيْن لعناصر مجتمع واحد لا يوجد بينهما علاقة، أي أنهما مستقلتان. وذلك في مقابل الفرضية البديلة التي تفيد بوجود علاقة بين الصفتيْن.

هذا وعند إجراء اختبار الاستقلال، تجدر الإشارة إلى النقاط التالية:

أ- جدول التوافق يحتوي على صفتيْن لعناصر مجتمع واحد، الأولى يمثلها صفوف الجدول بينمها يمثل الثانية أعمدة هذا الجدول. وتنص فرضية العدم على أن هاتيْن الصفتيْن مستقلتان. وبعبارة أخرى، فإن هاتيْن الصفتيْن لا تؤثر إحداهما على الأخرى. وأما الفرضية البديلة فإنها تفيد بأن هناك علاقة ما تربط ما بين هاتيْن الصفتيْن.

ب- درجات الحرية في اختبار الاستقلال يمكن إيجادها باستخدام الصيغة التالية:

درجات الحرية = (عدد الصفوف – ١) × (عدد الأعمدة – ١)

= (ص – ١)(د – ١)

ج- التكرار المتوقَّع لكل خلية من خلايا الجدول يتحدد كما يلي:

$$\textbf{التكرار المتوقَّع لخلية ما} = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلي}} \qquad \textbf{(٤ - ٢)}$$

والمقصود هنا هو الصف والعمود اللذان تقع فيهما الخلية موضع الاعتبار.

وهنا يلزم التأكيد على أن الصيغة (٤ - ٢) في حساب التكرار المتوقَّع إنما هي نتيجة لافتراض صحة فرضية العدم والتي تفيد باستقلالية الصفتيْن.

هذا ولتطبيق اختبار الاستقلال، فإن حجم العينة يجب أن يكون كبيراً بشكل كاف بحيث لا يقل التكرار المتوقَّع لأي خلية في جدول التوافق عن ٥. وفي الحالات التي يكون فيها التكرار المتوقَّع لخلية من الخلايا أقل من ٥ فإنه يجب حينئذ إما بزيادة حجم العينة أو بإدماج أكثر من قسم أو طبقة Category في قسم واحد أو طبقة واحدة.

د- إن إحصاء الاختبار مربع كاي يمكن إيجاد قيمته باستخدام الصيغة (٤ - ١) والتي سبق الإشارة إليها. وهنا يلزم التنويه إلى أن:

مجموع التكرارات المشاهَدة = مجموع التكرارات المتوقَّعة

أي أن: $\sum \text{ش} = \sum \text{ت}$

وهذه العلاقة صحيحة فيما يتعلق بجميع اختبارات مربع كاي موضع الاعتبار في هذا الكتاب وهي اختبارات الاستقلال والتجانس وجودة التوفيق. ولعل هذا يكون مؤشراً للتأكد من صحة حسابات التكرارات المتوقَّعة لخلايا جدول التوافق.

هـ - إن اختبارات الاستقلال والتجانس هي دائماً اختبارات ذات ذيل واحد جهة اليمين.

وفيما يلي بعض الأمثلة التي توضح كيفية إجراء اختبار مربع كاي للاستقلال.

مثال (٣):

باستخدام بيانات الجدول (٤ - ١)، هل تعطي نتائج العينة دليلاً كافياً للقول بأن وجهة النظر تجاه عمل المرأة لا تتأثر بكون الشخص رجلاً أو امرأة؟ أو بمعنى آخر، هل يمكن القول بأن الصفتيْن "وجهة النظر والجنس" مستقلتان؟ استخدم مستوى المعنوية ٠,٠٥.

الحل:

الصفتان التي يشملهما الجدول (٤ - ١) هما جنس الشخص من حيث كونه رجلاً أو امرأة ووجهة نظر الشخص تجاه عمل المرأة من حيث كونه مؤيداً أو معارضاً أو لا رأي له.

١- صياغة فرضية العدم والفرضية البديلة:

$ف_٠$: آراء الرجال والنساء واحدة تجاه عمل المرأة

$ف_١$: هناك اختلاف في الرأي بين الرجال والنساء فيما يتعلق بعمل المرأة.

٢- اختيار التوزيع الملائم:

التوزيع المستخدم هنا هو توزيع مربع كاي لأن الأمر يتعلق باختبار الاستقلال.

٣- تحديد منطقتي الرفض وعدم الرفض:

في هذا المثال، نجد أن:

ص = ٢ و د = ٣

وعليه فإن: درجات الحرية = (ص – ١)(د – ١)

= (٢ – ١)(٣ – ١)

= ١ × ٢ = ٢

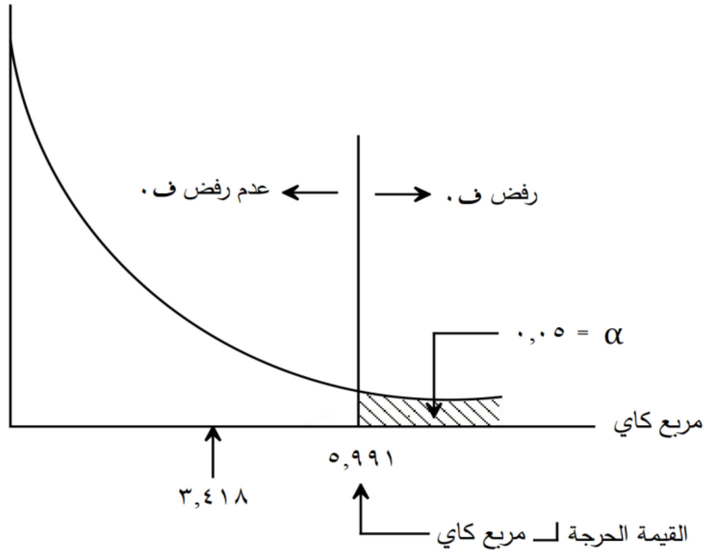
وحيث أن مستوى المعنوية (α) هو ٠,٠٥، فإنه بالكشف في جدول المساحة تحت منحنى توزيع مربع كاي (الملحق في نهاية الكتاب) عن قيمة مربع كاي بدرجات حرية ٢ ومستوى معنوية ٠,٠٥ (أي أن المساحة على الطرف الأيمن تساوي ٠,٠٥) فإننا نجد أن القيمة الحرجة لـ مربع كاي هي:

مربع كاي الجدولية (الحرجة) = ٥,٩٩١.

لذلك فإنه يمكن تحديد منطقتي الرفض وعدم الرفض كما هو موضَّح في الشكل (٤ - ٢).

٤- إيجاد قيمة إحصاء الاختبار:

لإيجاد قيمة إحصاء الاختبار فإنه يلزم حساب التكرار المتوقَّع لكل خلية من خلايا جدول التوافق – وذلك بافتراض صحة فرضية العدم – على النحو التالي:



الشكل (٤ - ٢)

ملحوظة:

في إيجادنا للتكرار المتوقَّع لكل خلية من خلايا جدول التوافق في هذا المثال والأمثلة التالية سوف نبدأ بالصفوف الأول فالثاني وهكذا. وفي مثالنا هذا، على سبيل المثال، سوف نعتبر أن الصف الأول يحتوي على الخلية الأولى والثانية والثالثة. وأما الصف الثاني فإنه يحتوي على الخلايا الرابعة والخامسة والسادسة. وسوف يُتَّبع نفس الأسلوب في الأمثلة القادمة.

الصف الأول:

التكرار المتوقَّع للخلية الأولى = $\frac{٦٠ \times ٥٢}{١٠٠}$ = ٣١,٢

التكرار المتوقَّع للخلية الثانية = $\frac{٦٠ \times ٣٤}{١٠٠}$ = ٢٠,٤

التكرار المتوقَّع للخلية الثالثة = $\frac{٦٠ \times ١٤}{١٠٠}$ = ٨,٤

الصف الثاني:

التكرار المتوقَّع للخلية الرابعة = $\frac{٤٠ \times ٥٢}{١٠٠}$ = ٢٠,٨

التكرار المتوقَّع للخلية الخامسة = $\frac{٤٠ \times ٣٤}{١٠٠}$ = ١٣,٦

التكرار المتوقَّع للخلية السادسة = $\frac{٤٠ \times ١٤}{١٠٠}$ = ٥,٦

وهنا يمكن للقارئ التأكد من صحة حسابات التكرارات المتوقَّعة وذلك بإيجاد مجموع التكرارات المتوقّعة والذي يجب أن يكون مساوياً ١٠٠.

الجدول (٤ - ٢) يوضح التكرار المشاهَد والتكرار المتوقَّع لكل خلية من خلايا الجدول، مع ملاحظة أن التكرار المتوقَّع موضوع بين قوسين وذلك تمييزاً له عن التكرار المشاهَد. وسوف يُتَّبع نفس الأسلوب في الأمثلة التالية:

الجدول (٤ - ٢)

وجهة النظر / الجنس	مؤيد	معارض	لا رأي له	المجموع
رجـال	٣٠ **(٣١,٢)**	٢٤ **(٢٠,٤)**	٦ **(٨,٤)**	٦٠
نســاء	٢٢ **(٢٠,٨)**	١٠ **(١٣,٦)**	٨ **(٥,٦)**	٤٠
المجموع	٥٢	٣٤	١٤	١٠٠

وحيث أن: مربع كاي = $\sum \frac{(ش - ت)^{٢}}{ت}$

$$\therefore \text{مربع كاي} = \frac{(٣٠ - ٣١,٢)^{٢}}{٣١,٢} + \frac{(٢٤ - ٢٠,٤)^{٢}}{٢٠,٤} + \frac{(٦ - ٨,٤)^{٢}}{٨,٤}$$

$$+ \frac{(٢٢ - ٢٠,٨)^{٢}}{٢٠,٨} + \frac{(١٠ - ١٣,٦)^{٢}}{١٣,٦} + \frac{(٨ - ٥,٦)^{٢}}{٥,٦}$$

$$= ٠,٠٤٦ + ٠,٦٣٥ + ٠,٦٨٦ + ٠,٠٦٩ + ٠,٩٥٣ + ١,٠٢٩ = ٣,٤١٨$$

٥- اتخاذ القرار:

حيث أن قيمة إحصاء الاختبار مربع كاي = ٣,٤١٨ أقل من القيمة الحرجة مربع كاي = ٥,٩٩١ وتقع في منطقة عدم الرفض (الشكل: ٤ - ٢) فإننا لا نرفض فرضية العدم.

٦- عدم رفض فرضية العدم يعني بأنه ليس هناك اختلاف بين آراء الرجال وآراء النساء تجاه عمل المرأة. أي أن آراء الرجال والنساء واحدة تجاه هذه القضية.

مثال (٤):

في دراسة للوقوف على أثر طبيعة المنطقة التي يقطن فيها طلاب المدارس الثانوية من حيث كونها ريفية أم حضرية على اتجاههم إلى التدخين. تم اختبار عينة مكوَّنة من ٢٠٠ طالب

يقطنون مناطق حضرية، وعينة أخرى مكوَّنة من ٣٠٠ طالب يقطنون مناطق ريفية. وبسؤالهم عن ما إذا كانوا يدخنون السجائر أم لا، كانت النتائج التالية:

البيان	منطقة حضرية	منطقة ريفية	المجموع
غير مدخن	١٢٠	٢٤٠	٣٦٠
مدخن	٨٠	٦٠	١٤٠
المجموع	٢٠٠	٣٠٠	٥٠٠

هل تعطي بيانات العينة دليلاً كافياً للحكم بأن حالة الشخص من حيث كونه مدخناً أم غير مدخن لا ترتبط بطبيعة المنطقة التي يقطن فيها الطالب من حيث كونها حضرية أم ريفية؟ استخدام مستوى المعنوية ٠,٠١.

الحل:

الصفتان اللتان يشملهما هذا المثال هو طبيعة المنطقة التي يقطن فيها الطلاب من حيث كونها ريفية أو حضرية وحالة الطالب من حيث كونه مدخناً أو غير مدخن.

١- صياغة فرضية العدم والفرضية البديلة:

ف. : ليس هناك علاقة ما بين حالة الطالب (مدخن – غير مدخن) وبين طبيعة المنطقة التي يقطن فيها (حضر – ريف).

ف$_{١}$: اتجاه الطالب إلى التدخين من عدمه يتوقف على كونه يقطن بمنطقة حضرية أم ريفية.

٢- اختيار التوزيع الملائم:

التوزيع المستخدم هنا هو توزيع مربع كاي لأن الأمر يتعلق باختبار الاستقلال.

٣- تحديد منطقتي الرفض وعدم الرفض:

في هذا المثال، نجد أن:

ص = ٢ و د = ٢

وعليه فإن: درجات الحرية = (ص – ١)(د – ١)

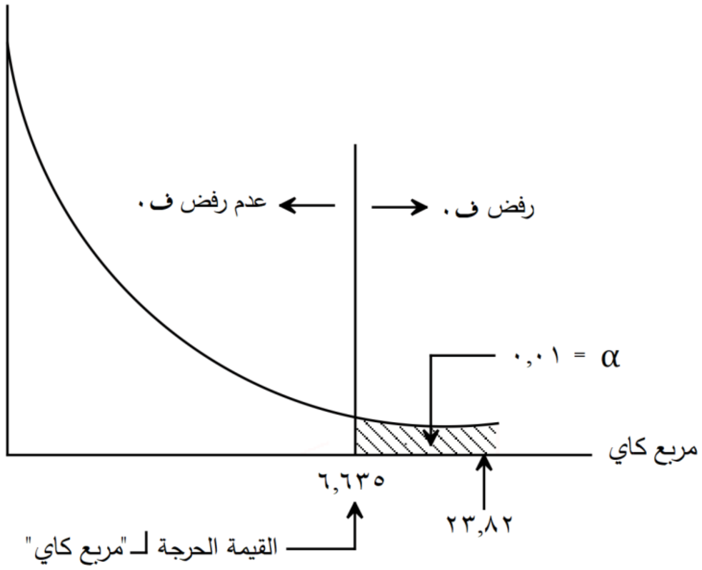
= (٢ – ١)(٢ – ١)

= ١ × ١ = درجة واحدة

وحيث أن مستوى المعنوية α = ٠,٠١، فإنه بالكشف عن القيمة الحرجة لـ مربع كاي في جدول المساحة تحت منحنى توزيع مربع كاي (الملحق ٣ في نهاية الكتاب) نجد أن قيمة مربع

كاي لدرجات حرية واحدة ومستوى معنوية (المساحة على الطرف الأيمن) ٠,٠١ هي مربع كاي = ٦,٦٣٥. أي أن قيمة مربع كاي الجدولية (الحرجة) = ٦,٦٣٥.

وبذلك تتحدد منطقتي الرفض وعدم الرفض كما هو موضح في الشكل (٤ - ٣)



الشكل (٤ - ٣)

٤- إيجاد قيمة إحصاء الاختبار:

لإيجاد قيمة إحصاء الاختبار فإنه يتعين في البداية حساب التكرار المتوقَّع لكل خلية من خلال جدول التوافق بافتراض صحة فرضية العدم.

الصف الأول:

التكرار المتوقَّع للخلية الأولى = $\frac{٣٦٠ \times ٢٠٠}{٥٠٠}$ = ١٤٤

التكرار المتوقَّع للخلية الثانية = $\frac{٣٦٠ \times ٣٠٠}{٥٠٠}$ = ٢١٦

الصف الثاني:

التكرار المتوقَّع للخلية الثالثة = $\frac{١٤٠ \times ٢٠٠}{٥٠٠}$ = ٥٦

التكرار المتوقَّع للخلية الرابعة = $\frac{١٤٠ \times ٣٠٠}{٥٠٠}$ = ٨٤

ويوضح الجدول (٤ - ٣) التكرار المشاهَد والتكرار المتوقَّع لكل خلية من خلايا الجدول (التكرارات المتوقَّعة موضوعة بين قوسينِ).

الجدول (٤ - ٣)

منطقة السكن / حالة الطالب	منطقة ريفية	منطقة حضرية	المجموع
غير مدخـــن	١٢٠ **(١٤٤)**	٢٤٠ **(٢١٦)**	٣٦٠
مدخـــــن	٨٠ **(٥٦)**	٦٠ **(٨٤)**	١٤٠
المجمـــوع	٢٠٠	٣٠٠	٥٠٠

$$\text{مربع كاي} = \frac{(١٢٠ - ١٤٤)^{٢}}{١٤٤} + \frac{(٢٤٠ - ٢١٦)^{٢}}{٢١٦} + \frac{(٨٠ - ٥٦)^{٢}}{٥٦} + \frac{(٦٠ - ٨٤)^{٢}}{٨٤}$$

$$= ٤ + ٢,٦٧ + ١٠,٢٩ + ٦,٨٦ = ٢٣,٨٢$$

٥- اتخاذ القرار:

حيث أن قيمة مربع كاي المحسوبة > قيمة مربع كاي الحرجة (٢٣,٨٢ > ٦,٦٣٥)، فإن ذلك يعني أن قيمة إحصاء الاختبار المحسوبة تقع في منطقة رفض فرضية العدم في الشكل (٤ - ٣). وعليه فإننا نقبل الفرضية البديلة.

٦- تفسير النتيجة: رفض فرضية العدم وقبول الفرضية البديلة يعني أن اتجاه الطلاب إلى التدخين من عدمه إنما يتوقف على كون الطالب يقطن منطقة حضرية أم ريفية. وبعبارة أخرى، يمكننا القول بأن طبيعة المنطقة التي يقطن فيها الطالب تؤثر سلباً أو ايجاباً على اتجاه الطالب للتدخين. أي أن الصفتيْن؛ طبيعة المنطقة (حضر وريف) وحالة الطالب (مدخن وغير مدخن) غير مستقلتيْن.

اختبار التجانس: A Test of Homogeneity

في اختبار التجانس يتم اختبار ما إذا كان مجتمعان (أو أكثر) متجانسيْن، أي متشابهيْن، وذلك فيما يتعلق بتوزيع صفة أو خاصية معينة. وبعبارة أخرى، فإن اختبار التجانس يعني اختبار فرضية العدم التي تفيد بأن نسبة العناصر ذات الصفات المعيَّنة هي واحدة في مجتمعيْن مختلفيْن أو أكثر. وعلى سبيل المثال، يمكننا اختبار فرضية العدم القائلة بأن توزيع الأسر على مستويات الدخول المختلفة (مرتفعة – متوسطة – منخفضة) هو واحد في مدينتيْن مختلفتيْن.

هذا ويتشابه اختبار التجانس مع اختبار الاستقلال في كل شيء، لذلك فإن ما قيل عن اختبار الاستقلال ممثَّلاً في النقاط من (أ) إلى (هـ) ينطبق أيضاً على اختبار التجانس. هذا بالإضافة إلى أن التكرار المتوقَّع لخلية من الخلايا لا يجب أن يقل عن ٥ في حالة اختبار التجانس، شأنه في ذلك شأن اختبار الاستقلال. وفي حالة وجود تكرار متوقَّع أقل من ٥ فإننا نتبع الإجراء المُشار إليه عند تناولنا لاختبار الاستقلال.

وهناك نقطة مهمة تجدر الإشارة إليها وهي أنه في جدول التوافق يمكن للباحث أن يميّز طبيعة القضية التي يتناولها الجدول، هل هي استقلال أم تجانس. فالاستقلال كما سبق أن أشرنا يعني استقلال صفتيْن، أي عدم وجود علاقة بينهما. وأما التجانس فيبحث تجانس مجتمعيْن أو أكثر فيما يتعلق بتوزيع صفة أو خاصية معينة.

وفي واقع الأمر، هناك حالات لجداول التوافق يمكن معها النظر إلى القضية المثارة – من زاوية معينة – على أنها قضية استقلال. وفي نفس الوقت يمكن النظر إليها من زاوية أخرى على أنها قضية تجانس. فعلى سبيل المثال، يمكن النظر إلى طبيعة المنطقة في مثال (٤) على أنها صفة للمنطقة (حضرية وريفية)، وبالتالي موضوع المثال يندرج تحت اختبار الاستقلال. وفي نفس الوقت يمكن النظر إلى طبيعة المنطقة (حضر وريف) على أنها تمثل مجتمعيْن مختلفيْن وأننا معنيين في هذه الحالة بتوزيع الأفراد في كل من هذيْن المجتمعيْن حسب صفة الشخص (مدخن أم غير مدخن). وهنا فإن موضوع المثال يندرج تحت اختبار التجانس. وفي كل الحالات طالما أن الباحث يفسر نتائجه بأسلوب صحيح، فإن ذلك ربما يخفف العبء على القارئ فيما يتعلق بتحديد طبيعة الموضوع المثار أمامه، هل هو استقلال أم تجانس.

والأمثلة التالية توضح إجراءات اختبار التجانس باستخدام جداول التوافق.

مثال (٥):

في دراسة لمعرفة نمط توزيع الدخول في مدينتي سوهاج وقنا، تم أخذ عينة مكونة من ١٠٠ أسرة من مدينة سوهاج وعينة أخرى مكونة من ١٥٠ أسرة من مدينة قنا. وبتوزيع هذه الأسر حسب مستويات الدخول المختلفة وفقاً لمعيار محدد، كانت النتائج التالية:

(البيانات افتراضية)

مستوى الدخل / المدينة	مرتفع	متوسط	منخفض	المجموع
سوهاج	٢٨	٣٢	٤٠	١٠٠
قنا	٣٤	٤٠	٧٦	١٥٠
المجموع	٦٢	٧٢	١١٦	٢٥٠

باستخدام مستوى المعنوية ٠,٠٥، هل يمكن القول بأن نمط توزيع الأسر على مستويات الدخول المختلفة واحد (متشابه) في المدينتيْن؟

الحل:

الصفة موضوع الاعتبار هنا هي مستوى الدخل الذي تتمتع بع الأسرة وأما المجتمعيْن فهما مدينتي سوهاج وقنا.

١- صياغة فرضية العدم والفرضية البديلة:

ف.: توزيع الأسر على مستويات الدخل واحد في المدينتيْن، أي أن نسبة الأسر في كل مستوى من مستويات الدخل متساوية في المدينتيْن.

ف$_١$: توزيع الأسر على مستويات الدخل مختلف في المدينتيْن، أي أن نسبة الأسر في كل مستوى من مستويات الدخل ليست متساوية في المدينتيْن.

٢- اختيار التوزيع الملائم:

حيث أننا بصدد اختبار أن نسبة الأسر في كل مستوى من مستويات الدخل واحدة في مدينتيْ سوهاج وقنا، فإن الاختبار الملائم هنا هو اختبار التجانس.

٣- تحديد منطقتي الرفض وعدم الرفض:

في هذا المثال: ص = ٢ و د = ٣

عدد درجات الحرية = (ص – ١)(د – ١) = (٢ – ١)(٣ – ١) = ١ × ٢ = ٢

وحيث أن مستوى المعنوية المطلوب هو α = ٠,٠٥، فإن القيمة الحرجة لـ مربع كاي هي ٥,٩٩١ وهي القيمة المناظرة لدرجات حرية ٢ ومستوى معنوية (مساحة على الطرف الأيمن) ٠,٠٥، وعليه فإن منطقتي الرفض وعدم الرفض تتحدان كما في الشكل (٤ - ٤).

٤- إيجاد قيمة إحصاء الاختبار:

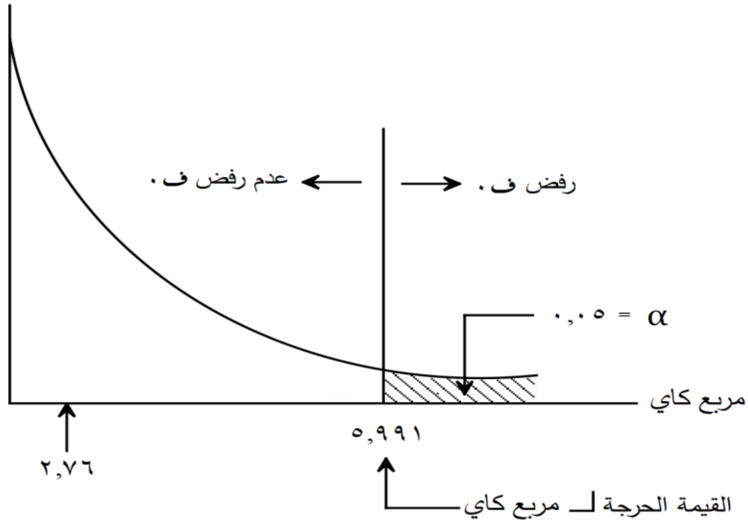
لإيجاد قيمة إحصاء الاختبار مربع كاي فإنه يتعين علينا حساب التكرار المتوقَّع لكل خلية من خلال جدول التوافق وذلك بافتراض صحة فرضية العدم.

الصف الأول:

التكرار المتوقَّع للخلية الأولى = $\frac{١٠٠ \times ٦٢}{٢٥٠}$ = ٢٤,٨

التكرار المتوقَّع للخلية الثانية = $\frac{١٠٠ \times ٧٢}{٢٥٠}$ = ٢٨,٨

التكرار المتوقَّع للخلية الثالثة = $\frac{١٠٠ \times ١١٦}{٢٥٠}$ = ٤٦,٤



الشكل (٤ - ٤)

الصف الثاني:

التكرار المتوقَّع للخلية الرابعة = $\frac{١٥٠ \times ٦٢}{٢٥٠}$ = ٣٧,٢

التكرار المتوقَّع للخلية الخامسة = $\frac{١٥٠ \times ٧٢}{٢٥٠}$ = ٤٣,٢

التكرار المتوقَّع للخلية السادسة = $\frac{١٥٠ \times ١١٦}{٢٥٠}$ = ٦٩,٦

هذا والجدول (٤ - ٤) يوضح التكرار المشاهَد والتكرار المتوقَّع لكل خلية من خلايا جدول التوافق (التكرارات المتوقعة موضوعة بين قوسين).

الجدول (٤ - ٤)

مستوى الدخل / المدينة	مرتفع	متوسط	منخفض	مجموع
غير مدخن	٢٨ (٢٤,٨)	٣٢ (٢٨,٨)	٤٠ (٤٦,٤)	١٠٠
مدخن	٣٤ (٣٧,٢)	٤٠ (٤٣,٢)	٧٦ (٦٩,٦)	١٥٠
المجموع	٦٢	٧٢	١١٦	٢٥٠

$$\text{مربع كاي} = \frac{(٢٨ - ٢٤,٨)^{٢}}{٢٤,٨} + \frac{(٣٢ - ٢٨,٨)^{٢}}{٢٨,٨} + \frac{(٤٠ - ٤٦,٦)^{٢}}{٤٦,٦}$$

$$+ \frac{(٣٤ - ٣٧,٢)^{٢}}{٣٧,٢} + \frac{(٤٠ - ٤٣,٢)^{٢}}{٤٣,٢} + \frac{(٧٦ - ٦٩,٩)^{٢}}{٦٩,٩}$$

$$= ٠,٤١ + ٠,٣٦ + ٠,٨٨ + ٠,٢٨ + ٠,٢٤ + ٠,٥٩ = ٢,٧٦$$

٥- اتخاذ القرار:

حيث أن قيمة المحسوبة لــ مربع كاي (٢,٧٦) أقل من القيمة الحرجة (٥,٩٩١)، وتقع في منطقة عدم الرفض (الشكل: ٤ - ٤)، فإننا لا نرفض فرضية العدم.

٦- عدم رفض فرضية العدم يفيد بأن توزيع الأمر حسب مستويات الدخل واحد في المدينتيْن. أي أنه يمكن القول بأن مجتمعي المدينتيْن لا يختلفان فيما بينهما وذلك فيما يتعلق بمستويات الدخل فيهما.

مثال (٦):

ثلاث آلات (أ ، ب ، **جـ**) تُستخدم في إنتاج منتَج معين. فإذا تم اختيار عينة عشوائية من إنتاج كل من الآلات الثلاثة، وكان حجم العينة ١٥٠، ١٠٠، ٢٥٠ وحدة من إنتاج الآلات الثلاثة أ ، ب ، **جـ** على الترتيب. وبتصنيف إنتاج كل آلة حسب درجة جودة المنتَج كانت النتائج التالية:

الآلـة / **درجة الجودة**	أ	ب	جـ	**المجموع**
أولى	٣٠	١٨	٤٥	٩٣
ثانية	١٠٠	٧٢	١٧٥	٣٤٧
ثالثة	٢٠	١٠	٣٠	٦٠
المجموع	١٥٠	١٠٠	٢٥٠	٥٠٠

هـل يمكن القول بـأن إنتـاج الآلات الثلاثـة على نفس الـدرجـة من الجودة؟ اسـتخـدم مستوى المعنوية ٠,٠١

الحل:

١- صياغة فرضية العدم والفرضية البديلة:

ف$_{٠}$: إنتاج الآلات الثلاثة على نفس الدرجة من الجودة.

ف$_{١}$: درجة جودة المنتَج تختلف من آلة لأخرى، أي أن الآلات تختلف فيما بينهما من حيث جودة منتَجاتها.

٢- اختيار التوزيع الملائم:

يتعلق هذا المثال بتوزيع الإنتاج حسب درجات الجودة وذلك بالنسبة لثلاث آلات مختلفة. أي أن الصفة هنا تتمثل في درجة جودة المنتَج، وأما الآلات أ ، ب ، **جـ** فهي تمثل ثلاثة مجتمعات مختلفة للوحدات المنتَجة. لذلك فإن الاختبار الملائم هنا هو اختبار التجانس، أي تجانس الإنتاج في الآلات الثلاثة. وعليه فإن التوزيع المستخدم هنا هو توزيع مربع كاي والاختبار الملائم هو اختبار مربع كاي للتجانس.

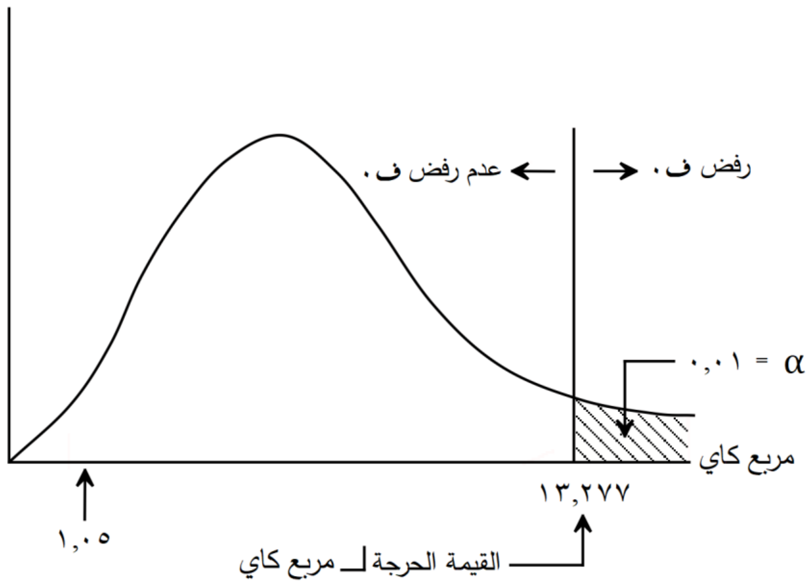
٣- تحديد منطقتي الرفض وعدم الرفض:

عدد درجات الحرية = (ص – ١)(د – ١) ؛ ص = ٣ و د = ٣

= (٣ – ١)(٣ – ١) = ٢ × ٢ = ٤

وحيث أن مستوى المعنوية α = ٠,٠١، فإن القيمة الحرجة لـ مربع كاي يمكن إيجادها من جدول المساحة تحت منحنى توزيع مربع كاي وذلك لدرجات حرية ٤ ومستوى معنوية (مساحة على الطرف الأيمن) ٠,٠١ وبالبحث في الجدول نجد أن: القيمة الحرجة لـ مربع كاي = ١٣,٢٧٧.

هذا ويوضح الشكل (٤ - ٥) منطقتي الرفض وعدم الرفض:



الشكل (٤ - ٥)

٤- إيجاد قيمة إحصاء الاختبار:

بافتراض صحة فرضية العدم، فإن التكرار المتوقَّع لكل خلية من خلايا الجدول يمكن حسابه على النحو التالي:

الصف الأول:

التكرار المتوقَّع للخلية الأولى = $\frac{٩٣ \times ١٥٠}{٥٠٠} = ٢٧,٩$

التكرار المتوقَّع للخلية الثانية = $\frac{٩٣ \times ١٠٠}{٥٠٠} = ١٨,٦$

التكرار المتوقَّع للخلية الثالثة = $\frac{٩٣ \times ٢٥٠}{٥٠٠} = ٤٦,٥$

الصف الثاني:

التكرار المتوقَّع للخلية الرابعة = $\frac{٣٤٧ \times ١٥٠}{٥٠٠} = ١٠٤,١$

التكرار المتوقَّع للخلية الخامسة = $\frac{٣٤٧ \times ١٠٠}{٥٠٠} = ٦٩,٤$

التكرار المتوقَّع للخلية السادسة = $\frac{٣٤٧ \times ٢٥٠}{٥٠٠} = ١٧٣,٥$

الصف الثالث:

التكرار المتوقَّع للخلية السابعة = $\frac{٦٠ \times ١٥٠}{٥٠٠} = ١٨$

التكرار المتوقَّع للخلية الثامنة = $\frac{٦٠ \times ١٠٠}{٥٠٠} = ١٢$

التكرار المتوقَّع للخلية التاسعة = $\frac{٦٠ \times ٢٥٠}{٥٠٠} = ٣٠$

ويوضح **الجدول (٤ - ٥)** التكرار المشاهَد والتكرار المتوقَّع لكل خلية من خلايا جدول التوافق.

مربع كاي = $\frac{(٣٠ - ٢٧,٩)^٢}{٢٧,٩} + \frac{(١٨ - ١٨,٦)^٢}{١٨,٦} + \frac{(٤٥ - ٤٦,٥)^٢}{٤٦,٥}$

$+ \frac{(١٠٠ - ١٠٤,١)^٢}{١٠٤,١} + \frac{(٧٢ - ٦٩,٤)^٢}{٦٩,٤} + \frac{(١٧٥ - ١٧٣,٥)^٢}{١٧٣,٥}$

$+ \frac{(٢٠ - ١٨)^٢}{١٨} + \frac{(١٠ - ١٢)^٢}{١٢} + \frac{(٣٠ - ٣٠)^٢}{٣٠}$

= ٠,١٦ + ٠,٠٢ + ٠,٠٥ + ٠,١٦ + ٠,١٠ + ٠,٠١ + ٠,٢٢

+ ٠,٣٣ + صفر = ١,٠٥

الجدول (٤ - ٤)

الآلة / درجة الجودة	أ	ب	جـ	المجموع
أولـى	٣٠ **(٢٧,٩)**	١٨ **(١٨,٦)**	٤٥ **(٤٦,٥)**	٩٣
ثانيــة	١٠٠ **(١٠٤,١)**	٧٢ **(٦٩,٤)**	١٧٥ **(١٧٣,٥)**	٣٤٧
ثالثــة	٢٠ **(١٨)**	١٠ **(١٢)**	٣٠ **(٣٠)**	٦٠
المجموع	١٥٠	١٠٠	٢٥٠	٥٠٠

٥- اتخاذ القرار:

حيث أن قيمة مربع كاي المحسوبة (١,٠٥) أقل من القيمة الحرجة لـ مربع كاي وهي ١٣,٢٧٧، وبالتالي فإنها تقع في منطقة عدم الرفض (الشكل: ٤ - ٥). أي أن بيانات العينة لا تعطي دليلاّ كافياً لرفض فرضية العدم.

٦- عدم رفض فرضية العدم يفيد بأن إنتاج الآلات الثلاثة على نفس الدرجة من الجودة. وبعبارة أخرى، توزيع الإنتاج على درجات الجودة المختلفة واحد للآلتيْن.

وفي نهاية تناولنا لاختبارات الاستقلال والتجانس، تجدر الإشارة إلى ما يلي:

أ- أنه يمكن استخدام صيغ أخرى لحساب قيمة مربع كاي، وهذه الصيغ أكثر سهولة ويسرا – من الناحية الحسابية – من الصورة (٤ - ١).

➡ في حالة جدول التوافق ٢ × ٢:

إذا افترضنا أن جدول التوافق ٢ × ٢ يأخذ الصورة:

أ	ب	أ + ب
جـ	د	**جـ** + د
أ + **جـ**	ب + د	أ + ب + **جـ** + د

فإن مربع كاي يمكن حسابها كما يلي:

$$\textbf{مربع كاي} = \frac{(\text{أد} - \text{بجـ})^{٢}\,(\text{أ} + \text{ب} + \text{جـ} + \text{د})}{(\text{أ} + \text{ب})(\text{جـ} + \text{د})(\text{أ} + \text{جـ})(\text{ب} + \text{د})} \qquad \textbf{(٤ - ٣)}$$

ففي مثال (٤)، على سبيل المثال، نجد أن:

$$\text{مربع كاي} = \frac{\left(١٢٠ \times ٦٠ - ٢٤٠ \times ٨٠\right)^{٢} \left(٥٠٠\right)}{\left(٣٦٠\right)\left(١٤٠\right)\left(٢٠٠\right)\left(٣٠٠\right)} = ٢٣,٨$$

وهي نفس النتيجة التي حصلنا عليها باستخدام الصيغة (٤ – ١). ومما لا شك فيه أن استخدام الصيغة (٤ - ٣) في حساب مربع كاي كان أيسر – حسابياً – من الصيغة (٤ - ١).

- في حالة جدول التوافق الذي يزيد فيه عدد الصفوف أو عدد الأعمدة أو كليهما عن ٢:

حيث أن: $\text{مربع كاي} = \sum \frac{\left(\text{ش} - \text{ت}\right)^{٢}}{\text{ت}}$

$$= \sum \frac{\left(\text{ش}^{٢} - ٢\text{ش ت} + \text{ت}^{٢}\right)}{\text{ت}}$$

$$= \sum \frac{\text{ش}^{٢}}{\text{ت}} - ٢\sum \frac{\text{ش ت}}{\text{ت}} + \sum \frac{\text{ت}^{٢}}{\text{ت}}$$

$$= \sum \frac{\text{ش}^{٢}}{\text{ت}} - ٢\sum \text{ش} + \sum \text{ت}$$

وحيث أن:

$$\sum \text{ش} = \sum \text{ت} = \text{ن}$$

فإن: $\text{مربع كاي} = \sum \frac{\text{ش}^{٢}}{\text{ت}} - ٢\sum \text{ش} + \sum \text{ش}$

$$= \sum \frac{\text{ش}^{٢}}{\text{ت}} - \sum \text{ش}$$

$$= \sum \frac{\text{ش}^{٢}}{\text{ت}} - \text{ن}$$

أي أن:

$$\textbf{مربع كاي} = \sum \frac{\text{ش}^{٢}}{\text{ت}} - \text{ن} \qquad \textbf{(٤ - ٤)}$$

حيث: **ش** التكرار المشاهَد ، ت التكرار المتوقَّع ، ن مجموع التكرارات.

وفي مثال (٥) على سبيل المثال:

مربع كاي = $\left(\frac{(٢٨)^{٢}}{٢٤,٨} + \frac{(٣٢)^{٢}}{٢٨,٨} + \frac{(٤٠)^{٢}}{٤٦,٤} + \frac{(٣٤)^{٢}}{٣٧,٢} + \frac{(٤٠)^{٢}}{٤٣,٢} + \frac{(٧٦)^{٢}}{٦٩,٦} \right)$ – ٢٥٠

= (٣١,٦ + ٣٥,٥٦ + ٣٤,٤٨ + ٣١,٠٨ + ٣٧,٠٤ + ٨٢,٩٩) – ٢٥٠

= ٢٥٢,٧٦ – ٢٥٠ = ٢,٧٦

وهي نفس النتيجة التي حصلنا عليها باستخدام الصيغة (٤ - ١). ومما لا شك فيه، فإن استخدام الصيغة (٤ - ٤) في إيجاد قيمة إحصاء الاختبار مربع كاي أيسر بكثير من استخدام الصيغة (٤ - ١). وتجدر الإشارة هنا إلى أن الصيغة (٤ - ٤) يمكن استخدامها أيضاً في حالة جدول التوافق ٢ × ٢.

ب- أنه عند استخدام اختبارات مربع كاي للاستقلال والتجانس فإن جدول التوافق يجب أن تحتوي خلاياه على تكرارات وليس نسب.

ج- أنه فيما يتعلق بالحد الأدنى الذي لا يجب أن يقل عنه التكرار المتوقَّع لأي خلية من خلايا جدول التوافق، فإن هناك الكثير من الآراء حول هذا الأمر. ولكننا اكتفينا بالرأي القائل بأن الحد الأدنى هو ٥ دون الدخول في كثير من التعقيدات وهو ما لا تحتمله هذه المرحلة من الدراسة.

ثانياً: اختبار جودة التوفيق A Goodness – of – Fit Test

يُستخدم اختبار مربع كاي أيضاً في إجراء الاختبارات حول التجارب التي لها أكثر من نتيجة Outcome والتي يُطلق عليها التجارب متعددة الحدود Multinomial وتتميز هذه التجارب بالخصائص التالية:

١- تحتوي التجربة على ن محاولة.

٢- كل محاولة ينتج عنها واحدة من النتائج الممكنة وعددها ك حيث ك > ٢.

٣- المحاولات مستقلة عن بعضها البعض.

٤- احتمال كل نتيجة من نتائج التجربة ثابت لكل محاولة.

ومن أمثلة التجارب متعددة الحدود رمي زهرة نرد عدة مرات، حيث أن هناك ٦ نتائج ممكنة وهي ظهور أي من الأرقام من ١ إلى ٦، وأن كل رمية مستقلة عن الرميات الأخرى. كما أن احتمالات النتائج الستة تظل ثابتة في كل رمية.

وفي هذا المثال، يمكن اختبار ما إذا كانت زهرة النرد متزنة (غير متحيزة) Fair. ويتم ذلك بإجراء اختبار مربع كاي على التكرارات المشاهَدة للأرقام من ١ إلى ٦ والتكرارات المتوقَّعة

لها. هذا والتكرار المتوقَّع لكل رقم يساوي – نظرياً – $\frac{١}{٦}$ إجمالي المحاولات (يُلاحظ أن احتمال ظهور أي رقم من الأرقام الستة عند رمي زهرة نرد هو $\frac{١}{٦}$). وبعبارة أخرى، فإن اختبار مربع كاي يبحث درجة تطابق التكرارات المشاهَدة للأرقام الستة مع التكرارات النظرية المناظرة والتي يتم إيجادها على أساس أن احتمال ظهور أي رقم عند رمي زهرة النرد هو دائماً $\frac{١}{٦}$ في كل رمية.

وفي اختبار جودة التوفيق، يتم اختبار فرضية العدم القائلة بأن التكرارات المشاهّدة لتجربة معينة تتوافق (تتطابق) مع التكرارات المتوقَّعة المناظرة لها وذلك في مقابل الفرضية البديلة التي تفيد بأن التكرارات المشاهَدة تختلف عن التكرارات المتوقَّعة.
هذا وفي اختبار جودة التوفيق تجدر الإشارة إلى النقاط التالية:

١- التكرار المتوقَّع (ت) لكل صفة أو طبقة Category يمكن الحصول عليه كما يلي:

$$\text{ت} = \text{ن}\,\text{ع} \qquad (٤ - ٥)$$

حيث: ن حجم العينة،
ع احتمال انتماء العنصر إلى الصفة أو الطبقة بافتراض صحة فرضية العدم.

٢- درجات الحرية هي عبارة عن عدد النتائج (الصفات) الممكنة للتجربة ناقصاً ١، أي أن:

درجات الحرية = ك – ١

حيث **ك** تمثل عدد النتائج (الصفات) الممكنة للتجربة.

٣- إن إحصاء الاختبار مربع كاي يُحسب باستخدام الصيغة (٤ - ١) أيضاً والمشار إليها في حالة اختبار الاستقلال والتجانس، كما يمكن حسابه أيضاً باستخدام الصيغة (٤ - ٤). أي أن:

$$\text{مربع كاي} = \sum \frac{(\text{ش} - \text{ت})^{٢}}{\text{ت}}$$

حيث: **ش** التكرار المشاهَد،
ت التكرار المتوقَّع المناظر والذي يعتمد في حسابه على افتراض صحة فرضية العدم؛ ت = ن ل

٤- أن اختبار مربع كاي للتجانس هو دائماً اختبار ذو ذيل واحد جهة اليمين.

وفيما يلي بعض الأمثلة التي توضح كيفية إجراء اختبار مربع كاي للتجانس.

مثال (٧):

في تجربة لإلقاء زهرة نرد على سطح أملس ٦٠ مرة، كانت النتائج التالية فيما يتعلق بعدد مرات ظهور كل رقم من الأرقام الستة:

الرقم	١	٢	٣	٤	٥	٦	المجموع
عدد مرات الظهور	٧	١١	٨	١٣	١٠	١١	٦٠

على ضوء نتائج هذه التجربة، هل يمكن القول بأن زهرة النرد هذه متزنة (غير متحيزة)؟ استخدم مستوى المعنوية ٠,٠١.

الحل:

١- صياغة فرضية العدم والفرضية البديلة:

$ف_٠$: الزهرة متزنة (غير متحيزة).

$ف_١$: الزهرة غير متزنة (متحيزة).

٢- اختيار التوزيع الملائم:

بما إننا هنا بصدد مطابقة تكرارات مشاهَدة بتكرارات متوقَّعة، فإن الاختبار الملائم هنا هو اختبار مربع كاي لجودة التوفيق.

٣- تحديد منطقتي الرفض وعدم الرفض:

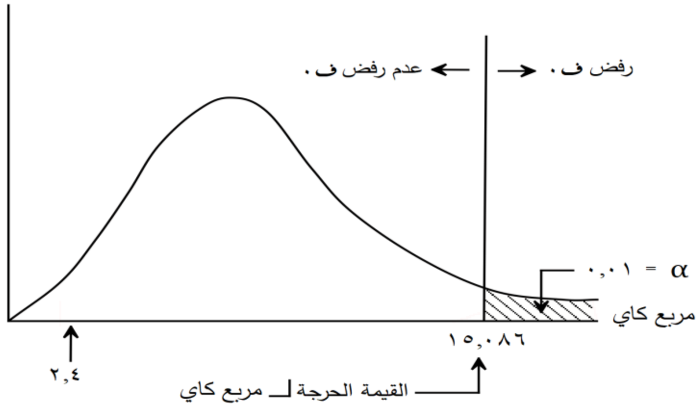
في هذا المثال، نجد أن: ك = ٦

وهي تمثل عدد النتائج الممكنة عند رمي زهرة نرد.

أي أن درجات الحرية هي: ك – ١ = ٦ – ١ = ٥

وحيث أن مستوى المعنوية المطلوب هو ٠,٠١، فإنه بالكشف عن قيمة مربع كاي الحرجة في جدول المساحة تحت منحنى توزيع مربع كاي عند درجات حرية ٥ ومساحة على الطرف الأيمن تساوي ٠,٠١ نجد أن: القيمة الحرجة لـ "مربع كاي" = ١٥,٠٨٦

وبالتالي يمكن تحديد منطقتي الرفض وعدم الرفض كما يتضح من الشكل (٤ - ٦).



الشكل (٤ - ٦)

٤- إيجاد قيمة إحصاء الاختبار:

لإيجاد قيمة إحصاء الاختبار فإنه يلزم حساب التكرار المتوقَّع لمرات ظهور كل رقم من الأرقام الستة، آخذين في الاعتبار أن احتمال ظهور أي رقم هو $\frac{١}{٦}$ دائماً في كل رمية لزهرة النرد (هذا لافتراضنا صحة فرضية العدم والتي تقول بأن الزهرة متزنة). والجدول (٤ - ٦) يوضح الحسابات اللازمة لذلك.

الجدول (٤ - ٦)

الرقم	التكرار المشاهَد (ش)	الاحتمال (ع)	التكرار المتوقَّع ت = ن ع	ش – ت	$\frac{(ش - ت)^{٢}}{ت}$
١	٧	$\frac{١}{٦}$	٦٠($\frac{١}{٦}$) = ١٠	-٣	٠,٩
٢	١١	$\frac{١}{٦}$	٦٠($\frac{١}{٦}$) = ١٠	١	٠,١
٣	٨	$\frac{١}{٦}$	٦٠($\frac{١}{٦}$) = ١٠	-٢	٠,٤
٤	١٣	$\frac{١}{٦}$	٦٠($\frac{١}{٦}$) = ١٠	٣	٠,٩
٥	١٠	$\frac{١}{٦}$	٦٠($\frac{١}{٦}$) = ١٠	صفر	صفر
٦	١١	$\frac{١}{٦}$	٦٠($\frac{١}{٦}$) = ١٠	١	٠,١
المجموع	٦٠	١	٦٠	صفر	٢,٤

من الجدول (٤ - ٦) يتضح أن قيمة مربع كاي المحسوبة (إحصاء الاختبار) هي ٢,٤.

٥- اتخاذ القرار:

حيث أن قيمة إحصاء الاختبار المحسوبة (٢,٤) أقل من القيمة الحرجة لـ مربع كاي (١٥,٠٨٦)، فإن ذلك يعني أن قيمة إحصاء الاختبار تقع في منطقة عدم الرفض الشكل (٤ - ٦). ويكون القرار هو عدم رفض فرضية العدم.

٦- عدم رفض فرضية العدم يعني أنه يمكن القول إن هذه الزهرة متزنة، أي غير متحيزة لرقم على حساب الآخر.

مثال (٨):

يدَّعي أحد علماء علم النبات بأن نباتاً من نوع معين يعطي أزهاراً من اللون الأبيض والأصفر والأسود بنسب ٥ : ٢ : ٣. فإذا تم زراعة ٢٠٠ نبات من هذا النوع، وكانت المشاهدات التالية:

اللون	أبيض	أصفر	أسود	المجموع
عدد الأزهار	٧٠	٥٠	٨٠	٢٠٠

هل تعطي هذه النتائج دليلاً كافياً للقول بأن نظرية هذا العالم صحيحة؟ استخدم مستوى المعنوية ٠,٠٥.

الحل:

١- صياغة فرضية العدم والفرضية البديلة:

$ف_{.}$: النظرية صحيحة.

$ف_{١}$: النظرية غير صحيحة.

٢- اختيار التوزيع الملائم:

نظراً لأن الإجابة على التساؤل المطلوب تتطلب معرفة التكرارات المتوقَّعة لكل لون وفقاً لنظرية هذا العالم، وهل هذه التكرارات تتوافق (تتطابق) مع التكرارات المشاهَدة؟ لذلك فإن التوزيع الملائم هنا هو توزيع مربع كاي، وأما الاختبار المناسب فهو اختبار جودة التوفيق.

٣- تحديد منطقتي الرفض وعدم الرفض:

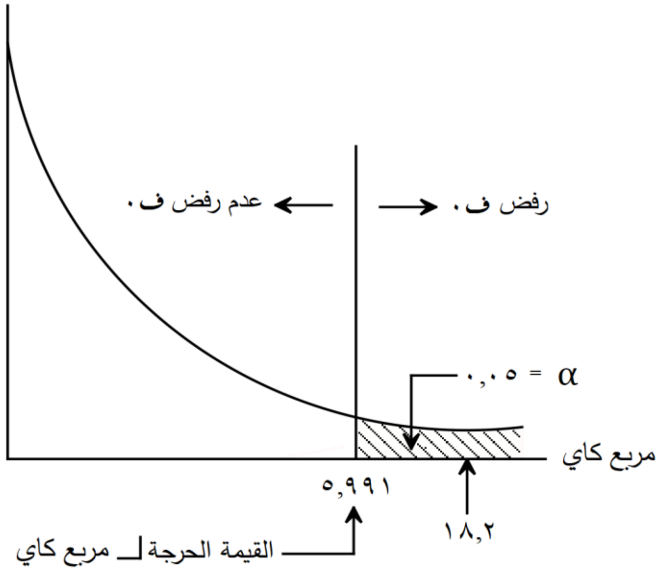
في هذا المثال، نجد أن: ك = ٣

وهي تمثل عدد الألوان الممكنة لزهور هذا النبات عند زراعته.

أي أن: درجات الحرية = ك – ١ = ٣ – ١ = ٢

وحيث أن مستوى المعنوية المطلوب هو α = ٠,٠٥، فإنه القيمة الحرجة لـ مربع كاي يمكن تحديدها من جدول المساحة تحت منحنى توزيع مربع كاي عند درجات حرية ٢ ومستوى معنوية (مساحة على الطرف الأيمن) ٠,٠٥ ومن الجدول نجد أن: القيمة الحرجة لـ مربع كاي تساوي ٥,٩٩١.

لذلك فإن منطقتي الرفض وعدم الرفض يمكن تحديدها كما في الشكل (٤ - ٧).



الشكل (٤ - ٧)

٤- إيجاد قيمة إحصاء الاختبار:

بافتراض صحة فرضية العدم، فإن احتمالات الحصول على الألوان الأبيض والأصفر والأسود تكون كما يلي:

$$\text{احتمال (اللون الأبيض)} = \frac{٥}{٥ + ٢ + ٣} = \frac{٥}{١٠} = ٠,٥$$

$$\text{احتمال (اللون الأصفر)} = \frac{٢}{٥ + ٢ + ٣} = \frac{٢}{١٠} = ٠,٢$$

$$\text{احتمال (اللون الأسود)} = \frac{٣}{٥ + ٢ + ٣} = \frac{٣}{١٠} = ٠,٣$$

لذلك فإن التكرار المتوقَّع لكل لون يمكن حسابه كما هو موضَّح في الجدول (٤ - ٧).

الجدول (٤ - ٧)

اللون	التكرار المشاهَد (ش)	الاحتمال (ع)	التكرار المتوقَّع (ت)	$\frac{(ش - ت)^{٢}}{ت}$
أبيض	٧٠	٠,٥	٢٠٠(٠,٥) = ١٠٠	٩,٠
أصفر	٥٠	٠,٢	٢٠٠(٠,٢) = ٤٠	٢,٥
أسود	٨٠	٠,٣	٢٠٠(٠,٣) = ٦٠	٦,٧
المجمــوع	٢٠٠	١,٠	٢٠٠	١٨,٢

من الجدول (٤ - ٧) نجد أن: قيمة إحصاء الاختبار = ١٨,٢.

٥- اتخاذ القرار:

حيث أن قيمة إحصاء الاختبار (١٨,٢) أكبر من القيمة الحرجة لـ مربع كاي فإنها تقع في منطقة الرفض (الشكل ٤ - ٧). لذلك نرفض فرضية العدم ونقبل الفرضية البديلة.

٦- رفض فرضية العدم وقبول الفرضية البديلة يفيد بأن نتائج التجربة لا تعطي دليلاً كافياً للتسليم بصحة نظرية عالم النبات.

وســوف نتوقف عند هذا الحد في دراســتنا لاختبارات مربع كاي للاســتقلال والتجانس وجودة التوفيق.

تمارين

الفصل الرابع

١- في دراسة لمعرفة آراء الرجال والنساء تجاه أحد القضايا، ثم أُخذت عينة حجمها ١٠٠ رجل و١٥٠ امرأة من أحد المجتمعات وكان آراؤهم كما يلي:

وجهة النظر / الجنس	موافق	غير موافق	لا رأي له	المجموع
رجل	٦٠	٣٠	١٠	١٠٠
امرأة	٩٠	٤٠	٢٠	١٥٠
المجموع	١٥٠	٧٠	٣٠	٢٥٠

هل يمكن القول بأنه في هذا المجتمع تختلف آراء الرجال عن آراء النساء تجاه هذه القضية؟ استخدم مستوى المعنوية ٠,٠١.

٢- تم أخذ عينة حجمها ٥٠٠ مولوداً من أحد مستشفيات الولادة وذلك خلال عام، حيث وُجد أن من بينهم ٢٨٠ ذكراً و٢٢٠ أنثى. هل تدعَّم بيانات العينة الرأي القائل بأن المواليد الإناث يمثلن ٥٠٪ من إجمالي المواليد في المجتمع ككل؟ استخدم مستوى المعنوية ٠,٠٥.

٣- أُلقيت زهرة نرد على سطح أملس ١٥٠ مرة فكانت النتائج التالية فيما يتعلق بعدد مرات ظهور كل رقم:

الرقم على الوجه	١	٢	٣	٤	٥	٦
عدد مرات الظهور	٢٥	٢٣	٢٦	٣١	٢٥	٢٠

على ضوء نتائج التجربة، هل ترى أن زهرة النرد متزنة (غير متحيزة)؟ استخدم مستوى المعنوية ٠,٠١.

٤- لدراسة العلاقة ما بين وقت العمل وإنتاجية العامل (معبَّراً عنها بعدد الوحدات المنتَجة) خلال فترة العمل، أُخذت عينة عشوائية مكونة من ٥٠٠ عامل من أحد المصانع الكبيرة فكانت النتائج التالية:

وقت العمل / الإنتاجية	صباحاً	ظهراً	مساءً	المجموع
٢٠ ــ ٢٤	٣٠	١٠٠	٧٠	٢٠٠
٢٥ ــ ٢٩	٤٠	٢٥	٤٥	١١٠
٣٠ ــ ٣٤	٥٠	٧٥	٦٥	١٩٠
المجموع	١٢٠	٢٠٠	١٨٠	٥٠٠

هل تعطي بيانات العينة دليلاً كافياً للقول بأن إنتاجية العامل في هذا المصنع تتأثر بوقت العمل؟ استخدم مستوى المعنوية ٠,٠٥.

٥- تم إلقاء قطعة عملة ١٠٠ مرة حيث ظهر الوجه الذي عليه الشعار ٦٠ مرة، هل يمكن القول بأن قطعة العملة متزنة (غير متحيزة)؟ استخدم مستوى المعنوية ٠,٠١.

٦- تم أخذ عينة حجمها ٢٠٠ طالباً من طلاب المدارس الثانوية في المناطق الحضرية، وعينة أخرى حجمها ٣٠٠ طالباً من طلال المدارس الثانوية في المناطق الريفية. وبسؤال كل طالب هل يدخن السجائر أم لا، كانت النتائج التالية:

البيان	مناطق حضرية	مناطق ريفية	المجموع
لم يدخن أبداً	١٢٠	٢٠٠	٣٢٠
يدخن	٨٠	١٠٠	١٨٠
المجموع	٢٠٠	٣٠٠	٥٠٠

هل يمكن القول بأن انتشار التدخين بين طلاب المدارس الثانوية يرتبط بطبيعة المنطقة التي تقع فيها المدرسة من حيث كونها حضرية أو ريفية؟ استخدم مستوى المعنوية ٠,٠٥.

٧- لدراسة العلاقة ما بين الرضا الوظيفي والمرتب الشهري للموظف تم أخذ عينة حجمها ٢٠٠ موظف من إحدى الجهات الحكومية. وبسؤال كل موظف عن المرتب الشهري (بالجنيه) الذي يتقاضاه وهل هو راضٍ – وظيفياً – أم لا، كانت النتائج التالية:

فئة المرتب الشهري / مستوى الرضا	أقل من ٢٥٠	٢٥٠ ــ ٥٠٠	٥٠٠ فأكثر	المجموع
مرتفع	٢٠	١٥	٥	٤٠
متوسط	٥٠	٣٥	١٥	١٠٠
منخفض	٣٠	٢٠	١٠	٦٠
المجموع	١٠٠	٧٠	٣٠	٢٠٠

هل تعطي بيانات العينة دليلاً كافياً للقول بأن هناك علاقة ما بين مستوى الرضا الوظيفي وبين المرتب الذي يتقاضاه الموظف في هذه الجهة الحكومية. استخدم مستوى المعنوية ٠,٠١.

٨- في دراسة لتقييم فعالية حملة تطعيم ضد مرض شلل الأطفال، كانت النتائج التالية فيما يتعلق بعينة مكونة من ١٠٠ طفل.

الحالة	**أُصيب بالمرض**	**لم يُصب بالمرض**	**المجموع**
تم التطعيم	١٠	٥٠	٦٠
لم يُطعَّم	٣٥	٥	٤٠
المجموع	٤٥	٥٥	١٠٠

على ضوء نتائج العينة، هل يمكن القول بأن حملة التطعيم كانت فعَّالة؟ استخدم مستوى المعنوية ٠,٠٥.

٩- لدراسة درجة التعليم في قرى محافظة قنا، أخذت أربع عينات عشوائية أحجامها ٨٠ ، ١٣٠ ، ١٠٠ ، ٩٠ من أربع قرى من المحافظة حيث وُجد أن عدد الأشخاص الذين لديهم إلمام بالقراءة والكتابة في العينات الأربع هو على الترتيب: ٥٣ ، ٩٧ ، ٩٠ ، ٦٠. هل تدل هذه البيانات على أن نسبة انتشار الأمية في القرى الأربعة متساوية؟ استخدم مستوى المعنوية ٠,٠٥.

١٠- لاحظ مدير إحدى فروع شركة عمر أفندي أن العملاء يفضلون بعض البائعين على البعض الآخر. وللتحقق من ذلك، تم تسجيل العملاء الذين دخلوا الفرع في أحد الأيام والبائعين الذين قاموا بخدمتهم فكانت النتائج التالية:

رقم البائع	١	٢	٣	٤	٥	المجموع
عدد العملاء	١٦٠	١٣٠	١٥٠	١٩٠	١٧٠	٨٠٠

هل تؤكد هذه البيانات صحة ملاحظة المدير؟ استخدم مستوى المعنوية ٠,٠١.

١١- الجدول التالي يوضح تقديرات نجاح عينة من طلاب إحدى كليات التجارة حسب نظام الدراسة وذلك في أحد الأعوام الدراسية.

نظام الدراسة / تقدير النجاح	انتظام	انتساب موجه	المجموع
جيد فأقل	٦٥	١٣٠	١٩٥
جيد جداً	٢٠	١٨	٣٨
ممتـاز	١٥	٢	١٧
المجموع	١٠٠	١٥٠	٢٥٠

هل تدل هذه البيانات على أن توزيع الطلاب المنتظمين وتوزيع الطلاب المنتسبين على تقديرات النجاح المختلفة متجانسان؟ استخدم مستوى المعنوية ٠,٠٥. وبعبارة أخرى، هل يمكن القول بأن التقدير الذي يحصل عليه الطالب يتوقف على كونه طالباً منتظماً أو منتسباً؟

١٢- تم تصنيف عينة مكونة من ١٠٠ شخص حسب لون العينيْن ولون الشعر، فكانت النتائج التالية:

لون العينيْن / لون الشعر	أزرق	أخضر	أسود	المجموع
فاتح	١٠	١٥	٥	٣٠
قاتم	٨	١٢	٥٠	٧٠
المجموع	١٨	٢٧	٥٥	١٠٠

باستخدام مستوى المعنوية ٠,٠٥، هل تستطيع القول بأنه ليست هناك علاقة بين لون العينيْن ولون الشعر؟

١٣- الجدول التالي يوضح التوزيع التكراري لعينة مكونة من ١٥٠ طالباً متغيباً عن محاضرات إحدى المواد في إحدى الكليات موزعين حسب أيام الأسبوع.

اليوم	**السبت**	**الأحد**	**الإثنين**	**الثلاثاء**	**الأربعاء**	**الخميس**	**المجموع**
عدد الغائبين	٣٠	٢٥	١٨	١٦	٢٦	٣٥	١٥٠

هل يمكن القول بأن غياب الطلاب موزع بالتساوي على أيام الأسبوع؟
استخدم مستوى المعنوية ٠,٠١.

الفصل الخامس

الإحصاءات الحيوية والصحية

مقدمـــة:

تُعتبر الإحصاءات الحيوية والصحية من العناصر الهامة في الدراسات الاجتماعية والسكانية. إذ أن تلك الإحصاءات تساعد الحكومات ومتخذي القرارات على معرفة أحوال وخصائص مواطني الدولة منذ ولادتهم وحتى وفاتهم مروراً بأحوالهم الصحية والاجتماعية. ولكون الإنسان عنصراً هاماً من عناصر الإنتاج، فإن دراسة خصائصه وأحواله صحياً واجتماعياً يضع الدولة على الطريق الصحيح في عملية اتخاذها للقرارات. كما يضمن لها فعالية خطط التنمية في المجالات المختلفة. هذا بالإضافة إلى استخدام الموارد البشرية المتاحة استخداماً أمثل يوفر للمجتمع سبل تقدمه وازدهاره بما يعود بالنفع على الشعب بأكمله.

وفي دراستنا للإحصاءات الحيوية والصحية سوف نتناول كلاً من الإحصاءات الحيوية والإحصاءات الصحية بشيء من التفصيل وذلك على النحو التالي:

الإحصاءات الحيوية

وهي البيانات المتعلقة بمجموعة الأحداث التي تصيب الإنسان منذ ولادته وحتى وفاته. لذلك فإنه من الطبيعي أن تشمل تلك الإحصاءات البيانات المتعلقة بالميلاد والزواج والطلاق إلى غير ذلك من الأحداث التي تنتهي بالوفاة. وفي تناولنا لتلك الإحصاءات. سوف نستعين في ذلك بأمثلة رقمية توضح مفهوم وكيفية حساب بعض المقاييس المتعلقة بهذا الأمر.

أولاً: إحصاءات المواليد

ويعتبر هذا النوع من الإحصاءات من العناصر الأساسية للوقوف على حركة السكان من حيث الزيادة او النقصان. وتتوافر البيانات اللازمة لتلك الإحصاءات في مكاتب الصحة وفي المستشفيات، حيث أنه عند تسجيل المولود الجديد عادة ما تطلب البيانات التالية:

تاريخ الميلاد – اسم المولود – نوع المولود – اسم الوالد ومهنته – اسم الأم – الجنسية – محل الميلاد – الديانة – حالة المولود من حيث كونه حياً أو ميتاً.

هذا وتستخدم تلك الإحصاءات في إيجاد العديد من المعدلات والمقاييس الحيوية. وهنا تلزم الإشارة إلى أن العدد (أو الرقم) الدال على أي معدل أو نسبة من المعدلات أو النسب التي سوف نتناولها بالأمثلة الرقمية إنما يعبِّر عن عدد مرات حدوث الشيء **في المتوسط** لكل ألف من الأفراد الممثّلين في مقام المعدل (أو النسبة). وسوف نخص من تلك المعدلات والنسب ما يلي:

١ـ معدل المواليد الخام:

ويمكن تعريف هذا المعدل كما يلي:

$$\textbf{معدل المواليد الخام} = \frac{\textbf{عدد المواليد الأحياء خلال العام}}{\textbf{عدد السكان في منتصف العام}} \times ١٠٠٠$$

ويبيِّن هذا المعدل عدد المواليد الأحياء لكل ألف من السكان في بلد معين في سنة محددة. ويُستخدم هذا المعدل للدلالة على قدرة السكان على التكاثر. وقد سُمي هذا المعدل بالخام أو (العام) وذلك لأنه لا يأخذ في الاعتبار الاختلافات في توزيع السكان حسب العمر والنوع. لذلك فإنه لا يصلح كمقياس للمقارنة بين دولة وأخرى أو بين المناطق المختلفة للدولة الواحدة إلا إذا تشابه التركيب العمري والنوعي لهذه الدول أو تلك المناطق.

مثال (١):

إذا كان عدد المواليد الأحياء لدولة ما خلال عام ٢٠١٨م هو ١٧٥٠٠٠٠ مولود حي. وكان تقدير عدد سكان هذه الدولة في منتصف نفس العام ٥٧٧٥٠٠٠٠ نسمة.
احسب معدل المواليد الخام.

الحـــل:

$$\text{معدل المواليد الخام} = \frac{\text{عدد المواليد الأحياء خلال عام ٢٠١٨}}{\text{عدد السكان في منتصف عام ٢٠١٨}} \times ١٠٠٠$$

$$= \frac{١٧٥٠٠٠٠}{٥٧٧٥٠٠٠٠} \times ١٠٠٠ \simeq ٣٠ \text{ في الألف}$$

أي أنه خلال عام ٢٠١٨ كان هناك ٣٠ مولوداً لكل ألف فرد من السكان في هذه الدولة.

٢ـ معدل الخصوبة العام:

ويعتبر هذا المعدل أكثر فائدة من سابقه من حيث التعرف على خصوبة السكان وذلك لأنه لا يتجاهل تركيب السكان حسب العمر والنوع. ويأخذ هذا المعدل الصورة التالية:

$$\textbf{معدل الخصوبة العام} = \frac{\textbf{عدد المواليد الأحياء خلال السنة}}{\textbf{عدد النساء في سن الحمل في منتصف السنة}} \times ١٠٠٠$$

ويبيِّن هذا المعدل مقدار ما تضيفه كل ١٠٠٠ أنثى في سن الحمل من المواليد الأحياء في سنة معينة.

ويرجع السبب إلى اعتبار النساء في سن الحمل في مقام هذا المعدل إلى أن تلك الفئة من السكان هي التي يتوقف عليها عدد المواليد في أية دولة. هذا وتجدر الإشارة الى أن تحديد سن الحمل عند النساء يختلف من دولة لأخرى وإن كانت غالبية الدولة تعتبر سن الحمل فيما بين ١٥ و ٤٩ سنة.

مثال (٢):

إذا كان عدد المواليد الأحياء في مدينة ما خلال سنة عام ٢٠١٩ هو ٧٥٠٠٠٠ مولود حي. وكان عدد النساء اللاتي في سن الحمل في منتصف هذا العام يقدر بـ ١٠١٥٠٠٠٠. احسب معدل الخصوبة العام.

الحـــل:

$$\text{معدل الخصوبة العام} = \frac{\text{عدد المواليد الأحياء خلال السنة}}{\text{عدد النساء في سن الحمل في منتصف السنة}} \times ١٠٠٠$$

$$= \frac{٧٥٠٠٠٠}{١٠١٥٠٠٠٠} \times ١٠٠٠ \simeq ٧٤ \text{ في الألف}$$

أي أن كل ١٠٠٠ أنثى في سن الحمل تضيف ٧٤ مولوداً حياً في عام ٢٠١٩م.

وتجدر الإشارة هنا إلى أن عدد النساء اللاتي في سن الحمل لا يكون معروفاً إلا في سنوات التعداد فقط. ونظراً للحاجة إلى معرفة هذا المعدل في غير سنوات التعداد، فإنه عادة ما يُفترَض ثبات نسبة النساء اللاتي في سن الحمل إلى عدد السكان في منتصف سنة التعداد وذلك في السنوات التي تلي هذه السنة وحتى إجراء تعداد آخر.

٣ــ معدل الخصوبة لفئة سن معينة (في المتوسط):

وهو المعدل الذي يُستخرَج لكل فئة من فئات سن الحمل. وعادة ما يكون طول الفئة هو خمس سنوات.

ويمكن تعريف هذا المعدل كما يلي:

معدل الخصوبة لفئة سن معينة (في المتوسط):

$$= \frac{\textbf{عدد المواليد الأحياء في هذه الفئة خلال السنة}}{\textbf{عدد النساء في هذه الفئة في منتصف السنة}} \times ١٠٠٠$$

ويبيِّن هذا المعدل عدد المواليد الأحياء لكل ١٠٠٠ أنثى في فئة معينة من فئات سن الحمل في سنة معينة.

وتبرز أهمية هذا المعدل في أنه يُعتبر مقياساً دقيقاً لمستوى الخصوبة حيث أنه لا يتأثر بالاختلاف في التركيب العمري بين النساء اللاتي في سن الحمل. بالإضافة إلى أنه يعكس السلوك الإنجابي للنساء اللاتي في سن الحمل في مختلف الأعمار.

هذا ويُقدَّر عدد النساء في فئة معينة من فئات سن الحمل كما يلي:

عدد النساء اللاتي في سن الحمل

= تقدير عدد السكان في منتصف العام × نسبة النساء اللاتي في سن الحمل وفقاً لآخر تعداد

عدد النساء في فئة معينة من فئات سن الحمل

= عدد النساء اللاتي في سن الحمل × نسبة النساء اللاتي في سن الحمل في هذه الفئة وفقاً لأخر تعداد

٤ـ معدل الخصوبة لفئة سن معينة:

وهو عبارة عن المعدل السابق مضروباً في طول فئة السن. أي أنه يعرف كما يلي:

معدل الخصوبة لفئة سن معينة:

$$= \frac{\text{عدد المواليد الأحياء في تلك الفئة خلال العام}}{\text{عدد النساء في تلك الفئة في منتصف العام}} \times \text{طول الفئة} \times ١٠٠٠$$

٥ـ معدل الخصوبة الأنثوية لفئة سن معينة:

ويبيِّن هذا المعدل عدد المواليد الإناث الأحياء لكل ١٠٠٠ أنثى في فئة معينة من فئات سن الحمل في سنة معينة. حيث نجد أنه يُعرّف كما يلي:

معدل الخصوبة الأنثوية لفئة سن معينة في المتوسط:

$$= \frac{\text{عدد المواليد الإناث الأحياء لفئة سن معينة خلال العام}}{\text{عدد النساء في هذه الفئة في منتصف العام}} \times ١٠٠٠$$

٦ـ معدل الخصوبة الأنثوية لفئة سن معينة:

وهو عبارة عن المعدل السابق مضروباً في طول الفئة. أي أن:

معدل الخصوبة الأنثوية لفئة سن معينة:

$$= \frac{\text{عدد المواليد الإناث الأحياء لفئة سن معينة خلال العام}}{\text{عدد النساء في هذه الفئة في منتصف العام}} \times \text{طول الفئة} \times ١٠٠٠$$

٧ـ معدل الخصوبة الكلى:

في كثير من الأحيان تصعب المقارنة بين الدول فيما يتعلق بمستويات الخصوبة الخاصة بفئات الأعمار المختلفة. لذلك فإنه من المفيد استخدام معدل واحد في عمليات المقارنة بحيث

يأخذ هذا المعدل في الاعتبار التغير في التركيب العمري للنساء اللاتي في سن الحمل. ومن ثم يُفضَّل هذا المعدل أيضاً على معدل الخصوبة العام. وهذا المعدل هو معدل الخصوبة الكلى.

ويشير معدل الخصوبة الكلى إلى عدد المواليد الذي تنجبه امرأة واحدة تبقى متزوجة وعلى قيد الحياة طوال مدة القدرة على الإنجاب. ولحساب هذا المعدل فإننا نحتاج إلى التوزيع العمري للنساء المتزوجات خلال فترة القدرة على الإنجاب. ويمكن التعبير عن معدل الخصوبة الكلى في الصورة التالية:

معدل الخصوبة الكلي

= مجموع حواصل ضرب (معدل الخصوبة لكل فئة عمرية × طول الفئة)

أي أن: معدل الخصوبة الكلي

$$= \sum \left(\frac{\text{عدد المواليد الإناث الأحياء لفئة سن معينة خلال العام}}{\text{عدد النساء في هذه الفئة في منتصف العام}} \times \text{طول الفئة} \right) \times ١٠٠٠$$

أي أن هذا المعدل هو عبارة عن مجموع المعدلات الخاصة بفئات الأعمار المختلفة والمعرفة في المعدل (٤).

هذا والمثال التالي يوضح كيفية حساب المعدلات المُشار إليها من قبل وهي المعدلات المعرفة في (٤): (٧).

مثال (٣):

فما يلي توزيع المواليد وعدد النساء حسب فئات الأعمار المختلفة:

فئات عمر الأم	عدد المواليد الأحياء (بالألف)	تقدير عدد النساء (بالألف)
١٥ ــ	١٢	٣٨٠
٢٠ ــ	١٨٦	٨١٠
٣٠ ــ	١١٥	٤٢٠
٣٥ ــ	٩٦	٥٢٠
٤٥ ــ ٥٠	٩	٢١٠
المجموع	٤١٨	٢٣٤٠

المطلوب:

أولاً: حساب معدلات الخصوبة التفصيلية (حسب فئات الأعمار) في المتوسط.

ثانياً: حساب معدلات الخصوبة التفصيلية (حسب فئات الاعمار).

ثالثاً: إيجاد معدل الخصوبة الكلى.

مع تفسير ما تحصل عليه من نتائج.

الحـــل:

الجدول (٥ – ١) يوضـــح كيفية حســـاب معدلات الخصـــوبة التفصـــيلية والمعدل الكلى للخصوبة.

الجدول (٥ – ١)

حساب معدلات الخصوبة حسب فئات السن ومعدل الخصوبة الكلى

(١)	(٢)	(٣)	(٤) $= \frac{(٢)}{(٣)} \times ١٠٠٠$	(٥)	(٦) = (٤) × (٥)
فئات عمر الام	عدد المواليد الأحياء (بالألف)	تقدير عدد النساء (بالألف)	معدل الخصوبة التفصيلي	طول الفئة	معدل الخصوبة التفصيلي × طول الفئة
١٥ –	١٢	٣٨٠	٣١,٥٨	٥	١٥٧,٩
٢٠ –	١٨٦	٨١٠	٢٢٩,٦٣	١٠	٢٢٩٦,٣
٣٠ –	١١٥	٤٢٠	٢٧٣,٨١	٥	١٣٦٩,١
٣٥ –	٩٦	٥٢٠	١٨٤,٦٢	١٠	١٨٤٦,٢
٤٥ – ٥٠	٩	٢١٠	٤٢,٨٦	٥	٢١٤,٣
المجموع	٤١٨	٢٣٤٠	١٧٨,٣		٥٨٨٣,٨

ويمكن تفسير النتائج الموضحة في الجدول (٥ – ١) على النحو التالي:

أننا إذا بدأنا بألف امرأة تتراوح أعمارهن ما بين ١٥ سنة إلى أقل من ٢٠ سنة فإننا نستطيع القول بأنهن أنجبن في المتوسط ٣١,٥٨ طفلاً كل سنة. وهذا يعنى أن إجمالي عدد مواليدهن خلال خمس ســنوات (١٥ – ٢٠) تبلغ ١٥٨ طفلاً (لاحظ ان: ٣١,٥٨ × ٥ ≈ ١٥٨). وبعد إنجابهن هذا العدد من الأطفال يكُنَّ قد بلغن فئة العمر (٢٠ – ٣٠) حيث معدل الإنجاب ٢٢٩,٦٣ طفلاً لكل ١٠٠٠ امرأة في العام. وعلى ذلك فإنهن ينجبن ٢٢٩٦ طفلاً خلال ١٠ سنوات. بعدها يكُنَّ قد بلغن فئة العمر (٣٠ – ٣٥) وهكذا. وبالتالي فإن إجمالي عدد مواليدهن من بداية ســن الخصــوبة (١٥ ســنة) وحتى تصــل أعمارهن إلى ٣٠ ســنة هي: ١٥٨ + ٢٢٩٦ = ٢٤٥٤ طفلاً. وهكذا يمكنّ تقدير إجمالي عدد مواليد الـ ١٠٠٠ امرأة منذ بدايـــة فترة الإنجاب وحتى نهايتها وهو ٥٨٨٤ طفلاً كما يظهر في الجدول. وهذا يعنى أن متوسـط عدد ما تنجبه ١٠٠٠ امرأة خلال فترة خصـوبتهن هو مجموع حواصـل ضـرب المعدلات التفصيلية للخصوبة في طول كل فئة وهو يساوي ٥٨٨٤.

∴ معدل الخصوبة الكلى = $\frac{٥٨٨٤}{١٠٠٠}$ = ٥,٩ ≈ ٦ أطفال لكل أم

أي أن كل أم متزوجة تظل على قيد الحياة وتمارس نفس معدلات الإنجاب الواردة في الجدول سوف تنجب في المتوسط ٦ أطفال أحياء خلال فترة العمر (١٥ إلى أقل من ٥٠ سنة).

٨ــ معدل التوالد العام:

في حساب معدلات الخصوبة المختلفة قمنا بقسمة عدد المواليد الأحياء على عدد الإناث اللاتي في سن الحمل. وفي واقع الأمر فإن مسئولية إنجاب الأطفال تقع فقط على النساء المتزوجات منهن، إذ أن هناك الكثير من الإناث في سن الحمل ممن لم يتزوجن بعد. لذلك فإنه لدراسة معدل التوالد في دولة ما فإنه تجب القسمة على عدد المتزوجات اللاتي في سن الحمل في منتصف العام المراد قياس معدل التوالد فيه. أي أن:

$$\text{معدل التوالد العام} = \frac{\text{عدد المواليد الأحياء خلال العام}}{\text{عدد المتزوجات في سن الحمل في منتصف العام}} \times ١٠٠٠$$

ويبيِّن هذا المعدل ما تضـــيفه كل ألف أنثى متزوجة وفي ســـن الحمل في ســـنة ما من المواليد الأحياء.

وللحصول على تقدير عدد المتزوجات اللاتي في سن الحمل في منتصف العام فإنه يُفترض أن نسبة المتزوجات اللاتي في سن الحمل ثابتة منذ آخر تعداد. أي أن هذه النسبة في آخر تعداد ثابتة بالنسبة للسنوات التالية له وحتى إجراء تعداد آخر. وفي ذلك فإننا نتبع أسلوباً مشابهاً لذلك الذي تم استخدامه عند تقدير عدد النساء، أي التقدير في منتصف العام.

مثال (٤):

لو افترضنا في مثال (٢) أن نسبة النساء المتزوجات من بين النساء في سن الحمل هي ٧٥٪. احسب معدل التوالد العام.

الحـــل:

$$\text{عدد النساء المتزوجات اللاتي في سن الحمل} = ١٠١٥٠٠٠٠ \times \frac{٧٥}{١٠٠}$$

$$= ٧٦١٢٥٠٠ \text{ امرأة}$$

$$\text{معدل التوالد العام} = \frac{\text{عدد المواليد الأحياء خلال العام}}{\text{عدد المتزوجات في سن الحمل في منتصف العام}} \times ١٠٠٠$$

$$= \frac{٧٥٠٠٠٠}{٧٦١٢٥٠٠} \times ١٠٠٠ \approx ٩٩ \text{ في الألف}$$

أي أن كل ألف أنثى متزوجة وفي سن الحمل تضيف ٩٩ مولوداً حياً في عام ١٩٩٦.

هذا ويمكن حساب هذا المعدل لكل فئة من فئات سن الحمل وذلك على نحو مشابه لما حدث في معدلات الخصوبة لفئات السن المختلفة.

٩ــ معدل التكاثر (التناسل) الإجمالي:

في تناولنا لمعدل التوالد العام تمت قسمة عدد المواليد الأحياء (ذكوراً وإناثاً) خلال العام على المتزوجات اللاتي في سن الحمل في منتصف العام. هذا ويُقاس تكاثر الشعوب بعدد المواليد الأحياء من الإناث. إذ أن الإناث من المواليد هن اللاتي تقع عليهن مسئولية الإنجاب في المستقبل وذلك عند بلوغهن سن الحمل. لذلك فإنه من الممكن تكوين معدل جديد يأخذ في الاعتبار عدد المواليد الأحياء من الإناث بدلاً من عدد المواليد الأحياء ذكوراً وإناثاً. وهذا المعدل الجديد يُعرف بمعدل التكاثر (التناسل) الإجمالي وهو يهتم بتقدير عدد الأطفال الإناث الذي يمكن أن تنجبه ألف سيدة تبقيْن على قيد الحياة طوال فترة الحمل.

ولحساب معدل التكاثر الإجمالي نتبع الخطوات التالية:

أ- يُحسب معدل التكاثر لكل فئة من فئات سن الحمل باستخدام الصيغة التالية:

معدل التكاثر لفئة معينة

$$= \frac{\text{عدد المواليد الأحياء من الإناث لهذه الفئة خلال العام}}{\text{عدد المتزوجات في هذه الفئة في منتصف العام}} \times \text{طول الفئة} \times ١٠٠٠$$

ب- بجمع معدلات التكاثر الخاصة بالفئات المختلفة نحصل على معدل التكاثر الإجمالي.

والمثال التالي يوضح ذلك.

مثال (٥):

الجدول التالي يوضح عدد المواليد الإناث (بالألف) وعدد المتزوجات (بالألف) وذلك للفئات المختلفة لعمر الأم في أحد البلدان (بيانات افتراضية):

فئات عمر الأم	١٥ ــ	٢٥ ــ	٣٠ ــ	٣٥ ــ	٤٥ ــ ٥٠
عدد المواليد الإناث (بالألف)	٢٨	٣٤	٥٢	١٤	١
عدد المتزوجات (بالألف)	٣٧٥	٢٠٠	١٥٠	٢٦٠	١٠٠

المطلـــوب:

حساب معدل التكاثر الإجمالي مع تفسير النتيجة التي تحصل عليها.

الحـــل:

لحساب معدل التكاثر الإجمالي يلزم حساب معدلات التكاثر التفصيلية لفئات سن الحمل المختلفة وذلك كما يوضحه الجدول التالي:

الجدول (٥ ــ ٢)
حساب معدلات التكاثر التفصيلية ومعدل التكاثر الإجمالي

(١)	(٢)	(٣)	(٤)	(٥) = $\frac{(٢)}{(٣)}$ × (٤) × ١٠٠٠
فئات عمر الأم	عدد المواليد الإناث (بالألف)	عدد المتزوجات (بالألف)	طول الفئة	معدل التكاثر التفصيلي
١٥ ــ	٢٨	٣٧٥	١٠	٧٤٦,٦
٢٥ ــ	٣٤	٢٠٠	٥	٨٥٠,٠
٣٠ ــ	٥٢	١٥٠	٥	١٧٣٣,٣
٣٥ ــ	١٤	٢٦٠	١٠	٥٣٨,٥
٤٥ ــ ٥٠	٢	١٠٥	٥	٩٥,٢
المجموع	١٣٠	١٠٩٠		٣٩٦٣,٦

ويمكننا تفسير النتائج الواردة في الجدول (٥ ــ ٢) على النحو التالي:
أننا إذا بدأنا بألف امرأة متزوجة تتراوح أعمارهن ما بين ١٥ إلى أقل من ٢٥ سنة، فإننا نستطيع القول بأنهن أنجبن ٧٤٧ طفلاً أنثى خلال عشـــر سنوات. وبعد إنجابهن هذا العدد يكُنّ قد بلغن فئة العمر (٢٥ ــ ٣٠) حيث ينجبن خلالها ٨٥٠ طفلاً أنثى آخرين. وبالتالي فإن إجمالي عدد مواليدهن من الإناث من بداية سن الخصوبة (١٥ سنة) وحتى تصل أعمارهن إلى ٣٠ سنة هي: ٧٤٧ + ٨٥٠ = ١٥٩٧ طفلاً أنثى. وهكذا يمكن الوصول إلى أن إجمالي عدد المواليد الإناث للألف امرأة متزوجة منذ بداية الإنجاب وحتى نهايتها هو ٣٩٦٤ طفلاً أنثى وذلك كما يوضحه مجموع العمود الخامس في الجدول (٥ ــ ٢). وهذا يعنى أن متوسط ما تنجبه المــرأة المتزوجة خلال فترة خصوبتها هو ما يُعبَّر عنه بـ "معدل التكاثر الإجمالي لكل أم" وذلك كما يلي:

$$\text{معدل التكاثر الإجمالي} = \frac{٣٩٦٤}{١٠٠٠} \approx ٤ \text{ أطفال إناث لكل أم}$$

أي أن كل أم متزوجة تظل على قيد الحياة وتمارس نفس معدلات إنجاب الإناث الواردة في الجدول سوف تنجب في المتوسط ٤ أطفال إناث خلال فترة العمر (١٥ سنة إلى أقل من ٥٠ سنة).

١٠ــ معدل التكاثر (التناسل) الصافي:

في حسابنا لمعدل التكاثر (التناسل) الإجمالي كان هناك افتراض بأن كل مولودة أنثى سوف تظل على قيد الحياة حتى تصل إلى العمر الذي وصلت إليه أمها. وهذا ــ بطبيعة الحال ــ افتراض غير صحيح دائماً، إذ أن عدد المواليد الإناث يتناقص بمرور السنوات وذلك

نتيجةً للوفيات. وعلى سبيل المثال قد نجد أنه من بين ألف مولودة أنثى لا يبقى على قيد الحياة حتى فئة العمر ١٥ ــ ٢٠ سوى ٧٥٠ فقط. وهذا العدد قد يتناقص إلى ٧٠٠ فقط في فئة العمر (٢٠ ــ ٢٥) وهكذا. وتتم معرفة عدد الباقين على قيد الحياة في فئات العمر المختلفة من المواليد الإناث بالاستعانة بما يعرف بـ "جداول الحياة". وهنا إذا ما أخذنا في الاعتبار تأثير عوامل الوفيات فإنه يمكننا الحصول على معدل جديد هو معدل التكاثر (التناسل) الصافي، وهو المعدل الذي يستبعد أثر الوفيات على المواليد الإناث. ويتم حساب هذا المعدل عن طريق ضرب معدل التكاثر الإجمالي لكل فئة من فئات سن الأم فـي نسبة الباقين على قيد الحياة من كل ألف مولودة أنثى (يتم الحصول علـى هذه النسبة باستخدام جداول الحياة) لكل فئة من فئات سن الأم. بعد ذلك نقوم بجمع حواصل الضرب للفئات المختلفة فنحصل على معدل التكاثر (التناسل) الصافي.

هذا ومعدل التكاثر (التناسل) الصافي يبيِّن مدى إحلال الجيل القادم محل الجيل الحاضر. أي أنه يوضح عدد المواليد الإناث لألف امرأة واللاتـي (أي المواليد الإناث) سوف يبلغن أعمار أمهاتهن. فإذا بلغ معدل التكاثر الصافي ١٠٠٠ (لكل ألف امرأة) فإن هذا يدل على أن الجيل يعوض نفسه. وأما إذا قل معدل التكاثر الصافي عن ١٠٠٠ فإن ذلك معناه أن الجيل لا يعوض نفسه وأن عدد السكان يتناقص بمرور الأجيال. وفي حالة زيادة هذا المعدل عن ١٠٠٠ فإن ذلك يعنى أن الأجيال تعوض ما يزيد على ما تفقده مما يدل على زيادة عدد السكان بمرور السنوات. وفي الواقع فإن معدل التكاثر (التناسل) الصافي هو أدق المقاييس السابقة في الحكم على قياس خصوبة الشعوب. ولعل المثال التالي يوضح كيفية حساب هذا المعدل وتفسير معناه.

مثال (٦):

باستخدام بيانات المثال السابق، وبافتراض أن عدد الباقيات على قيد الحياة من كل ألف مولودة أنثى لكل فئة من فئات السن الواردة في هذا المثال كان: ٦٨٠، ٦٥٠، ٦٢٥، ٦٠٠، ٥٠٠ على التوالي. احسب معدل التكاثر (التناسل) الصافي، مع تفسير النتيجة التي تحصل عليها.

الحـــل:

الجدول **(٥ - ٣)** يوضِّح الحسابات اللازمة لإيجاد معدل التكاثر (التناسل) الصافي وذلك وفقاً لما سبق شرحه.

هذا مع ملاحظة أن الحسابات المتعلقة بالعمود (٢) في الجدول (٥ ــ ٣) قد تم إعدادها في مثال (٥).

ومن الجدول **(٥ ــ ٣)** يمكننا استنتاج أن معدل التكاثر (التناسل) الصافي (لكل ١٠٠٠ امرأة) هو ٢٥١٤ طفلاً أنثى.

الجدول (٥ – ٣)
حساب معدل التكاثر (التناسل) الصافي

(١)	(٢)	(٣)	(٤) = (٣) ÷ ١٠٠٠	(٥) = (٢) × (٤)
فئات عمر الأم	معدل التكاثر التفصيلي	عدد الباقيات على قيد الحياة من ١٠٠٠ مولود أنثى	نسبة الباقيات على قيد الحياة من المواليد الإناث	معدل التكاثر الصافي التفصيلي
١٥ –	٧٤٦,٦	٦٨٠	٠,٦٨٠	٥٠٧,٧
٢٥ –	٨٥٠,٠	٦٥٠	٠,٦٥٠	٥٥٢,٥
٣٠ –	١٧٣٣,٣	٦٢٥	٠,٦٢٥	١٠٨٣,٣
٣٥ –	٥٣٨,٥	٦٠٠	٠,٦٠٠	٣٢٣,١
٤٥ – ٥٠	٩٥,٢	٥٠٠	٠,٥٠٠	٤٧,٦
المجموع				٢٥١٤,٢

أي أن معدل التكاثر الصافي لكل امرأة متزوجة هو:

معدل التكاثر الصافي = $\frac{٢٥١٤}{١٠٠٠}$ = ٢,٥ طفلاً أنثى في المتوسط

وهذا يعني أن كل أم متزوجة تمارس نفس معدلات إنجاب الإناث الواردة في مثال (٥) سوف تنجب في المتوسط ٢,٥ أنثى يُتوقَّع أن تعشن حتى سن الأم، وذلك بافتراض انطباق نسب البقاء على قيد الحياة الواردة في مثال (٦) عليهن.

١١– معدل المواليد موتى:

ويُعرَّف هذا المعدل كما يلي:

معدل المواليد الموتى = $\frac{\text{عدد المواليد موتى خلال العام}}{\text{عدد المواليد (أحياء + موتى) خلال العام}}$ × ١٠٠٠

ويبيّن هذا المعدل عدد المواليد موتى لكل ألف من إجمالي المواليد الأحياء والموتى معاً في سنة معينة.

١٢– نسبة المواليد موتى:

وهي تختلف عن معدل المواليد موتى في أن المقام هنا يشمل المواليد الأحياء فقط حيث تأخذ هذه النسبة الصورة:

نسبة المواليد الموتى = $\frac{\text{عدد المواليد موتى خلال العام}}{\text{عدد المواليد أحياء خلال العام}}$ × ١٠٠٠

وتعطي هذه النسبة عدد المواليد موتى لكل ألف مولود حي في سنة معينة.

هذا ومعدل المواليد موتى ونسبة المواليد موتى يعكسان الحالة الصحية للأم والتي تتوقف بدورها على مستوى الرعاية الطبية للأمهات أثناء الحمل.

مثال (٧):

لنفترض أن إجمالي عدد المواليد في مدينة ما في عام ٢٠١٧م كان ٢٠٠٠٠٠ مولوداً كان من بينهم ١٠٠٠٠ مولوداً ميتاً. احسب معدل المواليد أموات في هذه المدينة في هذا وكذلك نسبة المواليد موتى في نفس العام.

الحــــل

عدد المواليد الأحياء = ٢٠٠٠٠٠ – ١٠٠٠٠ = ١٩٠٠٠٠ مولود

$$\therefore \text{معدل المواليد الأموات} = \frac{\text{عدد المواليد موتى}}{\text{عدد المواليد (أحياء + موتى)}} \times ١٠٠٠$$

$$= \frac{١٠٠٠٠}{٢٠٠٠٠٠} \times ١٠٠٠ = ٥٠ \text{ في الألف}$$

أي أنه يولد ٥٠ مولوداً ميتاً لكل ألف مولود في هذه المدينة في عام٢٠١٧م.

$$\text{نسبة المواليد موتى} = \frac{\text{عدد المواليد موتى خلال العام}}{\text{عدد المواليد أحياء خلال العام}} \times ١٠٠٠$$

$$= \frac{١٠٠٠٠}{١٩٠٠٠٠} \times ١٠٠٠ = ٥٢,٦ \approx ٥٣ \text{ في الألف}$$

أي أنه يولد ٥٣ مولوداً ميتاً لكل ألف مولود حي في هذه المدينة في عام ٢٠١٧م.

ثانياً: إحصاءات الوفيات

تُعتبر معدلات الوفيات من المؤشرات الجيدة للمستوى الصحي في المجتمع. ليس هذا فقط بل إن هذه المعدلات قد تكون ذات دلالة على قياس المستوى الحضاري للدول.
هذا ودراسة الوفيات على قدر كبير من الأهمية ذلك لأنها تؤدى دوراً معاكساً لهذا الدور الذي تؤديه المواليد وذلك من حيث التأثير على الزيادة الطبيعية في أعداد السكان. إذ أنه من المعروف أن الزيادة الطبيعية في أعداد السكان ما هي إلا محصلة الفرق بين أعداد المواليد وأعداد الوفيات.
وكما كان نهجنا في دراسة المواليد، فإننا هنا نقوم بتناول بعض المعدلات المتعلقة بالوفيات مستعينين في ذلك بالبيانات المتوافرة عنها. هذا وعند التبليغ عن حالة وفاة فإن القانون يلزم بتسجيل البيانات المتعلقة بالشخص المتوفَّي ومنها:
سبب الوفاة – مكان الوفاة – تاريخ الوفاة – اسم المتوفي – العمر – المهنة – الجنسية – الديانة – الحالة الزواجية – عدد الأولاد – الحالة التعليمية، إلى غير ذلك من البيانات.

وتُستخدم البيانات والإحصاءات المتوافرة عن حالات الوفيات في إيجاد العديد من المعدلات والتي نتناول منها الآن ما يلي:

١ـ معدل الوفيات العام (الخام):

وهذا المعدل يشبه معدل المواليد العام (الخام)، فهو عبارة عن نسبة عدد الوفيات إلى عدد السكان مضروبة في ١٠٠٠. أي أن هذا المعدل يأخذ الصورة التالية:

$$\textbf{معدل الوفيات الخام} = \frac{\textbf{عدد الوفيات خلال السنة}}{\textbf{تقدير عدد السكان في منتصف السنة}} \times ١٠٠٠$$

أي أنه يعطي عدد الوفيات لكل ألف شخص من السكان في السنة. ويُلاحَظ هنا أن المواليد موتى لا يدخلون ضمن الوفيات.

وهذا المعدل وإن أفاد في التعرف على المستوى العام للوفيات، إلا أنه لا يصلح في عمليات المقارنة بين الدول أو بين مناطق الدولة الواحدة. ذلك لأنه ـ شأنه في ذلك شأن معدل المواليد العام ـ لا يأخذ في الاعتبار الاختلافات في التركيب العمري والنوعي للسكان.

مثال (٨):

إذا كان عدد سكان إحدى المدن في منتصف عام ٢٠١٨م هو ٢,٥ مليون نسمة وكان عدد الوفيات خلال نفس السنة هو ٦٢٥٠ فرداً. احسب معدل الوفيات الخام.

الحــل:

$$\text{معدل الوفيات الخام} = \frac{\text{عدد الوفيات خلال عام ٢٠١٨}}{\text{تقدير عدد السكان في منتصف عام ٢٠١٨}} \times ١٠٠٠$$

$$= \frac{٦٢٥٠}{٢٥٠٠٠٠٠} \times ١٠٠٠ = ٢,٥ \text{ في الألف}$$

أي أن هناك ٢,٥ حالة وفاة في المتوسط لكل ألف من السكان.

٢ـ معدل الوفيات في فئة عمر معينة:

وهذا المعدل يمكن أن يفيد في عمليات المقارنة المُشار إليها سابقاً. وهو يأخذ الصيغة التالية:

معدل الوفيات في فئة عمر معينة

$$= \frac{\textbf{عدد الوفيات في هذه الفئة خلال العام}}{\textbf{تقدير عدد السكان في هذه الفئة في منتصف العام}} \times ١٠٠٠$$

كما يمكن إيجاد هذا المعدل بالنسبة لكل من الذكور والإناث على حدة وهو ما يعبِّر عنه المعدلان التاليان.

٣ــ معدل الوفيات للذكور في فئة عمر معينة:

وهذا المعدل لا يفيد فقط في أمور المقارنات بين الدول المختلفة ولكن يفيد أيضاً في الوقوف على مستوى وفيات الذكور في فئات السن المختلفة داخل الدولة الواحدة. وهو يأخذ الصورة:

معدل الوفيات للذكور في فئة عمر معينة

$$= \frac{\text{عدد وفيات الذكور في هذه الفئة خلال السنة}}{\text{عدد الذكور في هذه الفئة في منتصف السنة}} \times ١٠٠٠$$

٤ــ معدل الوفيات للإناث في فئة عمر معينة:

ويمكن أن يُقال عن هذا المعدل ما سبق وأن قبل عن نظيره في حالة الذكور. ويُعبَّر عن هذا المعدل كما يلي:

معدل الوفيات للإناث في فئة عمر معينة

$$= \frac{\text{عدد وفيات الإناث في هذه الفئة خلال السنة}}{\text{عدد الإناث في هذه الفئة في منتصف السنة}} \times ١٠٠٠$$

مثال (٩):

الجدول التالي يوضح التوزيع العمري لسكان إحدى المدن حسب الجنس وأعدد الوفيات بها خلال عام ٢٠٢٠م (بيانات افتراضية):

فئات الأعمار	عدد السكان (بالآلاف)		عدد الوفيات
	ذكور	إناث	
صفر ــ	١٣٠	١١٦	٦٥٠٠
٥ ــ	١٠٨	١٠٠	٩٠٠
١٠ ــ	١١٦	١١٨	٤٨٠
١٥ ــ	١٠٠	٩٨	٧٠٠
٢٠ ــ	١٥٠	١٦٠	١٥٠٠
٣٠ ــ	١٢٦	١٢٠	١٦٠٠
٤٠ ــ	٩٠	٩٠	٢٠٠٠
٥٠ فأكثر	٨٠	٧٨	٤٠٠٠
المجموع	٩٠٠	٨٨٠	١٧٦٨٠

المطلــوب:

أولاً: حساب معدل الوفيات الخام.

ثانياً: إيجاد معدل الوفيات لفئة العمر (٢٠ – ٣٠) وذلك لكل مما يلي:

أ- الذكور **ب-** الإناث **ج-** الذكور والإناث معاً

وذلك بافتراض أن عدد وفيات الذكور في هذه الفئة هو ٨٠٠ شخصاً.

الحـــل:

أولاً: معدل الوفيات الخام $= \frac{\text{عدد الوفيات خلال عام ٢٠٢٠}}{\text{عدد السكان في منتصف عام ٢٠٢٠}} \times ١٠٠٠$

$$= \frac{١٧٦٨٠}{(٨٨٠٠٠٠ + ٩٠٠٠٠٠)} \times ١٠٠٠$$

$$= \frac{١٧٦٨٠}{١٧٨٠٠٠٠} \times ١٠٠٠$$

= ٩,٩ في الألف ≈ ١٠ في الألف

أي أن هناك ١٠ حالات وفاة بين كل ١٠٠٠ شخص من السكان وذلك في عام ٢٠٢٠م.

ثانياً:

أ- معدل وفيات الذكور في فئة العمر (٢٠ – ٣٠):

$$= \frac{\text{عدد وفيات الذكور في هذه الفئة}}{\text{عدد الذكور في هذه الفئة}} \times ١٠٠٠$$

$$= \frac{٨٠٠}{١٥٠٠٠٠} \times ١٠٠٠ \approx ٥ \text{ في الألف}$$

أي أن هناك ٥ حالات وفاة بين كل ألف من الذكور في هذه الفئة من العمر وذلك في عام ٢٠٢٠م.

ب- معدل وفيات الإناث في فئة العمر (٢٠ – ٣٠):

$$= \frac{\text{عدد وفيات الإناث في هذه الفئة}}{\text{عدد الإناث في هذه الفئة}} \times ١٠٠٠$$

$$= \frac{(١٥٠٠ - ٨٠٠)}{١٥٠٠٠٠} \times ١٠٠٠$$

$$= \frac{٧٠٠}{١٥٠٠٠٠} \times ١٠٠٠$$

≈ ٤ في الألف

أي أن هناك ٤ حالات وفاة بين كل ألف أنثى في هذه الفئة من العمر.

ج- معدل الوفيات في فئة العمر (٢٠ – ٣٠):

$$= \frac{\text{عدد الوفيات في هذه الفئة خلال العام}}{\text{تقدير عدد السكان في هذه الفئة في منتصف العام}} \times ١٠٠٠$$

$$= \frac{١٥٠٠}{(١٦٠٠٠٠+١٥٠٠٠٠)} \times ١٠٠٠ = \frac{١٥٠٠}{٣١٠٠٠٠} \times ١٠٠٠$$

≈ ٥ في الألف

وهذا يعني أن هناك ٥ حالات وفاة بين كل ألف من الأشخاص في هذه الفئة من العمر وذلك في عام ٢٠٢٠م.

هذا – وكما ذكرنا من قبل – فإن معدل المواليد الخام لا يصلح في عمليات المقارنات بين الدول المختلفة أو بين مناطق الدولة الواحدة. لذلك يتم اللجوء إلى حساب هذا المعدل لكل نوع ولفئات العمر المختلفة. ومع ذلك فإن إجراء المقارنات مستخدمين عدداً كبيراً من المعدلات يمثل عملية شاقة ومُجهِدة. ويصبح الأمر الأيسر والأسهل تطبيقاً هو استخدام معدل الوفيات الخام ولكن بعد تصحيحه وذلك بالشكل الذي يجعله ملائماً لعمليات المقارنات المُشار إليها. وهناك طريقتان لتصحيح معدل الوفيات الخام ولكنهما لا يدخلان ضمن نطاق هذا الكتاب.

٥– معدل وفيات الأطفال الرضع:

الأطفال الرضـــع هم الأطفال الذين تقل أعمارهم عن العام. ومعدل وفياتهم يأخذ الصورة التالية:

معدل وفيات الأطفال الرضع

$$= \frac{\textbf{عدد وفيات الأطفال دون العام خلال السنة}}{\textbf{عدد المواليد أحياء خلال السنة}} \times ١٠٠٠$$

وهذا المقياس على قدر كبير من الأهمية، إذ أنه يعكس مستوى التقدم في الدولة، ليس من الناحية الصحية فحسب بل أيضاً من الناحية الاقتصادية والاجتماعية، وذلك لأن الأطفال الرضع هم الأقل قدرة على تحمل سوء المستوى الاقتصادي والاجتماعي. هذا بالإضافة إلى أن وفاة الأطفال دون العام تعكس أيضاً الحالة الصحية والاجتماعية للأبويْن والتي تكون بدورها انعكاساً للأحوال الصحية والاجتماعية والاقتصادية في الدولة ككل.

مثال (١٠):

إذا كان عدد المواليد الأحياء في إحدى الدول في عام ٢٠١٩م هو ٩٠٠٠٠٠ مولود حي. وكان عدد وفيات الأطفال قبل بلوغهم العام الأول هو ٤٥٠٠. احسب معدل وفيات الأطفال الرضع.

الحــل:

معدل وفيات الأطفال الرضع

$$= \frac{\text{عدد وفيات الأطفال الرضع خلال السنة}}{\text{عدد المواليد أحياء في منصف السنة}} \times ١٠٠٠$$

$$= \frac{٤٥٠٠}{٩٠٠٠٠٠} \times ١٠٠٠ \approx ٥ \text{ في الألف}$$

أي أن هناك خمس حالات وفاة للأطفال قبل بلوغهم العام الأول وذلك من بين كل ١٠٠٠ مولود حي في عام ٢٠١٩م.

٦- معدل الوفيات حسب سبب الوفاة:

أحياناً ما نحتاج إلى إيجاد معدل الوفيات بحسب سبب الوفاة كالإصابة بمرض معين أو غير ذلك. وفي هذا الشأن يأخذ معدل الوفيات الصورة التالية:

معدل الوفيات حسب سبب الوفاة

$$= \frac{\textbf{عدد الوفيات الناشئة عن سبب معين خلال العام}}{\textbf{تقدير عدد السكان في منتصف العام}} \times ١٠٠٠$$

٧- معدل وفيات الأمومة:

ويعرَّف هذا المعدل كما يلي:

$$\textbf{معدل وفيات الأمومة} = \frac{\textbf{عدد وفيات النساء أثناء الحمل و الولادة}}{\textbf{عدد المواليد أحياء خلال العام}} \times ١٠٠٠$$

ويعكس هذا المعدل أيضاً مستوى الرعاية الصحية التي تُقدم للأمهات أثناء فترة الحمل والولادة.

ثالثاً: إحصاءات الزواج والطلاق

إحصاءات الزواج:

تُستخرج إحصاءات الزواج من سجلات الزوج والتي تشمل عادة البيانات التالية لكلل من الزوجيْن:

الاسم – العمر – الجنسية – عدد مرات الزواج السابقة – الحالة التعليمية – المهنة – الديانة... إلى غير ذلك من البيانات المتعلقة بهذا الشأن.

وتُستخدم هذه الإحصاءات في استخراج المعدلات المتعلقة بالزواج والتي من بينها ما يلي:

١– معدل الزواج الخام:

ويأخذ هذا المعدل الصورة التالية:

$$\textbf{معدل الزواج الخام} = \frac{\textbf{عدد الزيجات خلال السنة}}{\textbf{عدد السكان في منتصف السنة}} \times ١٠٠٠$$

ويبيّن هذا المعدل عدد الزيجات التي تمت خلال العام لكل ألف من السكان.

ولا يصلح هذا المعدل لأغراض المقارنة بين الدول إذ أنه لا يأخذ في الاعتبار الاختلاف في تركيب السكان ولابد من تعديله حتى يكون صالحاً في هذا الشأن.

٢– معدل الزواج المعدَّل:

في معدل الزواج الخام تمت قسمة عدد الزيجات على عدد السكان في منتصف العام. ومن هنا تأتى عدم دقة هذا المعدل، إذ أن هناك جزءاً كبيراً من السكان لا يكونون معرضين للزواج. فالأطفال مثلاً ليسوا معرضين لهذه الظاهرة ومن ثم يجب استبعادهم. أضف إلى ذلك أن من هم في سن الزواج فعلاً هناك جزء منهم ليسوا معرضين للزواج وهم المتزوجون بالفعل، إذ أن احتمال تعرضهم للزواج مرة أخرى قليل. لذلك يجب استبعاد هذه الفئة أيضاً من عدد السكان الذين هم في سن الزواج. وبذلك نجد أن معدل الزواج المعدل يأخذ الصورة التالية:

$$\textbf{معدل الزواج المعدَّل} = \frac{\textbf{عدد الزيجات خلال السنة}}{\textbf{عدد السكان في سن الزواج و غير متزوجين منتصف السنة}} \times ١٠٠٠$$

وهذا المعدل هو أكثر دقةً من سابقه، وهو يبيِّن عدد الزيجات لكل ألف شخص في سن الزواج وغير متزوجين. كما يصلح هذا المعدل لقياس درجـة الإقبال على الزواج. ويُطلق على هذا المعدل أحياناً معدل تكوين الأسر الجديدة.

هذا ويختلف سن الزواج من دولة لأخرى، وإن كان هذا السن يتحدد في غالبية الدول بـ ١٨ سنة بالنسبة للرجل و١٦ سنة بالنسبة للأنثى. ويتم تقدير عدد السكان في سن الزواج وغير متزوجين وذلك بضرب عدد السكان التقديري في منتصف العام في نسبة غير المتزوجين وفي سن الزواج في آخر تعداد. وأما غير المتزوجين فهم الذين لم يسبق لهم الزواج والمطلِّقون والمطلَّقات والأرامل.

وفي هذا المعدل أيضاً يمكن استخدام عدد المتزوجين (بدلاً من عدد الزيجات) في البسط. إلا أن المعدل في هذه الحالة يكون ضعف المعدل الأول إلا في الحالات التي تنتشر فيها عادة تعدد الزوجات.

مثال (١١):

إذا كان عدد الزيجات التي تمت في إحدى المدن ما خلال عام ٢٠١٨م هو ٢٥٠٠٠٠ حالة زواج. وكان تقدير عدد السكان في منتصف هذا العام هو ٥٠ مليون نسمة، بينهم ١٠ مليون شخص في سن الزواج وغير متزوجين.

المطلوب:

حساب المعدلات التالية:

أ- معدل الزواج الخام **ب-** معدل الزواج المعدل

الحـــل:

$$\text{أ- معدل الزواج الخام} = \frac{\text{عدد الزيجات خلال السنة}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$$

$$= \frac{٢٥٠٠٠٠}{٥٠٠٠٠٠٠٠} \times ١٠٠٠ = ٥ \text{ في الألف}$$

أي أن هناك خمس حالات زواج لكل ألف من سكان هذه الدولة في عام ٢٠١٨م.

$$\text{ب- معدل الزواج المعدل} = \frac{\text{عدد الزيجات خلال العام}}{\text{عدد السكان في سن الزواج و غير متزوجين منتصف العام}} \times ١٠٠٠$$

$$= \frac{٢٥٠٠٠٠}{١٠٠٠٠٠٠٠} \times ١٠٠٠ = ٢٥ \text{ في الألف}$$

وهذا يعني أن هناك ٢٥ حالة زواج لكل ألف من السكان في سن الزواج وغير متزوجين.

إحصاءات الطلاق:

تُستخرج إحصاءات الطلاق من سجلات الطلاق والتي تحتوي عادة على البيانات التالية لكل من المطلِّق والمطلَّقة:

الاسم – العمر – عدد مرات الزواج – الحالة التعليمية – المهنة – سبب الطلاق – مدة الحياة الزواجية – عدد الأولاد، إلى غير ذلك من البيانات.

وتُستخدم هذه الإحصاءات في استخراج العديد من المعدلات المتعلقة بالطلاق ومنها:

٣ـ معدل الطلاق الخام:

ويُحسب هذا المعدل بقسمة عدد حالات الطلاق خلال العام على تقدير السكان في منتصف العام ثم الضرب في ألف. وهو يبين عدد حالات الطلاق لكل ألف من السكان، ويأخذ الصورة التالية:

$$\text{معدل الطلاق الخام} = \frac{\text{عدد حالات الطلاق خلال العام}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$$

كما يُحسب هذا المعدل أيضاً باعتبار عدد المطلقين بدلاً من عدد حالات الطلاق وذلك في بسط هذا المعدل.

هذا ولا يصلح المعدل لقياس استقرار الحياة الزواجية لأن عدد السكان الذي يمثل مقام هذا المعدل يشمل غير المتزوجين والذين لا تنطبق عليهم حالات الطلاق، كما أنه يشمل الأطفال ومن هم دون سن الزواج. لذلك لابد من إجراء تعديل على هذا المعدل حتى يكون أكثر ملاءمة في قياس استقرار الحياة الزواجية. وفي عمليات المقارنة بين أكثر من دولة أو بين مناطق الدولة الواحدة.

٤ـ معدل الطلاق المعدل:

وفي هذا المعدل ينسب عدد حالات الطلاق أو عدد المطلقين إلى عدد المتزوجين في منتصف العام ثم يضرب الناتج في ألف.. أي أن:

$$\text{معدل الطلاق المعدَّل} = \frac{\text{عدد حالات الطلاق أو عدد المطلقين أثناء السنة}}{\text{عدد غير المتزوجين منتصف السنة}} \times ١٠٠٠$$

ويبيِّن هذا المعدل عدد حالات الطلاق أو عدد المطلقين لكل ألف شخص متزوج. ويمكن الحصول على تقدير لعدد المتزوجين في منتصف العام عن طريق تقدير عدد السكان في منتصف العام وضربه في نسبة المتزوجين المستخرَجة من آخر تعداد.

٥ـ معدل الطلاق الخاص:

في هذا المعدل ينسب عدد الذين طلِّقوا أثناء السنة إلى عدد المتزوجين فعلاً خلال السنة ثم ضرب الناتج في ألف. أي أن:

$$\text{معدل الطلاق الخاص} = \frac{\text{عدد الذين طلقوا خلال السنة}}{\text{عدد المتزوجين فعلاً خلال السنة}} \times ١٠٠٠$$

ويبين هذا المعدل عدد المطلقين لكلل ألف شخص متزوج فعلا خلال السنة.

هذا وعند إجراء مقارنة بين معدلات الطلاق في الدول المختلفة أو بين مناطق الدولة الواحدة يجب أن تراعَى الاختلافات بين العادات والديانات ونظم وقوانين الطلاق.

مثال (١٢):

إذا كان تقدير عدد السكان في منتصف أحد الأعوام في إحدى المدن هو ٥ مليون نسمة، وكانت عدد الزيجات التي تمت خلال العام هي ٤٠ ألف حالة زواج. كما بلغت عدد حالات الطلاق أثناء العام ١٠ آلاف حالة. فإذا علمت أنه حسب آخر تعداد كانت نسبة السكان في سن الزواج ٢٠٪ وأن غير المتزوجين يمثلون ١٥٪ من عدد السكان في سن الزواج.

المطلـــوب:

إيجاد المعدلات التالية:

أ- معدل الزواج الخام **ب-** معدل الزواج المعدل

ج- معدل الطلاق الخام **د-** معدل الطلاق المعدل

مع تفسير النتائج التي تحصل عليها.

الحـــل:

أ- معدل الزواج الخام = $\frac{\text{عدد الزيجات خلال العام}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$

$= \frac{٤٠٠٠٠}{٥٠٠٠٠٠٠} \times ١٠٠٠ = ٨$ في الألف

أي أن هناك ٨ حالات زواج لكل ألف من السكان.

ب- معدل الزواج المعدل = $\frac{\text{عدد الزيجات خلال العام}}{\text{عدد السكان في سن الزواج و غير متزوجين}} \times ١٠٠٠$

عدد السكان في سن الزواج وغير متزوجين = $٥٠٠٠٠٠٠ \times \frac{٢٠}{١٠٠} \times \frac{١٥}{١٠٠}$

= ١٥٠٠٠٠ شخص

$\therefore$ معدل الزواج المعدل = $\frac{٤٠٠٠٠}{١٥٠٠٠٠} \times ١٠٠٠ \approx ٢٦٧$ في الألف

مما أن هناك ٢٦٧ حالة زواج لكل ألف من السكان في سن الزواج وغير متزوجين.

ج- معدل الطلاق الخام = $\frac{\text{عدد حالات الطلاق خلال العام}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$

$= \frac{١٠٠٠٠}{٥٠٠٠٠٠٠} \times ١٠٠٠ = ٢$ في الألف

أي أن هناك حالتي طلاق لكل ألف من السكان.

د- معدل الطلاق المعدَّل = $\frac{\text{عدد حالات الطلاق أثناء السنة}}{\text{عدد المتزوجين في منتصف السنة}}$ × ١٠٠٠

وحيث أن: عدد المتزوجين = ٥٠٠٠٠٠٠ × $\frac{٢٠}{١٠٠}$ × $\frac{٨٥}{١٠٠}$

= ٨٥٠٠٠٠ شخص

∴ معدل الطلاق المعدَّل = $\frac{١٠٠٠٠}{٨٥٠٠٠٠}$ × ١٠٠٠ ≈ ١٢ في الألف

وهذا يعني أن هناك ١٢ حالة طلاق بين كل ألف من المتزوجين.

مثال (١٣):

بافتراض أن عدد المتزوجين فعلاً خلال إحدى السنوات في بلد ما هو ٥٠٠٠ شخص، وأن عدد الذين طلقوا خلال نفس السنة هو ٢٠٠شخص. احسب معدل الطلاق الخاص.

الحـــــل:

معدل الطلاق الخاص = $\frac{\text{عدد الذين طلقوا خلال السنة}}{\text{عدد المتزوجين فعلاً خلال العام}}$ × ١٠٠٠

= $\frac{٢٠٠}{٥٠٠٠}$ × ١٠٠٠ = ٢ في الألف

أي أن هناك ٤٠ شخصاً شملتهم حالات الطلاق لكل ألف ممن تزوجوا فعلاً خلال العام.

الإحصاءات الصحية

إن الاهتمام بصحة الإنسان هو هدف تسعى إليه جميع الحكومات. وفي سبيل ذلك تقيم المستشفيات من أجل تقديم خدمات صحية للمواطنين. ذلك لأن الإنسان كمورد بشري للإنتاج يمثل ركيزةً أساسية من ركائز التقدم الذي تنشده جميع الدول. لهذا نرى كيف تحاول الدولة جاهدة توفير العناصـــر الأساسية التي تضمن تقديم خدمة طبية متميزة لجميع مواطنيها. ويتمثل ذلك في توفير التجهيزات والمعدات الطبية اللازمة إلى جانب الأطباء وهيئة التمريض وغير ذلك من الأجهزة الإدارية والفنية اللازمة.

ومما لا شك فيه أن الزيادة المضطردة في أعداد السكان عاما بعد آخر يُلقى على كاهل الحكومات مسئوليةً كبيرةً من حيث تطوير الخدمة الصحية كماً وكيفاً كي تواكب الزيادة الكبيرة في أعداد طالبي هذه الخدمة.

هذا والإحصاءات الصحية ماهي إلا عبارة عن البيانات الخاصة بكل ماله علاقة بصحة الإنسان والعوامل المؤثرة عليها. وهذه الإحصاءات هي التي تمكِّننا من استخراج المعدلات الصحية والتي تلقى بدورها الضوء على مواطن القوة والضعف في عملية تقديم الخدمة الصحية للمواطنين. وبذلك تمكِّن المسئولين من العمل دائماً على تحسين تلك الخدمة وتطويرها بما يواكب منطق التطور الحادث في جميع المجالات في عصرنا الحاضر.

كما تأتي أهمية الإحصاءات الصحية في كونها الأداة المُثلى والتي لا غنى عنها في رفع خدمات المستشفيات. حيث أن الأمر يتعلق بتوظيف موارد متنوعة لخدمة أهداف متكاملة ومتشعبة بأقل تكلفة ممكنة وبأعلى كفاءة. ومن ثم فلابد من تحليل تلك الموارد والأهداف. ولتحقيق ذلك يلزم ما يلي:

- التعرف على أنواع الخِدمات التي يطلبها المرضى، ويتطلب ذلك معرفة أنواع الأمراض الأكثر انتشاراً والتعرف على حاملي تلك الأمراض.
- الوقوف على مستوى الإمكانات المتاحة من خِدمات المستشفيات لمقابلة احتياجات المرضى. ويشمل ذلك من بين ما يشمل القوى العاملة والأسِرَّة.
- توظيف الإمكانات لمواجهة الخدمات الصحية المطلوبة.

مما تقدم وفي دراستنا للإحصاءات الصحية سوف يقتصر تناولنا لها على مجاليْن هما:

- إحصاءات الأمراض.
- إحصاءات الخِدمات الصحية.

وسوف نتناول كل نوع من هذه الإحصاءات بشيء من التفصيل وذلك على النحو الموضَّح فيما يلي:

أولاً: إحصاءات الأمراض

تؤثر الأمراض بشكل خطير على صحة أفراد المجتمع لا سيما تلك الأمراض الخطيرة والتي تتميز بالانتشار وسرعة العدوى. لذلك أصبح من المهم الحصول على بيانات تلك الأمراض واستخراج المعدلات المتعلقة بها وذلك بهدف الوصول إلى مستوى انتشارها بين أفراد المجتمع والأضرار الناجمة عنها حتى يمكن تلاشيها. ومن المعدلات التي نتناولها في هذا الشأن ما يلي:

١ــ معدل الإصابة بالمرض:

ويشير هذا المعدل إلى عدد الأشخاص المصابين بالمرض خلال سنة ما لكل ألف شخص من السكان في هذه السنة. أي أن:

$$\text{معدل الإصابة بالمرض} = \frac{\text{عدد الأشخاص الذين أصيبوا بالمرض خلال السنة}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$$

٢ــ معدل وفيات المرض:

ويأخذ هذا المعدل الصورة التالية:

$$\text{معدل وفيات المرض} = \frac{\text{عدد المتوفين بسبب المرض}}{\text{عدد المصابين بهذا المرض}} \times ١٠٠٠$$

وهو يبيِّن عدد المتوفين نتيجة للإصابة بهذا المرض لكل ألف شخص مصاب به.

٣ــ معدل حاملي المرض:

حاملو المرض هم الأشخاص الذين ينقلون المرض دون أن تظهر عليهم أعراض المرض العادية. ويُعرّف هذا المعدل كما يلي:

$$\text{معدل حاملي المرض} = \frac{\text{عدد حاملي المرض}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$$

ويشير هذا المعدل إلى عدد حاملي المرض لكل ألف من السكان.

٤ــ معدل انتشار المرض:

ويختلف هذا المعدل عن معدل الإصابة بالمرض في أن هذا المعدل يأخذ في الاعتبار عدد الإصابات القديمة بالمرض بالإضافة إلى عدد الإصابات الجديدة والتي يقتصر عليها معدل الإصابة بالمرض. أي أن:

$$\text{معدل انتشار المرض} = \frac{\text{عدد الإصابات (الجديدة + القديمة)}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$$

وهو يوضح عدد المصابين بالمرض لكل ألف من السكان.

٥ــ معدل المعرضين للخطر:

المعرضون للخطر هم سكان منطقة حدث بها المرض فتم الحجر عليهم خشية انتشاره. ويأخذ معدل المعرضين للخطر الصورة التالية:

$$\text{معدل المعرضين للخطر} = \frac{\text{عدد المعرضين للخطر}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$$

ويعطي هذا المعدل عدد المعرضين لخطر الإصابة بالمرض وذلك لكل ألف من السكان.

مثال (١٤):

إذا افترضنا أن مرض الكوليرا قد انتشر في إحدى السنوات في بلد ما فكان عدد المصابين به هو ٦٥٠٠ شخصاً وكان تقدير عدد السكان في منتصف هذه السنة هو ٢٥ مليون نسمة. فإذا مات ٥٠٠٠ شخص ممن أصيبوا بالمرض، وأنه تم الحجر على سكان منطقة يبلغ تعداد سكانها ١,٥ مليون شخص. وبإجراء الفحوصات الطبية للسكان تبيَّن أن هناك ٧٥٠٠ شخص حاملين للمرض.

المطلـــوب:

حساب المعدلات التالية:

أ- معدل الإصابة بالمرض **ب-** معدل وفيات المرض

ج- معدل حاملي المرض **د-** معدل المعرضين للخطر

هـ- معدل انتشار المرض، وذلك بافتراض أن هناك ٢٠٠ حالة إصابة قديمة بهذا المرض.

الحـــــل:

أ- معدل الإصابة بمرض الكوليرا = $\frac{\text{عدد المصابين بالمرض خلال السنة}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$

$= \frac{٦٥٠٠}{٢٥٠٠٠٠٠٠} \times ١٠٠٠ = ٠,٢٦$ في الألف

أي أن معدل الإصابة بمرض الكوليرا = ٢٦ لكل ١٠٠ ألف.

وهذا يعنى أنه في كل ١٠٠ ألف من السكان أُصيب ٢٦ بمرض الكوليرا في هذه السنة.

ب- معدل وفيات الكوليرا = $\frac{\text{عدد المتوفين بسبب مرض الكوليرا}}{\text{عدد المصابين بهذا المرض}} \times ١٠٠٠$

$= \frac{٥٠٠٠}{٦٥٠٠} \times ١٠٠٠ \approx ٧٦٩$ في الألف

أي أنه من بين كل ألف شخص أصيب بمرض الكوليرا في هذا العام تُوفِّي ٧٦٩ شخصاً من جرّاء إصابتهم به.

ج- عدد حاملي مرض الكوليرا = $\frac{\text{عدد حاملي مرض الكوليرا}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$

$= \frac{٧٥٠٠}{٢٥٠٠٠٠٠٠} \times ١٠٠٠ = ٠,٣$ في الألف

أي أن معدل حاملي مرض الكوليرا هو ٣ لكل ١٠ ألاف. وهذا يعنى أنه من بين كل ١٠ ألاف من السكان هناك ثلاثة أشخاص حاملين لمرض الكوليرا ولا تظهر عليهم أعراضه العادية.

$$\text{د- عدد المعرضين للخطر} = \frac{\text{عدد المعرضين للخطر}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$$

$$= \frac{١٥٠٠٠٠٠}{٢٥٠٠٠٠٠٠} \times ١٠٠٠ = ٦٠ \text{ في الألف}$$

وهذا يعني أن هناك ٦٠ شخصاً معرضون للإصابة بمرض الكوليرا من بين كلُ ألف من السكان.

$$\text{هـ- معدل وفيات الكوليرا} = \frac{\text{عدد الإصابات (الجديدة + القديمة)}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$$

$$= \frac{٦٥٠٠ + ٢٠٠}{٢٥٠٠٠٠٠٠} \times ١٠٠٠ = ٠,٢٧ \text{ في الالف}$$

أي أن معدل انتشارا المرض هو ٢٧ لكل ١٠٠ ألف من السكان. وهذا يعنى أنه من بين كل ١٠٠ ألف من السكان هناك ٢٧ شخصاً مصاباً بمرض الكوليرا.

ثانياً: إحصاءات الخِدمات الصحية

الخِدمات الصحية هي تلك الخِدمات التي تقدمها المؤسسات الصحية كالمستشفيات لأفراد المجتمع بدون مقابل أو بمقابل. ولا شك أن التمتع بالخدمة الصحية يتم من خلال دخول المريض إلى المستشفى وإقامته به إذا تطلب الأمر ذلك، واستعماله للسرير في حالة الإقامة، وتلقيه الرعاية الطبية من قِبَل الأطباء والهيئة التمريضية، وأخيراً خروجه من المستشفى. وبذلك يمكننا أن نستنتج أنه في تناولنا للإحصاءات المتعلقة باستعمال الخدمات الصحية سوف نكون في حاجة إلى الوقوف على مستوى المعدلات الصحية في هذا الشأن والتي نخص منها ما يلي:

١ـ نسبة إشغال الأسرة في فترة زمنية:

وتعَّرف هذه النسبة كما يلي:

$$\textbf{نسبة إشغال الأسِرَّة} = \frac{\textbf{عدد أيام التمريض خلال فترة زمنية معينة}}{\textbf{مجموع أيام التمريض خلال نفس الفترة}} \times ١٠٠$$

وهنا تلزم الإشارة إلى أن المقصود بأيام التمريض في البسط هو عدد أيام التمريض الفعلية خلال الفترة الزمنية. وأما أيام التمريض في المقام فالمقصود بها عدد أيام التمريض المتاحة لدى المستشفى خلال نفس الفترة.

وتعطي هذه النسبة عدد الأسرة التي تم إشغالها من قِبَل المرضى لكل ١٠٠ سرير من أسِرّة المستشفى وذلك خلال فترة زمنية معينة.

هذا ويتم حساب مجموع أيام التمريض خلال فترة زمنية معينة (الموجود في المقام) عن طريق ضرب عدد الأسِرّة المتاحة في المستشفى والمُعدَّة لاستقبال المرضى في عدد أيام الفترة الزمنية. أي أن:

مجموع أيام التمريض = عدد الأسِرَّة × عدد أيام الفترة الزمنية

والمقصود بيوم التمريض أو اليوم العلاجي هو اليوم الذي يقضيه المريض بالمستشفى ويتلقى خلاله خدمة تمريضية. هذا ولا يُحسب يوم خروج المريض من المستشفى بينما يعتبر يوم الدخول يوماً علاجياً.

وتفيد هذه النسبة في التعرف على ما إذا كان هناك نقص أم لا في عدد الأسِرَّة المتاحة في المستشفيات خلال فترة زمنية محددة، إذ أنها تعكس العلاقة ما بين عدد الأسِرَّة التي تم إشغالها بالمرضى وبين عدد الأسِرّة المتاحة في المستشفى خلال تلك الفترة الزمنية.

٢ـ نسبة إشغال الأسِرَّة اليومي:

وهذه النسبة هي حالة خاصة من النسبة السابقة وهي الحالة التي تكون فيها الفترة الزمنية يوماً واحداً. أي أن:

$$\textbf{نسبة إشغال الأسِرّة اليومي} = \frac{\textbf{عدد أيام التمريض خلال يوم معين}}{\textbf{مجموع الأسِرّة خلال نفس اليوم}} \times ١٠٠$$

٣ـ معدل الدخول اليومي (التنويم):

وهو يشير إلى عدد حالات دخول المرضى (التنويم) في اليوم الواحد في المتوسط. أي أن:

$$\textbf{معدل الدخول اليومي} = \frac{\textbf{مجموع عدد المرضى الداخلين خلال فترة معينة}}{\textbf{عدد أيام تلك الفترة}}$$

٤ـ معدل الخروج اليومي:

وهذا المعدل يأخذ الصورة التالية:

$$\textbf{معدل الخروج اليومي} = \frac{\textbf{مجموع عدد حالات الخروج خلال فترة معينة}}{\textbf{عدد أيام تلك الفترة}}$$

أي أنه يبيِّن عدد الخارجين من المستشفى في اليوم الواحد في المتوسط.

٥ـ متوسط مدة الإقامة (البقاء):

ويشير هذا المتوسط إلى عدد الأيام في المتوسط التي كان المريض فيها منوَّماً داخل المستشفى خلال فترة زمنية. أي أن:

$$\text{متوسط مدة الإقامة (البقاء)} = \frac{\text{مجموع أيام تمريض المرضى الذين خرجوا خلال فترة معينة (مجموع مدد البقاء)}}{\text{عدد المرضى الذين خرجوا خلال نفس الفترة}}$$

ويختلف هذا المتوسط باختلاف نوع المستشفى وباختلاف الأقسام داخل المستشفى الواحد. حيث نجد أن هذا المتوسط يميل إلى الارتفاع في حالة الأمراض الصدرية والنفسية على سبيل المثال. وعند حساب متوسط مدة الإقامة (البقاء) في أحد المستشفيات فإنه يُؤخذ في الاعتبار جميع حالات الخروج خلال الفترة الزمنية موضع الاعتبار حتى ولو كانت أيامهم العلاجية ممتدة من فترات سابقة، بينما تُهمل أيام التمريض خلال الفترة الزمنية والتي تمتد لفترات لاحقة.

٦ـ معدل دوران السرير:

ويشير هذا المعدل إلى عدد المرضى (في المتوسط) الذين شغلوا سريراً واحداً خلال فترة زمنية معينة. ويعرّف هذا المعدل كما يلي:

$$\text{معدل دوران السرير} = \frac{\text{مجموع حالات الخروج (بما فيهم الوفيات) خلال فترة معينة}}{\text{عدد الأسِرّة خلال تلك الفترة}}$$

٧ـ فترة خلو السرير:

وهي المدة بالأيام (في المتوسط) والتي يكون فيها السرير في المستشفى شاغراً ما بين ترك مريض له وشغله بواسطة مريض آخر. أي هي متوسط عدد الأيام التي تتخلل الفترة ما بين خروج المريض الذي يشغل السرير وبين دخول مريض آخر فيه. يمكن تحديد هذه الفترة باستخدام الصيغة التالية:

$$\text{فترة خلو السرير} = \frac{\text{عدد الأسِرّة} \times \text{عدد أيام الفترة} - \text{(مجموع مدد البقاء لمرضى الخروج بما فيهم الوفيات خلال الفترة)}}{\text{عدد حالات الخروج بما فيهم الوفيات خلال نفس الفترة}}$$

وبطبيعة الحال فإن هذه الفترة تقل كلما كان معدل دوران السرير كبيراً، وكذلك كلما كان متوسط مدة البقاء كبيراً، والعكس صحيح. كما أنها تناسب عكسياً مع نسبة إشغال الأسرّة خلال نفس الفترة.

٨ـ النسبة المئوية لخلو السرير:

من الطبيعي أن تكون هذه النسبة هي المكمِّلة لنسبة إشغال السرير. أي أن:

النسبة المئوية لخلو السرير خلال فترة معينة

= ١٠٠٪ ـ نسبة إشغال السرير خلال نفس الفترة

مثال (١٦):

إذا توافرت البيانات التالية عن أحد المستشفيات في بلد ما في عام ٢٠٢٠م:

ـ عدد الأسِرّة بالمستشفى = ٤٥٠ سريراً.

ـ عدد المرضى الخارجين من المستشفى خلال العام = ٥٤٠٠ مريضاً.

ـ متوسط مدة البقاء في هذا المستشفى لهذا العام = ٥ أيام.

المطلـــوب:

إيجاد المؤشرات الصحية التالية:

أ- معدل دوران السرير　　**ب-** معدل الخروج اليومي

ج- نسبة إشغال الأسّرة　　**د-** النسبة المئوية لخلو السرير

هـ- فترة خلو السرير

الحـــل:

أ- معدل دوران السرير $= \dfrac{\text{عدد حالات الخروج خلال عام ٢٠٢٠}}{\text{عدد الأسِرّة في هذا العام}}$

$= \dfrac{٥٤٠٠}{٤٥٠} =$ ١٢ مريضاً

أي أنه خلال عام ٢٠٢٠م تم إشغال السرير بواسطة ١٢ مريضاً في المتوسط.

ب- معدل الخروج اليومي $= \dfrac{\text{عدد حالات الخروج خلال عام ٢٠٢٠}}{\text{عدد أيام عام ٢٠٢٠}}$

$= \dfrac{٥٤٠٠}{٣٦٦} =$ ١٤,٧٥ $\approx$ ١٥ مريضاً

أي أنه في كل يوم من أيام عام ٢٠٢٠م كان يخرج من المستشفى ١٥ مريضاً في المتوسط.

ج- نسبة إشغال الأسِرّة $= \dfrac{\text{عدد أيام التمريض خلال عام ٢٠٢٠}}{\text{مجموع أيام التمريض خلال نفس العام}} \times$ ١٠٠

$$= \frac{\text{عدد أيام التمريض خلال عام ٢٠٢٠}}{\text{عدد الأسِرّة} \times \text{عدد أيام عام ٢٠٢٠}} \times ١٠٠$$

$$= \frac{\text{عدد أيام التمريض خلال عام ٢٠٢٠}}{٤٥٠ \times ٣٦٦} \times ١٠٠$$

وحيث أن:

$$\text{متوسط مدة البقاء} = \frac{\text{مجموع أيام التمريض خلال عام ٢٠٢٠}}{\text{عدد المرضى الخارجين خلال هذا العام}}$$

فإننا نجد أن:

$$\text{متوسط مدة البقاء} = \frac{\text{مجموع أيام التمريض خلال عام ٢٠٢٠}}{٥٤٠٠} = ٥$$

أي أن: مجموع أيام التمريض خلال عام ٢٠٢٠ = ٥ × ٥٤٠٠ = ٢٧٠٠٠ يوم تمريض

$$\therefore \text{نسبة إشغال الأسرة} = \frac{٢٧٠٠٠}{٤٥٠ \times ٣٦٦} \times ١٠٠ = ١٦,٤\%$$

وهذا يعني أنه خلال عام ٢٠٢٠م تم إشغال ١٦,٤٪ فقط من إجمالي عدد الأسِرّة في هذا المستشفى. وهي بلا شك نسبة ضعيفة.

د- النسبة المئوية لخلو السرير = ١٠٠٪ – نسبة إشغال الأسِرّة

= ١٠٠٪ – ١٦,٤٪ = ٨٣,٦٪

أي أنه يمكن القول بأن ٨٣,٦٪ من أسِرّة هذا المستشفى كانت شاغرة خلال عام ٢٠٢٠م.

$$\text{هـ - فترة خلو السرير} = \frac{\text{عدد الأسِرّة} \times \text{عدد أيام عام ٢٠٢٠ - مجموع مدد البقاء (أيام التمريض)}}{\text{عدد حالات الخروج خلال عام ٢٠٢٠}}$$

$$= \frac{٤٥٠ \times ٣٦٦ - ٢٧٠٠٠}{٥٤٠٠} = ٢٥,٥ \text{ يوماً}$$

وهذا يعني أن المدة ما بين ترك المريض للسرير الذي يشغله وبين دخول مريض آخر في نفس السرير هي ٢٥,٥ يوماً في المتوسط.

مثال (١٧):

في المثال السابق، كيف يمكن الربط بين النتيجة في (أ) والنتيجة في (هـ) وصولاً إلى النتيجة في (ج).

الحـــــل:

حيث أن فترة خلو السرير هي ٢٥,٥ يوماً (النتيجة في هـ) وأن السرير قد شغله ١٢ مريضاً خلال العام (النتيجة في أ). فإنه يمكن استنتاج أن السرير الواحد كان شاغراً خلال العام لمدة (٢٥,٥ × ١٢ = ٣٠٦ يوم). وحيث أن عام ٢٠٢٠م يتكون من ٣٦٦ يوماً، فإن:

$$\text{نسبة إشغال السرير} = \frac{٣٦٦ - ٣٠٦}{٣٦٦} \times ١٠٠ = ١٦,٤٪$$

وهي نفس النتيجة التي توصلنا إليها في (ج).

٩ــ معدل الأطباء لعدد السكان:

ويبيِّن هذا المعدل عدد الأطباء (في المتوسط) لكل ألف من السكان ويعرّف كما يلي:

$$\textbf{معدل الأطباء لعدد السكان} = \frac{\textbf{عدد الأطباء خلال العام}}{\textbf{عدد السكان في منتصف العام}} \times ١٠٠٠$$

ويفيد هذا المعدل في التعرف على مدى الاستفادة من الخدمات الطبية. فكلما قل هذا المعدل كان ذلك مؤشراً على أن الطبيب لا يكون لديه الوقت الكافي لكل مريض والعكس صحيح.

١٠ــ معدل التمريض لعدد السكان:

وهو شبيه بالمعدل السابق إذ أنه يشير إلى عدد أفراد الهيئة التمريضية (في المتوسط) لكل ألف من السكان. ويأخذ الصورة التالية:

$$\textbf{معدل التمريض لعدد السكان} = \frac{\textbf{عدد أفراد الهيئة التمريضية خلال العام}}{\textbf{عدد السكان في منتصف العام}} \times ١٠٠٠$$

١١ــ معدل الأسِرّة للتمريض:

ويوضح هذا المعدل عدد الأسِرّة (في المتوسط) لكل فرد من أفراد الهيئة التمريضية. أي أن:

$$\textbf{معدل الأسِرّة للتمريض} = \frac{\textbf{عدد الأسِرّة في المستشفى}}{\textbf{عدد أفراد الهيئة التمريضية في المستشفى}}$$

وكلما زاد هذا المعدل دل ذلك توافر الخدمة التمريضية للمرضى والعكس صحيح.

هذا ويفضل استخراج هذا المعدل لكل قسم على حدة داخل المستشفى إذ أن مدى الحاجة إلى الخدمات التمريضية يختلف من قسم لآخر داخل المستشفى الواحد.

مثال (١٨):

إذا توافرت البيانات التالية عن إحدى المدن في عام ٢٠١٨م:

- عدد السكان = ٥٠٠٠٠ نسمة
- عدد الأسِرّة بالمستشفى الموجود بها = ٤٥٠ سريراً
- عدد أفراد الهيئة التمريضية بالمستشفى = ٣٥٠ ممرضة وممرضاً
- عدد الأطباء بالمستشفى = ٢٥٠ طبيباً

المطلوب:

احسب المعدلات التالية:

أ- معدل الأسِرّة لعدد السكان **ب-** معدل التمريض لعدد السكان

ج- معدل الأسِرّة للتمريض

الحـــــل:

أ- معدل الأطباء لعدد السكان = $\frac{\text{عدد الأطباء خلال عام ٢٠١٨}}{\text{عدد السكان في عام ٢٠١٨}} \times ١٠٠٠$

$= \frac{٢٥٠}{٥٠٠٠٠} \times ١٠٠٠ = ٥$ في الألف

أي أن هناك خمسة أطباء لكل ألف من السكان في عام ٢٠١٨م.

ب- معدل التمريض لعدد السكان = $\frac{\text{عدد أفراد الهيئة التمريضية في عام ٢٠١٨}}{\text{عدد السكان في نفس العام}} \times ١٠٠٠$

$= \frac{٣٥٠}{٥٠٠٠٠} \times ١٠٠٠ = ٧$ في الألف

أي أن هناك ٧ أفراد من الهيئة التمريضية لكل ألف من السكان في عام ٢٠١٨م.

ج- معدل الأسِرّة للتمريض = $\frac{\text{عدد الأسِرّة في المستشفى}}{\text{عدد أفراد الهيئة التمريضية في المستشفى}}$

$= \frac{٤٥٠}{٣٥٠} = ١,٣$ سرير / ممرض أو ممرضة

أي أن كل ممرضة (أو ممرض) تقوم بتقديم الخدمة التمريضية لمرضى ١,٣ سريرا في المتوسط. وبمعنى آخر فإن كل ١٠ ممرضات وممرضين يقومون بتقديم الخدمة التمريضية لمرضى ١٣ سريراً.

وسوف نكتفي بهذا القدر من الإحصاءات الصحية. إذ أن التعرض للمزيد من التفصيلات في هذا الشأن قد ينحو بهذا الكتاب بعيداً عن مضمون الموضوعات التي عادة ما تشملها كتب الإحصاء التطبيقي. فلا زال هناك العديد من المؤشرات الصحية والتي تُضفي تغطيتها صيغةً يغلب عليها الطابع الصحي أكثر من كونها إحصائية بالمفهوم البحت لعلم الإحصاء المُراد تناول بعض جوانبه في هذا الكتاب.

تمارين

الفصل الخامس

١- إذا توافرت البيانات التالية عن نتائج تعداد تم إجراؤه في عام ٢٠١٦م في إحدى الدول:

- عدد السكان = ٤٠ مليون
- عدد المواليد أحياء خلال عام ٢٠١٦ = ١,١ مليون مولود حي.
- نسبة النساء اللاتي في سن الحمل = ٣٢٪ من عدد السكان.
- نسبة المتزوجات اللاتي في سن الحمل = ٧٥٪ من عدد النساء اللاتي في سن الحمل.

المطلوب:

حساب المعدلات التالية:

أ- معدل المواليد الخام **ب-** معدل الخصوبة العام **ج-** معدل التوالد العام

مع شرح ما يعنيه كل معدل مستعيناً بالنتائج التي حصلت عليها.

٢- الجدول التالي يوضح معدلات الخصوبة التفصيلية حسب فئات عمر الأمهات:

فئات العمـــــر	١٥-	٢٠-	٢٥-	٣٥-	٤٥- ٥٠
معدل الخصوبة	٨٠	١٩٠	٢٢٥	١٤٠	٤٠

احسب معدل الخصوبة الكلي شارحاً مدلول النتيجة التي حصلت عليها.

٣- الجدول التالي يوضح بعض البيانات عن أحد البلدان خلال سنة ما:

فئات سن الحمل	١٥-	٢٠-	٣٠-	٣٥-	٤٥- ٥٠
عدد المواليد أحياء (بالألف)	٦٠	١٧٥	٩٠	٥٠	١٥
عدد المواليد الإناث (بالألف)	٢٥	٨٠	٤٠	٢٠	٥
عدد النساء (بالألف)	٥٧٠	١١٥٠	٦٥٠	٦٢٥	٩٠٠
عدد المتزوجات (بالألف)	٤٧٥	٩٥٠	٥٠٠	٤٠٠	٨٠٠

المطــلوب:

أولاً: احسب المعدلات التالية

أ- معدل الخصوبة الكلي **ب-** معدل التوالد العام

ج- معدل التكاثر (التناسلي) الإجمالي

مع شرح النتائج التي حصلت عليها

ثانياً: قارن بين المعدلات الثلاثة من حيث أفضليتها كمقياس للخصوبة.

٤- فيما يلي البيانات الخاصة بالتوزيع العمري لسكان إحدى المدن (ذكوراً وإناثاً) وكذلك عدد المواليد وعدد الوفيات حسب الجنس وذلك في سنة ما:

فئات الأعمار	عدد السكان (بالألف)		عدد المواليد حسب عمر الأم		عدد الوفيات	
	ذكور	إناث	ذكور	إناث	ذكور	إناث
صفر –	٦	٩	–	–	٢٠٠	٢٥٠
١ –	٢٦	٢٥	–	–	٥٠٠	٥٥٠
٥ –	٢٤	٢٦	–	–	١٢٠	١٣٠
١٠ –	٢٦	٢٨	–	–	٧٠	٨٠
١٥ –	٢٤	٢٥	٨٠	٧٠	١٢٥	١٢٥
٢٠ –	٣٢	٣٤	٣٠٥٠	٣٠٧٠	١٨٠	٢٢٠
٣٠ –	٢٦	٢٨	٢١٢٠	٢١٠٠	٢٠٠	٢٥٠
٤٠ –	٢٢	٢٧	١٣٥٠	١٣٠٠	٢٣٠	٢٧٠
٥٠ – فأكثر	١٦	١٨	–	–	٥٠٠	٤٠٠
المجموع	٢٠٠	٢٢٠	٦٦٠٠	٦٥٥٠	٢١٢٥	٢٢٧٥

المطلــوب: حساب المعدلات التالية:

- معدل المواليد الخام
- معدل الخصوبة العام
- معدل الخصوبة لفئة السن (٢٠–٣٠) في المتوسط
- معدل الخصوبة الأنثوية لفئة السن (١٠ – ١٥)
- معدل الخصوبة الكلى
- معدل التوالد العام
- معدل التكاثر الإجمالي: بافتراض أن المتزوجات يمثلن ٩٥٪ من إجمالي عدد النساء في كل فئة من فئات سن الحمل.
- معدل الوفيات الخام
- معدل الوفيات في فئة العمر (٣٠ – ٤٠)
- معدل وفيات كل من الذكور والإناث في فئة العمر (١٥ –٢٠) مع المقارنة بينهما
- معدل وفيات الأطفال الرضع
- معدل وفيات الأمومة، وذلك بافتراض أن عدد وفيات النساء أثناء الحمل والولادة كان ٤٠ حالة وفاة.

٥- الجدول التالي يوضح التوزيع العمري لسكان احدى المدن وعدد المواليد وعدد الوفيات لهذه المدينة في سنة ما:

فئات الأعمار	عدد السكان (بالألف)		عدد المواليد أحيــاء	عدد الوفيات
	ذكور	اناث		
صفر –	٥٠	٥١	–	٣١٠٠
٥ –	٤٤	٤١	–	٤٤٠
١٠ –	٤٨	٤٤	–	٢٥٠
١٥ –	٤٠	٣٨	٥٨٠	٣٣٠
٢٠ –	٦٢	٥٤	١٠٥٠٠	٧٤٣
٣٠ –	٥٢	٤٨	٧٤٠٠	٨٥٠
٤٠ –	٣٤	٣٧	٤٦٥٠	٩٤٢
٥٠ فأكثر	٣٠	٢٧	–	١٩٤٥
المجموع	٣٦٠	٣٤٠	٢٣١٣٠	٨٦٠٠

المطلـــوب:

أ- حساب معدلات المواليد ومعدلات الوفيات المختلفة.

ب- حساب معدلات الخصوبة التفصيلية ومعدل الخصوبة الكلى.

٦- إذا توافرت بيانات الزواج والطلاق التالية عن إحدى المدن في سنة ما:

- عدد حالات الزواج التي تمت خلال السنة = ١٢٠٠٠ حالة زواج.
- تقدير عدد السكان في منتصف السنة = ٢ مليون نسمة.
- عدد السكان في سن الزواج = ٨٠٠٠٠٠ شخص.
- أن المتزوجين يمثلون ٦٥٪ من عدد السكان في سن الزواج.
- عدد حالات الطلاق التي تمت خلال السنة = ٢٠٠٠ حالة طلاق.
- عدد المتزوجين خلال السنة يمثل ٥٪ من إجمالي عدد المتزوجين.
- عدد المطلقين خلال السنة = ٤٠٠٠ شخص.

المطلـــوب: احسب المعدلات التالية

أ- معدل الزواج الخام

ب- معدل الزواج المعدل

ج- معدل الطلاق الخام

د- معدل الطلاق المُعَدل

هـ- معدل الطلاق الخاص

مع شرح النتائج التي تحصل عليها.

٧- بافتراض أن مرض الجدري قد انتشر في إحدى الدول في سنة ما فظهر أن عدد الأفراد المصابين بالمرض هو ٢٠٠٠ شخص، وأن عدد الذين ماتوا متأثرين بالمرض هو ١٢٠ شخصاً. فإذا كان تقدير عدد السكان في هذه السنة هو ١٢ مليون نسمة، وأنه عند إجراء الفحص الوقائي تَبيَّن أن هناك ٣٠٠ شخص حاملين للمرض وأن عدد المعرضين للخطر هو ٢٥٠٠٠٠ فرد.

المطلـــوب: حساب المعدلات التالية

أ- معدل الإصابة بالمرض ب- معدل وفيات المرض

ج- معدل حاملي المرض د- معدل المعرضين للخطر

هـ- معدل انتشار المرض، وذلك بافتراض أن هناك ١٠٠ حالة إصابة قديمة بهذا المرض.

٨- انتشر مرض الكوليرا في بلديْن متجاوريْن (أ و ب) في أحد الأعوام فكانت البيانات التالية:

البيـــــان	**البلد (أ)**	**البلد (ب)**
عدد السكـــان (بالآلاف)	٢٠٠	٢٥٠
عدد المصابيــن بالمرض	٥٠٠٠	٧٥٠٠
عدد حاملي المـــــرض	٨٠٠	١٢٠٠

المطلــوب:

أولاً: حساب المعدلات التالية:

أ- معدل الإصابة بالمرض في كل من البلديْن أ، ب.

ب- معدل حاملي المرض في كل من البلديْن أ، ب.

ثانياً: إذا بلغ معدل وفيات المرض في البلد (أ) ٨٠ في الألف وهو يعادل ضعف نفس المعدل في البلد (ب). احسب عدد الوفيات بسبب هذا المرض في كل من البلديْن أ و ب.

٩- إذا افترضنا أن دولة ما تتكون من منطقتيْن (أ و ب) وأن البيانات الصحية التالية كانت متاحة عن كل منطقة على حدة خلال سنة معينة

بيـــــان	**المنطقــة (أ)**	**المنطقة (ب)**	**إجمالي الدولة**
عـدد السكـــان (بالملاييـن)	٣	٤,١٦	٧,١٦
عـدد المستشفيات	١٩٢	٢٠٨	٤٠٠
عـدد الأسرّة بالمستشفيات	٩٦٠٠	١٣٠٠٠	٢٢٦٠٠
عـدد الأطباء	١٨٠٠٠	٢٠٨٠٠	٣٨٨٠٠
عـدد أفراد الهيئة التمريضية	٤٥٠٠٠	٤١٦٠٠	٨٦٦٠٠
عـدد المرضى الخارجين (أحياء)	٥٦٠٠	٥٠٠٠	١٠٦٠٠

المطلـــوب:

أولاً: قارن بين المنطقتيْن (أ و ب) من حيث المؤشرات التالية مفسِّراً مدلول كل مؤشر:

- معدل الأسِرّة لعدد السكان
- معدل الأطباء لعدد السكان
- معدل التمريض لعدد السكان
- متوسط ما يخدمه المستشفى من السكان
- متوسط الأسِرّة لكل مستشفى

ثانياً: مستعيناً بالبيانات التالية بالإضافة إلى البيانات السابقة عن كل من المنطقتيْن:

- معدل الخروج اليومي للمنطقة (أ) = ١٦ مريض / يوم
- معدل الخروج اليومي للمنطقة (ب) = ١٤ مريض / يوم
- نسبة انشغال الأسِرّة في المنطقة (أ) = ١٥٪
- نسبة انشغال الأسِرّة في المنطقة (ب) = ١٠٪

أوجد ما يلي:

أ- إجمالي عدد الخارجين (المتوفين) من المستشفيات خلال العام في الدولة ككل.

ب- متوسط مدة البقاء في المنطقة (أ)

ج- نسبة خلو الأسِرّة في المنطقة (أ)

د- نسبة انشغال الأسِرّة في الدولة ككل

هـ- معدل دوران السرير في المنطقة (ب)

ملحوظة: اعتبر عدد أيام السنة = ٣٦٠ يوماً.

١٠- إذا توافرت البيانات التالية عن أحد المستشفيات في بلد ما في عام ٢٠١٧.

- عدد سكان البلد = ٢٠٠٠٠ نسمة
- عدد الأسِرَّة في المستشفى = ٢٧٠ سريراً
- عدد أيام التمريض = ٣٦٠٠٠ يوم
- عدد حالات الخروج = ٤٨٠٠
- عدد الأطباء = ١٢٠ طبيباً
- عدد أفراد الهيئة التمريضية = ٣٦٠ فرداً

المطلـــوب: إيجاد المؤشرات التالية

– نسبة إشغال الأسِرَّة – النسبة المئوية لخلو السرير – فترة خلو السرير

– معدل الأسِرَّة للسكان – معدل التمريض لعدد السكان – متوسط مدة البقاء

– معدل دوران السرير – معدل الأطباء لعدد السكان

١١- إذا افترضنا أن عدد الأسِرّة أحد المستشفيات خلال شهر ما هو ٥٠٠ سرير. فإذا بلغ عدد أيام التمريض خلال نفس الشهر ٤٠٠٠ يوم، وانه خرج من المستشفى خلال نفس الشهر ٨٠٠ مريض (بما فيهم الوفيات).

المطلـــوب: احسب ما يلي

أ- نسبة إشغال الأسِرّة

ب- متوسط مدة البقاء

مع شرح النتائج التي تحصل عليها.

تطبيقات إحصائية
في
المجالات المختلفة

تطبيقات
الفصل الأول
(الاحتمالات والتوزيعات الاحتمالية)

التطبيق (١)

إذا أُلقيت زهرتا نرد على سطح أملس، احسب الاحتمالات التالية:

أ- الحصول على الرقم ٦ على وجه الزهرة الأولى.

ب- الحصول على رقميْن مجموعهما ٩.

الحـــل:

من المعروف أن عدد الحالات الكلية الممكنة عند رمى زهرتي نرد هو ٣٦ حالة ($٦^٢$ = ٣٦)، وعلى ذلك يمكن حساب الاحتمالات المختلفة المطلوبة كما يلي:

أ- هناك ست حالات يظهر فيها الرقم ٦ على وجه الزهرة الأولى، وهذه الحـالات هي:

(٦،١) ، (٦،٢) ، (٦،٣) ، (٦،٤) ، (٦،٥) ، (٦،٦)

أي أن الاحتمال المطلوب = $\frac{٦}{٣٦}$ = $\frac{١}{٦}$

ب- الحالات التي يكون فيها مجموع الرقميْن مساوياً ٩ هي:

(٣،٦)، (٤،٥)، (٥،٤)، (٦،٣)، حيث يشير الرقم الأول إلى وجه الزهرة الأولى بينما يشير الرقم الثاني إلى وجه الزهرة الثانية.

∴ الاحتمال المطلوب = $\frac{٤}{٣٦}$ = $\frac{١}{٩}$

التطبيق (٢)

صندوق يحتوي على ٤ كرات بيضاء، ٦ سوداء وآخر يحتوي على ٣ كرات بيضاء، ٥ سوداء. أُختير أحد الصندوقيْن عشوائياً ثم سُحبت منه كرة.

المطلــوب:

أ- احتمال أن تكون الكرة المسحوبة بيضاء.

ب- إذا كانت الكرة المسحوبة سوداء، ما هو احتمال أن تكون قد سُحِبت من الصندوق الأول.

الحـــل:

سوف نسمي الأحداث على النحو التالي:

الصندوق الأول أ ، الصندوق الثاني ب ،

الكرة بيضـــاء ض ، الكرة الســـوداء س

وعليه يمكن حساب الاحتمالات المطلوبة على النحو التالي:

أ- ع(الكرة بيضاء) = ع(ض)

ويمكن أن تكون الكرة بيضاء في حالتيْن هما:
اختيار الصندوق الأول وسحب كرة بيضاء منه، أو:
اختيار الصندوق الثاني وسحب كرة بيضاء منه،

$$\therefore\ ع(ض) = ع(أ)\,ع(ض \mid أ) + ع(ب)\,ع(ض \mid ب)$$

$$= \frac{١}{٢} \times \frac{٤}{١٠} + \frac{١}{٢} \times \frac{٣}{٨} = ٠,٣٨٧٥$$

ب- الاحتمال المطلوب هو:

$$ع(\text{الصندوق الأول} \mid \text{الكرة سوداء}) = ع(أ \mid س)$$

$$ع(أ \mid س) = \frac{ع(أ س)}{ع(س)}$$

$$= \frac{ع(أ)\,ع(أ \mid س)}{ع(أ)\,ع(س \mid أ) + ع(ب)\,ع(س \mid ب)}$$

$$= \frac{\frac{١}{٢} \times \frac{٦}{١٠}}{\frac{١}{٢} \times \frac{٦}{١٠} + \frac{١}{٢} \times \frac{٥}{٨}} = \frac{٠,٣}{٠,٦١٢٥} = ٠,٤٩$$

التطبيق (٣)

إذا كان احتمال نجاح الطالب في مادة الإحصاء هو ٠,٨٥ واحتمال نجاحه في مادة الاقتصاد هو ٠,٧، احسب الاحتمالات التالية:

أ- رسوبه في المادتيْن.

ب- نجاحه في مادة الإحصاء أو مادة الاقتصاد.

ج- نجاحه في مادة واحدة على الأقل.

الحـــل:

سوف نسمي الأحداث على النحو التالي:

النجاح في مادة الإحصاء $أ$ ، النجاح في مادة الاقتصاد $ب$

الرسوب في مادة الإحصاء $\overline{أ}$ ، الرسوب في مادة الاقتصاد $\overline{ب}$

أ- $ع(\text{رسوبه في المادتيْن}) = ع(\overline{أ}\ \overline{ب})$

وحيث أن هذيْن الحدثيْن مستقلان،

∴ ع($\overline{أ}$ $\overline{ب}$) = ع($\overline{أ}$)ع($\overline{ب}$)

= ٠,١٥ × ٠,٣ = ٠,٠٤٥

ب- ع(النجاح في مادة الإحصاء أو مادة الاقتصاد) = ع(ا أو ب)

وحيث أن هذيْن الحدثيْن مستقلان وغير متنافييْن،

ع(ا أو ب) = ع(ا) + ع(ب) – ع(ا ب)

= ع(ا) + ع(ب) – ع(ا)ع(ب)

= ٠,٨٥ + ٠,٧ – ٠,٨٥ × ٠,٧

= ١,٥٥ – ٠,٥٩٥ = ٠,٩٥٥

ج- حل أول:

ع(النجاح في مادة واحدة على الأقل) = ١ – ع(رسوبه في المادتيْن)

= ١ – ٠,٠٤٥ = ٠,٩٥٥

حل آخر:

ع(النجاح في مادة واحدة على الأقل)

= ع($\overline{ب}$ ا) + ع($\overline{أ}$ ب) + ع($\overline{أ}$ $\overline{ب}$)

= ع(ا)ع($\overline{ب}$) + ع($\overline{أ}$)ع(ب) + ع(ا)ع(ب)

= ٠,٨٥ × ٠,٣ + ٠,١٥ × ٠,٧ + ٠,٨٥ × ٠,٧

= ٠,٢٥٥ + ٠,١٠٥ + ٠,٥٩٥ = ٠,٩٥٥

وهي نفس النتيجة التي حصلنا عليها من قبل.

التطبيق (٤)

أظهرت التجارب السابقة أنه من بين كل ١٠٠٠ وحدة منتَجة بواسطة الآلة (ا) في أحد المصانع هناك ٥٠ وحدة معيبة، وأنه من بين كل ٢٠٠٠ وحدة منتَجة بواسطة الآلة ب في نفس المصنع هناك ٢٠٠ وحدة معيبة. فإذا كان الإنتاج اليومي في المصنع يتمثل في ٩٠٠ وحدة منتَجة بواسطة الآلة (ا)، ٦٠٠ وحدة منتَجة بواسطة الآلة (ب)، وأنه تم اختيار وحدة من الإنتاج اليومي للمصنع. احسب الاحتمالات التالية:

أ- أن تكون من إنتاج الآلة (ا).

ب- أن تكون من إنتاج الآلة (ب) ومعيبة.

ج- أن تكون معيبة.

د- إذا كانت هذه الوحدة معيبة، احسب احتمال أن تكون من إنتاج الآلة (ب).

الحــل:

سوف نسمي الأحداث على النحو التالي:

- الوحدة من إنتاج الآلة (أ) أ
- الوحدة من إنتاج الآلة (ب) ب
- الوحدة المعيبة م

أ- ع(أن تكون من إنتاج الآلة أ) = ع(أ)

$$= \frac{٩٠٠}{٩٠٠ + ٦٠٠} = ٠,٦$$

ب- ع(أن تكون من إنتاج الآلة ب ومعيبة) = ع(ب م)

= ع(ب) ع(م | ب)

$$= \frac{٦٠٠}{١٥٠٠} \times \frac{٢٠٠}{٢٠٠٠}$$

$$= ٠,٤ \times ٠,١ = ٠,٠٤$$

ج- الاحتمال المطلوب هو ع(من إنتاج أ| معيبة)

$$ع(أ | م) = \frac{ع(أم)}{ع(م)} = \frac{ع(أ)ع(م|أ)}{ع(أ)ع(م|أ) + ع(ب)ع(م|ب)}$$

$$= \frac{\frac{٩٠٠}{١٥٠٠} \times \frac{٥٠}{١٠٠٠}}{\frac{٩٠٠}{١٥٠٠} \times \frac{٥٠}{١٠٠٠} + \frac{٦٠٠}{١٥٠٠} \times \frac{٢٠٠}{٢٠٠٠}}$$

$$= \frac{٠,٠٣}{٠,٠٣ + ٠,٠٤} = \frac{٠,٠٣}{٠,٠٧} = ٠,٤٣$$

التطبيق (٥)

فيما يلي التوزيع الاحتمالي للمتغير العشوائي المنفصل (س):

س	صفر	١	٢	٣
ع(س)	$\frac{١}{٨}$	$\frac{١}{٤}$	$\frac{١}{٢}$	$\frac{١}{٨}$

المطلــوب: حساب ما يلي:

أ- التوقع والتباين للمتغير س.

ب- ت(س – ١)، ت(س + ٤).

ج- الانحراف المتوسط للمتغير العشوائي س.

الحـــل:

أ- $ت(س) = \sum_{س=٠}^{٣} س_ر \, ع(س_ر)$

$$= صفر \times \frac{١}{٨} + ١ \times \frac{١}{٤} + ٢ \times \frac{١}{٢} + ٣ \times \frac{١}{٨}$$

$$= صفر + ٠,٢٥ + ١ + ٠,٣٧٥ = ١,٦٢٥$$

$تبا(س) = ت\,[س – ت\,(س)]^٢$

$= ت\,(س^٢) - [ت\,(س)]^٢$

وحيث أن:

$ت(س^٢) = \sum س^٢ \, ع(س)$

$$= صفر \times \frac{١}{٨} + (١) \times \frac{١}{٤} + (٢)^٢ \times \frac{١}{٢} + (٣)^٢ \times \frac{١}{٨}$$

$$= صفر + ٠,٢٥ + ٢ + ١,١٢٥ = ٣,٣٧٥$$

$\therefore تبا\,(س) = ٣,٣٧٥ - (١,٦٢٥)^٢ = ٠,٧٣٤$

ب- $ت(س – ١) = ت(س) – ١$

$= ١,٦٢٥ – ١ = ٠,٦٢٥$

$ت(س + ٤) = ت(س) + ٤$

$= ١,٦٢٥ + ٤ = ٥,٦٢٥$

ج- الانحراف المتوسط $= \sum |س - \bar{س}|\, ع(س)$

$$= |صفر – ١,٦٢٥| \times \frac{١}{٨} + |١ – ١,٦٢٥| \times \frac{١}{٤}$$

$$+ |٢ – ١,٦٢٥| \times \frac{١}{٢} + |٣ – ١,٦٢٥| \times \frac{١}{٨}$$

$$= ١,٦٢٥ \times \frac{١}{٨} + ٠,٦٢٥ \times \frac{١}{٤} + ٠,٣٧٥ \times \frac{١}{٢} + ١,٣٧٥ \times \frac{١}{٨}$$

$$= ٠,٧١٩$$

التطبيق (٦)

إذا كانت دالة كثافة الاحتمال للمتغير العشوائي س هي:

$$\text{ف}(\text{س}) = \begin{cases} \frac{٣}{٤}(٢\text{س} - \text{س}^{٢}) & \text{صفر} \leq \text{س} \leq ٢ \\ \text{صفر} & \text{خلاف ذلك} \end{cases}$$

المطلـــوب:

أ- أوجد القيمة المتوقّعة والوســـيط والمنوال لهذا المتغير، ثم علِّق على ما تحصـــل عليه من إجابة.

ب- احسب الاحتمالات التالية:

- $\text{ع}(\text{س} \leq ١)$
- $\text{ع}(\text{س} > ١{,}٥)$
- $\text{ع}(\text{س} \leq ٠{,}٨)$

الحــل:

أ- $\text{ت}(\text{س}) = \int \text{س}\,\text{ف}(\text{س}) = \bar{\text{س}}$

$$= \int_{٠}^{٢} \text{س} \times \frac{٣}{٤}(٢\text{س} - \text{س}^{٢})\, \text{ءس}$$

$$= \frac{٣}{٤}\int_{٠}^{٢} (٢\text{س}^{٢} - \text{س}^{٣})\, \text{ءس}$$

$$= \frac{٣}{٤}\left[\frac{٢\text{س}^{٣}}{٣} - \frac{\text{س}^{٤}}{٤}\right]_{٠}^{٢}$$

$$= \frac{٣}{٤}\left(\frac{٢ \times ٨}{٣} - \frac{١٦}{٤}\right) = ١$$

إيجاد الوسيط: الوسيط هو قيمة و التي تحقق: $\int_{٠}^{\text{و}} \text{ف}(\text{س})\, \text{ءس} = \frac{١}{٢}$

$$\frac{٣}{٤}\int_{٠}^{\text{و}} (٢\text{س} - \text{س}^{٢})\, \text{ءس} = \frac{١}{٢}$$

$$\therefore \frac{٣}{٤}\left[\frac{٢\text{س}^{٢}}{٢} - \frac{\text{س}^{٣}}{٣}\right]_{٠}^{\text{و}} = \frac{١}{٢}$$

$$\frac{٣}{٤}\left[\frac{٢و^{٢}}{٢} - \frac{و^{٣}}{٣}\right] = \frac{١}{٢}$$

$$\frac{٣}{٤}\left(و^{٢} - \frac{و^{٣}}{٣}\right) = \frac{١}{٢}$$

$$\therefore \frac{٣}{٤} و^{٢} - \frac{١}{٤} و^{٣} = \frac{١}{٢}$$

أي أن: $و^{٣} - ٣و^{٢} + ٢ = صفر$

$\therefore (و - ١)(و^{٢} - ٢و - ٢) = صفر$

أي أن أحد جذور هذه المعادلة هو: و = ١

ويمكن التأكد من أن الجذريْن الآخريْن يقعان خارج المدى المعرف به المتغير العشوائي وبالتالي فهما مرفوضان.

أي أن: **الوسيط = ١**

إيجاد المنوال:

لإيجاد قيمة المنوال والتي تجعل دالة كثافة الاحتمال نهاية عظمى نتبع الآتي:

حيث $ف(س) = \frac{٣}{٤}(٢س - س^{٢})$

$$\therefore \frac{د ف(س)}{د س} = \frac{٣}{٢} - \frac{٣}{٢}س = صفر \qquad \therefore س = ١$$

$$\frac{د^{٢} ف(س)}{د س^{٢}} = - \frac{٣}{٢} < صفر$$

أي أنه عند س = ١ هناك نهاية عظمى

$\therefore$ قيمة المنوال = ١

وحيث أن: الوسط الحسابي = الوسيط = المنوال = ١

فإنه يمكننا استنتاج أن دالة كثافة الاحتمال لهذا المتغير هي دالة متماثلة.

ب- $د(س_{١}) = \int_{-\infty}^{س_{١}} \frac{٣}{٤}(٢س - س^{٢}) ءس$

$$= \int_{-\infty}^{س_{١}} ف(س) ءس$$

$$= \frac{٣}{٤}\left[\frac{٢س^٢}{٢} - \frac{س^٣}{٣}\right]_{٠}^{س_١}$$

$$= \left[س_١^٢ - \frac{س_١^٣}{٣}\right] = \frac{س_١^٢}{٤}\left(٣ - س_١\right)$$

أي أن:

$$د(س_١) = ع(س \le س_١) = \frac{س_١^٢}{٤}\left(٣ - س_١\right)$$

$$\therefore\ ع(س \le ١) = د(١) = \frac{١}{٤}(٣ - ١) = \frac{١}{٢}$$

$$ع(س > ١,٥) = ١ - ع(س \le ١,٥) = ١ - د(١,٥)$$

$$= ١ - \frac{(١,٥)^٢}{٤}(٣ - ١,٥) = ٠,١٥٦$$

وكذلك:

$$ع(س \le ٠,٨) = د(٠,٨)$$

$$= \frac{(٠,٨)^٢}{٤}(٣ - ٠,٨) = ٠,٣٥٢$$

التطبيق (٧)

إذا كانت نسبة المعيب في إنتاج أحد المصانع هي ١٠٪، وأُخذت عينة مكونة من خمس وحدات من إنتاج هذا المصنع. احسب الاحتمالات التالية:

أ- ألا يكون بالعينة أية وحدات معيبة.

ب- أن يكون بالعينة وحدتان معيبتان.

ج- ألا يزيد عدد الوحدات السليمة عن أربع وحدات.

د- احسب العدد المتوقَّع لعدد الوحدات المعيبة بين مفردات العينة.

الحـــل:

حيث أن احتمال وجود وحدة معيبة بين مفردات العينة هو احتمال ثابت لجميع الوحدات ويساوى ٠,١ فإن التوزيع الملائم هنا هو توزيع ثنائي الحدّين

أ- حيث أن: $ن = ٥$ ، $ر =$ صفر ، $ل = ٠,١$ ، $ك = ٠,٩$

$$ع(ر) = {}^{ن}ق_{ر}\ ل^{ر}\ ك^{ن-ر}$$

$\therefore$ ع(صفر) = ^{٥}ق$_{.}$(٠,١)^{صفر}(٠,٩)^{٥} = ٠,٥٩

ب- ن = ٥ ، ر = ٢ ، ل = ٠,١ ، ك = ٠,٩

ع(٢) = ^{٥}ق$_{٢}$(٠,١)^{٢}(٠,٩)^{٣}

= ١٠(٠,٠١)(٠,٧٢٩) = ٠,٠٧٣

ج- ظهور وحدة سليمة هنا هي حالة النجاح بينما ظهور وحدة معيبة هي حالة الفشل وعلى ذلك نجد أن:

ن = ٥ ، ر = ٠ ، ١، ٢ ، ٣ ، ٤ ، ل = ٠,٩ ، ك = ٠,١

ع(ر $\leq$ ٤) = ١ – ع(ر > ٤) = ١ – ع(ر = ٥)

= ١ – ^{٥}ق$_{٥}$(٠,٩)^{٥}(٠,١)^{صفر}

= ١ – ٠,٥٩ = ٠,٤١

د- العدد المتوقَّع للوحدات المعيبة = $\sum_{ر=٠}^{٥}$ ر ع(ر)

حيث أن ذلك يتطلب حسابات كثيرة فإنه يمكن الحصول على العدد المتوقَّع للوحدات المعيبة بطريقة أيسر على النحو التالي:

حيث أن الوحدات المعيبة تتبع التوزيع ثنائي الحديْن وأن القيمة المتوقَّعة لهذا التوزيع تساوى **ن ل** فإننا نحصل على:

العدد المتوقَّع للوحدات المعيبة = **ن ل** = ٥ × ٠,١ = ٠,٥

التطبيق (٨)

أُلقيت قطعة عمله أربع مرات، **المطلوب:**

أ- أوجد دالة كثافة الاحتمال للمتغير العشوائي س والذي يمثل ظهــور الوجه الذي عليه الشعار.

ب- احسب احتمال ظهور الشعار مرة واحدة على الأقل.

ج- احسب القيمة المتوقَّعة للمتغير س.

الحـــل:

احتمال ظهور الوجه الذي عليه الشعار يساوى $\frac{١}{٢}$ دائماً ونجد بالتالي أن التوزيع الملائم في هذه الحالة هو التوزيع ثنائي الحديْن.

أ- $ع(ر) = ن ق_ر \; ل^ر \; ك^{ن-ر}$

$\therefore \; ع(ر) = {}^{٤}ق_ر \left(\frac{١}{٢}\right)^ر \left(\frac{١}{٢}\right)^{٤-ر}$

أي أن دالة كثافة الاحتمال المطلوبة هي:

$ع(ر) = {}^{٤}ق_ر \left(\frac{١}{٢}\right)^{٤}$

ب- $ع(ر \leq ١) = ١ - ع(ر = \text{صفر})$

$= ١ - {}^{٤}ق. \left(\frac{١}{٢}\right)^{٤}$

$= ١ - ٠{,}٠٦٢٥ = ٠{,}٩٣٧٥$

ج- حيث أن القيمة المتوقعة للتوزيع ثنائي الحديْن = ن ل

$\therefore$ القيمة المتوقعة لعدد مرات ظهور الوجه الذي عليه الشعار

$= ٤\left(\frac{١}{٢}\right) = ٢$

التطبيق (٩)

إذا أظهرت الخبرة الماضية أن ١٪ من الأطفال بمدينة ما يتعرضون للإصابة بمرض معين كل عام. فإذا أخذنا عينة عشوائية مكونة من ١٠٠ طفل.

المطلــوب:

أ- احتمال أن خمسة أطفال سوف يصابون بهذا المرض.

ب- القيمة المتوقَّعة لعدد الأطفال الذين يصابون بالمرض خلال العام.

الحــــل:

حيث أن: ن > ٣٠ ، ن ل = ١٠٠(٠,٠١) = ١ وهي أقل من ٥ فإن التوزيع الملائم هنا هو توزيع بواسون والذي يأخذ الصورة التالية:

$$ع(ر) = \frac{هـ^{-\lambda} \lambda^{ر}}{ر!} \quad ، \quad ر = ٠، ١، \ldots، \infty$$

أ- $\lambda = ن ل = ١٠٠(٠{,}٠١) = ١$

$$\therefore \; ع(ر = ٥) = \frac{هـ^{-١} (١)^{٥}}{٥!}$$

$$= \frac{٠{,}٣٦٧٩}{١٢٠} = ٠{,}٠٠٣١$$

ب- حيث أن القيمة المتوقَّعة لتوزيع بواسون = λ

∴ القيمة المتوقَّعة لعدد الأطفال الذين يتعرضون للإصابة بهذا المرض

= مريض واحد (حيث أن قيمة λ = ١).

التطبيق (١٠)

من واقع سجلات إحدى محطات تزويد السيارات بالوقود وُجِد أنه تتوقف ثلاث سيارات في المتوسط خلال نصف ساعة وذلك للتزود بالوقود.

المطلـــوب:

أ- احتمال توقف ٥ سيارات للتزود بالوقود خلال نصف ساعة.

ب- احتمال توقف سيارة واحدة على الأقل للتزود بالوقود خلال عشر دقائق.

ج- العدد المتوقَّع للسيارات التي تتوقف للتزويد بالوقود خلال ساعة.

الحـــل:

متوسط عدد السيارات في الدقيقة = $\frac{٣}{٣٠}$ = ٠,١ سيارة

أ- متوسط عدد السيارات في نصف ساعة = ٣٠ × ٠,١ = ٣ سيارات

أي أن: λ = ٣

وحيث أن: $ر$ = ٥

∴ الاحتمال المطلوب = $\frac{هـ^{-٣} (٣)^{٥}}{٥!}$ = $\frac{١٢,٠}{١٢٠}$ = ٠,١

ب- متوسط عدد السيارات في عشر دقائق = ١٠ × ٠,١ = سيارة واحدة

أي أن: λ = ١

وحيث أن: $ر \geq ١$

∴ $ع(ر \geq ١) = ١ - ع(ر = ٠)$

$= ١ - \frac{هـ^{-١} (١)^{٠}}{٠!}$

= ١ – ٠,٣٦٧٩ = ٠,٦٣٢١

ج- متوسط عدد السيارات في الساعة = ٦٠ × ٠,١ = ٦ سيارات

أي أن: λ = ٦

وحيث أن القيمة المتوقَّعة لتوزيع بواسون = λ

∴ العدد المتوقَّع للسيارات التي تقف خلال ساعة للتزود بالوقود = ٦ سيارات.

التطبيق (١١)

مجتمع مكون من عشرة طلاب منهم ٦ طلاب ذكور، ٤ طالبات. فإذا أُختيرت مجموعة مكونة من خمس طلاب بطريقة عشوائية وبدون إحلال، احسب احتمال أن تتكون المجموعة المختارة من:

أ- ثلاثة طلاب وطالبتيْن. **ب-** طالب واحد وأربع طالبات.

الحـــل:

التوزيع فوق الهندسي هو التوزيع الملائم في مثل تلك الحالات وهو يأخذ الصورة التالية:

$$\text{ع}(\text{س} = \text{ر}) = \frac{{}^{\text{ن}_1}\text{ق}_{\text{ر}} \times {}^{\text{ن}_2}\text{ق}_{\text{ں} - \text{ر}}}{{}^{\text{ن}}\text{ق}_{\text{ں}}} \text{ ، ر = ٠، ١،.... ں}$$

أ- حيث أن:

$\text{ن} = ١٠$ ، $\text{ں} = ٥$ ، $\text{ن}_1 = ٦$ ، $\text{ن}_2 = ٤$ ، $\text{ر} = ٣$

$$\therefore \text{المطلوب} = \frac{{}^{٦}\text{ق}_{٣} \times {}^{٤}\text{ق}_{٢}}{{}^{١٠}\text{ق}_{٥}} = ٠,٤٧٦$$

ب- $\text{ن} = ١٠$ ، $\text{ں} = ٥$ ، $\text{ن}_1 = ٦$ ، $\text{ن}_2 = ٤$ ، $\text{ر} = ١$

$$\therefore \text{الاحتمال المطلوب} = \frac{{}^{٦}\text{ق}_{١} \times {}^{٤}\text{ق}_{٤}}{{}^{١٠}\text{ق}_{٥}} = ٠,٠٠٤$$

التطبيق (١٢)

إذا كانت **ي** متغير معتاد معياري، احسب الاحتمالات التالية باستخدام الجداول:

أ- ع(ي > ١,٥) **ب-** ع(ي ≤ ٢,٤)

ج- ع(٠,٨ ≤ ي ≤ ٢,٧) **د-** ع(ي > - ١,٦)

هـ- ع(-٠,٣ ≤ ي ≤ ١,٨)

الحـــل:

أ- ع(ي > ١,٥) = ١– ع(ي ≤ ١,٥) = ١– ϕ(١,٥)

= ١ – ٠,٩٣٣٢ = ٠,٠٦٦٨

ب- ع(ي ≤ ٢,٤) = ϕ(٢,٤) = ٠,٩٩١٨

ج- ع(٠,٨ ≤ ي ≤ ٢,٧) = ϕ(٢,٧) – ϕ(٠,٨)

= ٠,٩٩٦٥ – ٠,٧٨٨١ = ٠,٢٠٨٤

د- ع(ي > - ١,٦) = ١ – ع(ي ≤ - ١,٦) = ١ – ϕ(- ١,٦)

وحيث أن: ϕ(- ي) = ١ – ϕ(ي)

∴ ع(ي > - ١,٦) = ١ – [١ – ϕ(١,٦)]

= ϕ(١,٦) = ٠,٩٤٥٢

هـ- ع(-٠,٣ ≤ ي ≤ ١,٨) = ϕ(١,٨) – ϕ(- ٠,٣)

= ϕ(١,٨) – [١ – ϕ(٠,٣)]

= ϕ(١,٨) + ϕ(٠,٣) – ١

= ٠,٩٦٤١ + ٠,٦١٧٩ = ٠,٥٨٢

التطبيق (١٣)

إذا كانت أوزان مجموعة كبيرة من الناس تتبع توزيعاً معتاداً بمتوسط ٧٥ كجم وانحراف معياري ٥ كجم. فإذا تم اختيار شخص واحد بطريقة عشوائية من بين مجموعة الأشخاص، احسب ما يلي:

أ- احتمال أن يزيد وزن الشخص عن ٨٠ كجم.

ب- احتمال أن يتراوح وزن الشخص ما بين ٦٥، ٨٥ كجم.

ج- ما هو الوزن الذي يزيد عليه ٦٥٪ من إجمالي عدد الأشخاص؟

د- ما هو الوزن الذي يقل عنه أو يساويه ٨٥٪ من مجموعة الأشخاص؟

الحـــل:

أ- ع(س > ٨٠) = ١ – ع(س ≤ ٨٠)

$$= ١ - \text{ع}\left(\frac{\text{س} - ٧٥}{٥} \le \frac{٨٠ - ٧٥}{٥}\right)$$

= ١ – ع(ي ≤ ١) = ١ – ϕ(١)

= ١ – ٠,٨٤١٣ = ٠,١٥٨٧

ب- $\text{ع}(٦٥ \le \text{س} \le ٨٥) = \phi\left(\frac{٨٥ - ٧٥}{٥}\right) - \phi\left(\frac{٦٥ - ٧٥}{٥}\right)$

= ϕ(٢) – ϕ(-٢)

$$= \phi(٢) - [١ - \phi(٢)]$$
$$= ٢\,\phi(٢) - ١$$
$$= ٢ \times ٠{,}٩٧٧٣ - ١ = ٠{,}٩٥٤٦$$

ج- المطلوب هو تحديد قيمة الوزن (س)، ولتكن (أ)، الذي يحقق:

$$ع(س > أ) = ٠{,}٦٥$$

أي أن: $١ - ع(س \leq أ) = ٠{,}٦٥$

$$\therefore\ ع(س \leq أ) = ٠{,}٣٥$$

وحيث أن $ع(س \leq أ) = ٠{,}٣٥$

$$\therefore\ ع\left(\frac{س - ٧٥}{٥} \leq \frac{أ - ٧٥}{٥}\right) = ع\left(ي \leq \frac{أ - ٧٥}{٥}\right) = ٠{,}٣٥$$

$$\therefore \phi\left(\frac{أ - ٧٥}{٥}\right) = ٠{,}٣٥$$

وحيث أن $\left(\frac{أ - ٧٥}{٥}\right)$ هو مقدار سالب فإنه يمكننا استنتاج ما يلي:

$$\phi\left(\frac{أ - ٧٥}{٥}\right) = ١ - \phi\left(\frac{٧٥ - أ}{٥}\right) = ٠{,}٣٥$$

أي أن: $\phi\left(\frac{٧٥ - أ}{٥}\right) = ٠{,}٦٥$

وبالبحث في جدول المســـاحات تحت المنحنى المعتاد المعياري عن قيمة ي المناظرة لـ $\phi(ي) = ٠{,}٦٥$ نجد أن:

$$\frac{٧٥ - أ}{٥} = ٠{,}٣٨٥$$

$$\therefore\ أ = ٧٣{,}٠٧٥ \text{ كجم}$$

أي أن الوزن الذي يزيد عليه ٦٥٪ من إجمالي عدد الأشخاص هو ٧٣ كجم.

د- المطلوب هو إيجاد قيمة الوزن (س)، ولتكن (أ)، الذي يحقق:

$$ع(س \leq أ) = ٠{,}٨٥$$

$$\therefore\ ع\left(\frac{س - ٧٥}{٥} \leq \frac{أ - ٧٥}{٥}\right) = ٠{,}٨٥$$

$$ع\left(ي \leq \frac{أ - ٧٥}{٥}\right) = ٠{,}٨٥$$

أي أن: $\phi\left(\frac{أ - ٧٥}{٥}\right) = ٠{,}٨٥$

وحيث أن $\left(\frac{\text{أ} - ٧٥}{٥}\right)$ مقدار موجب فإنه بالبحث في جدول المساحات تحت المنحنى المعتاد المعياري عن قيمة ي المناظرة لـ ϕ(ي) = ٠,٨٥ نجد أن:

$$\frac{\text{أ} - ٧٥}{٥} = ١,٠٤$$

∴ أ = ٧٥ + ٥ × ١,٠٤ = ٨٠,٢ كجم

التطبيق (١٤)

إذا كان عدد الطلاب الذين يرغبون في الالتحاق بكلية الشرطة هو ١٠٠٠٠ طالب، وكانت أطوالهم تتبع توزيعاً معتاداً متوسطه ١٧٠ سم وانحرافه المعياري ١٥ سم. فإذا علمت أن كلية الشرطة تقبل الأشخاص الذين تتراوح أطوالهم ما بين ١٤٥، ١٧٥ سم (بافتراض نجاحهم في بقية الاختبارات)، ما هو عدد الطلاب المتوقع قبولهم في كلية الشرطة.

الحـــل:

قبل ايجاد العدد المتوقَّع يلزم حساب الاحتمال: ع(١٤٥ ≤ س ≤ ١٧٥) وذلك كما يلي:

$$\text{ع}(١٤٥ \leq \text{س} \leq ١٧٥) = \phi\left(\frac{١٧٥ - ١٧٠}{١٥}\right) - \phi\left(\frac{١٤٥ - ١٧٠}{١٥}\right)$$

$$= \phi(٠,٣٣) - \phi(-١,٦٧)$$

$$= \phi(٠,٣٣) - [١ - \phi(١,٦٧)]$$

$$= \phi(٠,٣٣) + \phi(١,٦٧) - ١$$

$$= ٠,٦٢٩٣ + ٠,٩٥٢٥ - ١$$

$$= ٠,٥٨١٨$$

∴ عدد الطلاب المتوقَّع قبولهم = ١٠٠٠٠ × ٠,٥٨١٨

= ٥٨١٨ طالباً

تطبيقات
الفصل الثاني
(التقدير الاحصائي)

التطبيق (١٥)

إذا كان مجتمع البحث يتكون من ٣٥٠٠ مفردة بمتوسط ١٥٠ وانحراف معياري ٩٠، احسب التوقَّع والخطأ المعياري لتوزيع المعاينة للوسط الحسابي للعينة عندما يكون حجم العينة:

أ- ١٢٠ مفردة **ب -** ٢٧٠ مفردة

الحــــل:

أ- حيث أن: $ن = ٣٥٠٠$ ، $\text{ں} = ١٢٠$

$\therefore \text{ں} < ٠,٠٥ ن$

لذلك يجب عدم أخذ معامل التصحيح في الاعتبار عند حساب الخطأ المعياري.

ت($\bar{س}$) = $م_{\bar{س}}$ = م = ١٥٠ ،

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{\text{ں}}} = \frac{٩٠}{\sqrt{١٢٠}} = ٨,٢٢$$

ب- حيث أن: $ن = ٣٥٠٠$ ، $\text{ں} = ٢٥٠$

$\therefore \text{ں} > ٠,٠٥ ن$

وبالتالي يجب أخذ معامل التصحيح في الاعتبار عند حساب الخطأ المعياري

ت($\bar{س}$)= م = ١٥٠ ،

$$\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{\text{ں}}}\sqrt{\frac{ن - \text{ں}}{ن - ١}} = \frac{٩٠}{\sqrt{٢٥٠}}\sqrt{\frac{٣٥٠٠ - ٢٥٠}{٣٥٠٠ - ١}} = ٥,٤٩$$

التطبيق (١٦)

سُحبت عينة عشوائية مكونة من ٣٦ مصباحاً كهربائياً من إنتاج مصنع ما للمصابيح الكهربائية. فإذا علمت أن عمر التشغيل للمصابيح الكهربائية المنتَجة في هذا المصنع تتبع توزيعاً معتاداً متوسطه ٣٠٠ ساعة وانحرافه المعياري ٤٥ ساعة.

المطلـــوب:

أ- احتمال أن يزيد متوسط العينة عن ٣٢٠ ساعة.

ب- احتمال أن يتراوح متوسط العينة ما بين ٢٨٠، ٣١٠ ساعة.

الحــــل:

أ- $ع(\bar{س} > ٣٢٠) = ١ - ع(\bar{س} \leq ٣٢٠)$

$$= \text{١} - \text{ع}\left(\frac{\bar{\text{س}} - \text{٣٠٠}}{\frac{\text{٤٥}}{\sqrt{\text{٣٦}}}} \le \frac{\text{٣٢٠} - \text{٣٠٠}}{\frac{\text{٤٥}}{\sqrt{\text{٣٦}}}}\right)$$

$$= \text{١} - \text{ع}(\text{ي} \le \text{٢,٦٧}) = \text{١} - \phi(\text{٢,٦٧})$$

$$= \text{١} - \text{٠,٩٩٦٢} = \text{٠,٠٠٣٨}$$

ب- $\text{ع}(\text{٢٨٠} \le \text{س} \le \text{٣١٠}) = \phi\left(\frac{\text{٣١٠} - \text{٣٠٠}}{\frac{\text{٤٥}}{\sqrt{\text{٣٦}}}}\right) - \phi\left(\frac{\text{٢٨٠} - \text{٣٠٠}}{\frac{\text{٤٥}}{\sqrt{\text{٣٦}}}}\right)$

$$= \phi(\text{١,٣٣}) - \phi(\text{-٢,٦٧})$$

$$= \phi(\text{١,٣٣}) - [\text{١} - \phi(\text{٢,٦٧})]$$

$$= \phi(\text{١,٣٣}) + \phi(\text{٢,٦٧}) - \text{١}$$

$$= \text{٠,٩٠٨٢} + \text{٠,٩٩٦٢} - \text{١} = \text{٠,٩٠٤٤}$$

التطبيق (١٧)

إذا كان توزيع أجور العمال في أحد المصانع يتبع توزيعاً معتاداً انحرافه المعياري ١٥ جنيهاً. وبأخذ عينة عشوائية من عمال المصنع والبالغ عددهم ١٢٠٠ عاملاً وُجد أن متوسط أجور العمال يساوى ٨٥ جنيهاً، أوجد بفترة ثقة ٩٥٪ متوسط أجور العمال في المصنع إذا كان:

أ- حجم العينة ٨٠ عاملاً **ب-** حجم العينة ٢٠ عاملاً

الحـــل:

حيث أن المجتمع يتبع توزيعاً معتاداً تباينه معروف فإنه – وبغض النظر عن حجم العينة صغيراً كان أم كبيراً – يمكن استخدام التوزيع المعتاد في تقدير الوسط الحسابي للمجتمع.

(يُلاحَظ أن الحالتيْن أ، ب تناظران على الترتيب الحالتيْن ١، ٢ في الجدول (٢ – ٤)):

أ- حيث أن: $\text{ن} = \text{١٢٠٠}$ ، $\text{ں} = \text{٨٠}$

$\therefore \text{ں} > \text{٠,٠٥}\,\text{ن}$

$$\therefore \sigma_{\bar{\text{س}}} = \frac{\sigma}{\sqrt{\text{ں}}}\sqrt{\frac{\text{ن} - \text{ں}}{\text{ن} - \text{١}}} = \frac{\text{١٥}}{\sqrt{\text{٨٠}}}\sqrt{\frac{\text{١٢٠٠} - \text{٨٠}}{\text{١٢٠٠} - \text{١}}} = \text{١,٦٢}$$

وحيث أن: $\alpha = \text{٠,٠٥}$ ، $\text{ي}_{\frac{\alpha}{\text{٢}}} = \text{١,٩٦}$

∴ الحد الأدنى لفترة الثـقة = $\bar{س} - ي_{\frac{\alpha}{٢}} \frac{\sigma}{\sqrt{ن}} \sqrt{\frac{ن - ن}{ن - ١}}$

= ٨٥ ــ ١,٩٦ × ١,٦٢ = ٨١,٨٢ جنيهاً ،

الحد الأعلى لفترة الثـقة = ٨٥ + ١,٩٦ × ١,٦٢ = ٨٨,١٨ جنيهاً

أي أن متوسط أجور العمال في المصنع ككل يتراوح ما بين ٨١,٨٢ ، ٨٨,١٨ جنيهاً، وثـقتنا في ذلك تبلغ ٩٥٪.

ب- حيث أن: ن = ١٢٠٠ ، ن = ٢٠

∴ ن > ٠,٠٥ ن

∴ $\sigma_{\bar{س}} = \frac{\sigma}{\sqrt{ن}} = \frac{١٥}{\sqrt{٢٠}} = ٣,٣٥$

∴ الحد الأدنى لفترة الثـقة = $\bar{س} - ي_{\frac{\alpha}{٢}} (\frac{\sigma}{\sqrt{ن}})$

= ٨٥ ــ ١,٩٦ × ٣,٣٥ = ٧٨,٤٣ جنيهاً ،

الحد الأعلى لفترة الثـقة = ٨٥ + ١,٩٦ × ٣,٣٥ = ٩١,٥٧ جنيهاً

التطبيق (١٨)

تهتم كلية التجارة بإحدى الجامعات بدراسة متوسط عدد الساعات التي يقضيها الطلاب في مذاكرة دروسهم. ومن أجل ذلك تم اختيار عينة من طلاب الكلية، وبتسجيل الوقت الذي يقضيه الطالب في المذاكرة وُجِد أن متوسط عدد ساعات المذاكرة أسبوعياً هو ٢٤ ساعة، كما أن الانحراف المعياري هو ٤ ساعات. فإذا علمت أن الوقت الذي يقضيه الطلاب في المذاكرة يتبع توزيعاً معتاداً مجهولةُ معلماته. المطلوب إيجاد فترة الثـقة ٩٩٪ لمتوسط ساعات المذاكرة اسبوعياً إذا كان:

أ- حجم العينة ١٠٠ طالب   **ب-** حجم العينة ٢٥ طالب.

الحـــل:

أ- حيث أن المجتمع يتبع توزيعاً معتاداً غير معروف وسطه الحسابي وانحرافه المعياري وحيث أن حجم العينة كبيراً (ن = ١٠٠ > ٣٠) فإنه يمكن تطبيق نظرية الحد المركزية لاستخدام التوزيع المعتاد في تقدير الوسط الحسابي للمجتمع (يُلاحَظ أن هذه الحالة تمثل الحالة ٣ في الجدول (٢ ــ ٤)).

$ع_{\bar{س}} = \frac{ع}{\sqrt{ن}} = \frac{٤}{\sqrt{١٠٠}} = ٠,٤$

وحيث أن درجة الثـقة المطلوبة هي ٩٩٪ فإننا نجد أن:

α = ٠,٠١ ، $ي_{\frac{\alpha}{٢}}$ = ٢,٥٨

∴ الحد الأدنى لفترة الثقة = $\bar{س} - ي_{\frac{\alpha}{٢}} (\frac{ع}{\sqrt{ن}})$

= ٢٤ – ٢,٥٨ × ٠,٤ = ٢٢,٩٧ ساعة،

الحد الأعلى لفترة الثقة = ٢٤ + ٢,٥٨ × ٠,٤ = ٢٥,٠٣ ساعة

أي أن فترة الثقة ٩٩٪ لمتوسط ساعات المذاكرة أسبوعياً في المجتمع هي ٢٢,٩٧، ٢٥,٠٣ ساعة.

ب- حيث أن المجتمع يتبع توزيعاً معتاداً مجهولُ تباينه، كما أن حجم العينة صغير (ن = ٢٥ < ٣٠) فإن التوزيع الملائم في هذه الحالة هو توزيع ت (هذه الحالة تمثل الحالة (٤) في الجدول (٢ – ٤)).

الخطأ المعياري $ع_{\bar{س}} = \frac{ع}{\sqrt{ن}} = \frac{٤}{\sqrt{٢٥}}$ = ٠,٨

وحيث أن α = ٠,٠١ نجد أن:

$ت_{(ن-١)، \frac{\alpha}{٢}} = ت_{٢٤، ٠,٠٠٥}$ = ٢,٧٩٧ باستخدام جدول (ت)

∴ الحد الأدنى لفترة الثقة = $\bar{س} - ت_{٢٤، ٠,٠٠٥} (\frac{ع}{\sqrt{ن}})$

= ٢٤ – ٢,٧٩٧ × ٠,٨ = ٢١,٧٦ ساعة ،

∴ الحد الأعلى لفترة الثقة = ٢٤ + ٢,٧٩٧ × ٠,٨ = ٢٦,٢٤ ساعة

أي أن فترة الثقة ٩٩٪ لمتوسط ساعات المذاكرة أسبوعياً هي ٢١,٧٦ و ٢٦,٢٤ ساعة.

التطبيق (١٩)

في دراسة إحصائية لمعرفة معدل استهلاك نوع معين من السيارات للوقود تم اختيار عشر سيارات من هذا النوع من السيارات وتم تشغيلها وسُجلت المسافة التي تقطعها كل سيارة لكل لتر من البنزين. فكانت المسافات التي قطعتها السيارات العشر كما يلي:

٩ ١٤ ١٥ ١١ ١٣ ١٢ ١٣ ٨ ١٥ ١٠

المطلـوب:

إيجاد فترة الثقة ٩٠٪ لمتوسط المسافة التي تقطعها السيارة المنتَجة من هذا النوع لكل لتر واحد من البنزين وذلك بافتراض أن مُعدل استهلاك هذا النوع من السيارات يتبع التوزيع المعتاد.

الحـــل:

المجتمع هنا يتبع توزيعاً معتاداً غير معروف تباينه، وحيث أن حجم العينة صغير ($ن = ١٠ < ٣٠$) فإن التوزيع الملائم لتقدير متوسط المجتمع هو توزيع (ت) والذي يمكن استخدامه على النحو التالي:

الخطأ المعياري = $\frac{ع}{\sqrt{ن}}$، حيث ع هي الانحراف المعياري للعينة والذي يستخدم كتقدير للانحراف المعياري للمجتمع (σ) والذي يأخذ الصورة التالية:

$$ع = \sqrt{\frac{١}{ن - ١}\left[\sum س^٢ - \frac{(\sum س)^٢}{ن}\right]}$$

وبحساب مجموع قيم ($\sum س$) ومجموع مربعاتها ($\sum س^٢$) نجد ما يلي:

$$\sum س = ١٢٠ \quad ، \quad \sum س^٢ = ١٤٩٤$$

$$\therefore ع = \sqrt{\frac{١}{٩}\left[١٤٩٤ - \frac{(١٢٠)^٢}{١٠}\right]}$$

$$= \sqrt{\frac{٥٤}{٩}} = \sqrt{٦} = ٢,٤٥$$

وحيث أن $\alpha = ٠,١٠$ ،

$ت_{(ن-١)،\frac{\alpha}{٢}} = ت_{٩، ٠,٠٥} = ١,٨٣٣$ (جدول ت)

الخطأ المعياري = $\frac{ع}{\sqrt{ن}} = \frac{٢,٤٥}{\sqrt{١٠}} = ٠,٧٧$ ،

$$\bar{س} = \frac{\sum س}{ن} = \frac{١٢٠}{١٠} = ١٢$$

$\therefore$ الحد الأدنى لفترة الثقة = $\bar{س} - ت_{٩، ٠,٠٥} \frac{ع}{\sqrt{ن}}$

$= ١٢ - ١,٨٣٣ \times ٠,٧٧ = ١٠,٥٩$ ،

$\therefore$ الحد الأعلى لفترة الثقة = $١٢ + ١,٨٣٣ \times ٠,٧٧ = ١٣,٤١$

أي أن متوسط المسافة التي يقطعها هذا النوع من السيارات لكل لتر واحد من البنزين يتراوح ما بين ١٠,٥٧، ١٣,٤٣ كم وذلك بدرجة ثقة ٩٠٪.

التطبيق (٢٠)

إذا كان متوسط أجر العامل في الساعة لعينة مكونة من ٦٠ عاملاً في إحدى الشركات الكبرى هي ٧٥ جنيهاً. وكان الانحراف المعياري لهذا الأجر في الشركة هو ١٢ جنيهاً. أوجد فترة الثقة ٩٥٪ لمتوسط أجر العامل في الساعة في هذه الشركة.

الحـــل:

حيث أن حجم العينة كبير ($ن > ٣٠$) والانحراف المعياري للمجتمع معلوم، فإنه يمكن استخدام التوزيع المعتاد في تقدير متوسط المجتمع أياً كان توزيع المجتمع المسحوب منه العينة وذلك تطبيقاً لنظرية الحد المركزية.

حيث أن درجة الثقة المطلوبة هي ٩٥٪، فإننا نجد أن:

$$\alpha = ٠,٠٥ \quad ، \quad ي_{\frac{\alpha}{٢}} = ١,٩٦$$

$$\text{الخطأ المعياري} = \frac{\sigma}{\sqrt{ن}} = \frac{١٢}{\sqrt{٦٠}} = ١,٥٥$$

$$\therefore \text{الحد الأدنى لفترة الثقة} = \bar{س} - ي_{\frac{\alpha}{٢}} \left(\frac{\sigma}{\sqrt{ن}}\right)$$

= ٧٥ – ١,٩٦ × ١,٥٥ = ٧١,٩٦ جنيهاً ،

الحد الأعلى لفترة الثقة = ٧٥ + ١,٩٦ × ١,٥٥ = ٧٨,٠٤ جنيهاً

أي أن متوسط أجر العامل في الساعة في هذه الشركة يتراوح ما بين ٧٤,٩٦ و ٧٨,٠٤ جنيهاً وذلك بفترة ثقة ٩٥٪.

التطبيق (٢١)

في دراسة لتقدير متوسط دخل الأسرة في إحدى المحافظات، أُختيرت عينة مكونة من ٢٠٠ أسرة من هذه المحافظة فوجد أن المتوسط والانحراف المعياري لدخول الأسرة في العينة ٥٠٠٠ جنيهاً و ٢٠٠ جنيهاً على الترتيب. أوجد تقديراً لمتوسط دخل الأسرة في المحافظة مستخدماً في ذلك:

أ- درجة ثقة ٩٥٪ **ب-** درجة ثقة ٩٩٪

قارن بين نتائجك في الحالتيْن مُبدياً ما تراه مناسباً من تعليق.

الحــــل:

هنا يمكننا استخدام التوزيع المعتاد في تقدير متوسط المجتمع حيث نجد أن حجم العينة كبير ($ن$ = ٢٠٠ > ٣٠) وذلك تطبيقاً لنظرية الحد المركزية

أ- حيث أن درجة الثقة المطلوبة هي ٩٥٪، فإننا نجد أن:

$$\alpha = ٠,٠٥ \quad ، \quad ي_{\frac{\alpha}{٢}} = ١,٩٦$$

$$\text{الخطأ المعياري} = \frac{ع}{\sqrt{ن}} = \frac{٢٠٠}{\sqrt{٢٠٠}} = ١٤,١٤$$

$$\therefore \text{الحد الأدنى لفترة الثقة} = \bar{س} - ي_{\frac{\alpha}{٢}} \frac{ع}{\sqrt{ن}}$$

= ٥٠٠٠ – ١,٩٦ × ١٤,١٤ = ٤٩٧٢,٢٩ جنيهاً ،

الحد الأعلى لفترة الثقة = ٥٠٠٠ + ١,٩٦ × ١٤,١٤ = ٥٠٢٧,٧١ جنيهاً

أي أن فترة الثقة ٩٥٪ لمتوسط دخول الأسرة في المحافظة هي ٤٩٧٢,٢٩ و ٥٠٢٧,٧١ جنيهاً.

ب- في هذه الحالة نجد أن فترة الثقة المطلوبة هي ٩٩٪، وبالتالي ينتج أن:

$$\alpha = ٠,٠١ \quad ، \quad ي_{\frac{\alpha}{٢}} = ٢,٥٨$$

∴ الحد الأدنى لفترة الثقة = ٥٠٠٠ – ٢,٥٨ × ١٤,١٤ = ٤٩٦٣,٥٢ جنيهاً ،

الحد الأعلى لفترة الثقة = ٥٠٠٠ + ١,٩٦ × ١,٤١ = ٥٠٣٦,٤٨ جنيهاً

أي أن فترة الثقة ٩٩٪ لمتوسط دخول الأسرة في المحافظة هي ٤٩٦٣,٥٢ و ٥٠٣٦,٤٨ جنيهاً.

وبمقارنة فترة الثقة في أ، ب نجد أن فترة الثقة ٩٩٪ أكبر من فترة الثقة ٩٥٪. وهذا أمر طبيعي ومتوقَّع، إذ أنه كلما زادت درجة الثقة المطلوبة كلما زاد طول فترة الثقة (بافتراض ثبات العوامل الأخرى).

التطبيق (٢٢)

سُحبت عينة من ٢٠٠ مصباح كهربائي من النوع (أ) فكان الوسط الحسابي لعمر المصباح ١٥٠٠ ساعة بانحراف معياري ١٠٠ ساعة. كما سُحبت عينة أخرى حجمها ١٦٠ مصباحاً من النوع (ب) فكان الوسط الحسابي لعمر المصباح ١١٠٠ ساعة بانحراف معياري ٩٠ ساعة. قدِّر بفترة ثقة ٩٩٪ الفرق بين الوسطيْن الحسابييْن لأعمار المصابيح في النوعيْن أ و ب إذا كان من المعروف أن عمر المصابيح يتبع التوزيع المعتاد.

الحـــل:

حيث أن كلاً من المجتمعيْن يتبع التوزيع المعتاد فإنه يمكننا استخدام التوزيع المعتاد في تقدير الفرق بين المتوسطيْن أياً كان حجم كل من العينتيْن وذلك على النحو التالي:

α = ٠,٠١ ، $ي_{\frac{\alpha}{٢}}$ = ٢,٥٨ ،

الخطـأ المعياري للفرق بين متوسطيْن = $\sqrt{\frac{ع_١^٢}{ن_١} + \frac{ع_٢^٢}{ن_٢}}$

$$= \sqrt{\frac{(١٠٠)^٢}{٢٠٠} + \frac{(٩٠)^٢}{١٦٠}} = \sqrt{٥٠ + ٥٠,٣٢٥}$$

$$= \sqrt{١٠٠,٦٢٥} = ١٠,٠٣$$

∴ الحد الأدنى لفترة الثـقة = $(\bar{س}_١ - \bar{س}_٢) - ي_{\frac{\alpha}{٢}} \, ع_{(\bar{س}_١ - \bar{س}_٢)}$

= (١٥٠٠ – ١١٠٠) – ٢,٥٨ × ١٠,٠٣ = ٣٧٤,١٢ ساعة

الحد الأعلى لفترة الثـقة = (١٥٠٠ – ١١٠٠) + ٢,٥٨ × ١٠,٠٣ = ٤٢٥,٨٨ ساعة

وعلى ذلك فإن الفرق بين متوسطي أعمار المصابيح في النوعين أ و ب يتراوح ما بين ٣٧٤,١٢ و ٤٢٥,٨٨ ساعة وذلك بدرجة ثـقة ٩٩٪.

التطبيق (٢٣)

في التطبيق السابق، قدر بفترة ثـقة ٩٥٪ الفرق بين الوسطيْن الحسابييْن لأعمار المصابيح في النوعين أ و ب إذا كان حجم العينة من النوع (أ) ١٢ مصباحاً وحجم العينة من النوع (ب) ١٥ مصباحاً.

الحـــل:

حيث أن كلّاً من المجتمعيْن يتبع التوزيع المعتاد، كما أن حجم كل من العينتيْن صغيراً، فإن التوزيع الملائم لتقدير الفرق بين المتوسطين هو توزيع (ت). هذا وبافتراض أن المجتمعيْن لهما نفس التباين (المجهول) فإن الخطأ المعياري للفرق بين المتوسطيْن يأخذ الصورة التالية:

$ع_{(\bar{س}_١ - \bar{س}_٢)} = \sqrt{\frac{ع_م^٢}{ن_١} + \frac{ع_م^٢}{ن_٢}}$ وحيث:

$$ع_م^٢ = \frac{(ن_١ - ١)ع_١^٢ + (ن_٢ - ١)ع_٢^٢}{ن_١ + ن_٢ - ٢}$$

∴ الخطأ المعياري للفرق بين المتوسطيْن يمكن حسابه على النحو التالي:

$$ع_م^{٢} = \frac{(١٢-١)(١٠٠)^{٢} + (١٥-١)(٩٠)^{٢}}{١٢+١٥-٢} = ٨٩٣٦$$

$$∴ \text{الخطأ المعياري} = \sqrt{\frac{٨٩٣٦}{١٢} + \frac{٨٩٣٦}{١٥}} = ٣٦,٦١$$

$$∴ \text{الحد الأدنى لفترة الثقة} = (\bar{س}_١ - \bar{س}_٢) - ي_{\frac{\alpha}{٢}} ع_{(\bar{س}_١ - \bar{س}_٢)}$$

$$= (١٥٠٠ - ١١٠٠) - ١,٩٦ \times ٣٦,٦١ = ٣٢٨,٢٤ \text{ ساعة}$$

$$\text{الحد الأعلى لفترة الثقة} = (١٥٠٠ - ١١٠٠) + ١,٩٦ \times ٣٦,٦١ = ٤٧١,٧٦ \text{ ساعة}$$

وهذا يعني أن الفرق بين متوسطي أعمار المصابيح في النوعيْن أ و ب يقع ما بين ٣٢٨,٢٤، ٤٧١,٧٦ ساعة وذلك بدرجة ثقة ٩٥٪.

التطبيق (٢٤)

أُختيرت عينة حجمها ٨٠ طالباً من بين طلاب إحدى كليات التجارة. أوجد احتمال أن تزيد نسبة المدخنين في العينة عن ٢٥٪ إذا علمت أن نسبة المدخنين بين طلاب الكلية تبلغ ٣٠٪.

الحـــل:

$$\text{الخطأ المعياري} = \sqrt{\frac{لك}{ن}} = \sqrt{\frac{٠,٣ \times ٠,٧}{٨٠}} = ٠,٠٥١$$

$$ع(\hat{ل} > ٠,٢٥) = ١ - ع(\hat{ل} \le ٠,٢٥)$$

$$ع(\hat{ل} \le ٠,٢٥) = ع\left(\frac{\hat{ل} - ل}{\sqrt{\frac{لك}{ن}}} \le \frac{٠,٢٥ - ل}{\sqrt{\frac{لك}{ن}}}\right)$$

$$= ع\left(ي \le \frac{٠,٢٥ - ٠,٣}{٠,٠٥١}\right)$$

$$= ع(ي \le -٠,٩٨) = \phi(-٠,٩٨)$$

$$= ١ - \phi(٠,٩٨) = ١ - ٠,٨٣٦٥$$

$$= ٠,١٦٣٥$$

$$∴ ع(\hat{ل} > ٠,٢٥) = ١ - ٠,١٦٣٥ = ٠,٨٣٦٥$$

التطبيق (٢٥)

ترغب هيئة لاستطلاع الرأي العام في أن تقدر بدرجة ثقة ٩٥٪ نسبة الناخبين في المجتمع الذين يُتوقَّع أن يعطوا أصواتهم لمرشح معين. فقامت بأخذ عينة حجمها ١٢٠ ناخباً حيث وجدت أن منهم ٨٤ يؤيدون هذا المرشح. أوجد فترة الثقة لنسبة الناخبين المؤيدين لهذا المرشح في الدائرة الإنتخابية ككل؟

الحـــل:

نسبة المؤيدين لهذا المرشح في العينة $(\hat{ل}) = \frac{٨٤}{١٢٠} = ٠{,}٧$

الخطأ المعياري $= \sqrt{\frac{\hat{ل}\hat{ك}}{ن}} = \sqrt{\frac{٠{,}٣ \times ٠{,}٧}{١٢٠}} = ٠{,}٠٤٢$

وحيث أن درجة الثقة المطلوبة هي ٩٥٪ فإننا نجد أن:

$\alpha = ٠{,}٠٥$ ، $ي_{\frac{\alpha}{٢}} = ١{,}٩٦$

$\therefore$ الحد الأدنى لفترة الثقة $= \hat{ل} - ي_{\frac{\alpha}{٢}} \sqrt{\frac{\hat{ل}\hat{ك}}{ن}}$

$= ٠{,}٧ - ١{,}٩٦ \times ٠{,}٠٤٢ = ٠{,}٦٢$

الحد الأعلى لفترة الثقة $= ٠{,}٧ + ١{,}٩٦ \times ٠{,}٠٤٢ = ٠{,}٧٨$

أي أن نسبة المؤيدين لهذا المرشح في المجتمع تتراوح ما بين ٦٢٪ و ٧٨٪ وذلك بدرجة ثقة ٩٥٪.

التطبيق (٢٦)

في دراسة لمعرفة نسبة المصابين بمرض البلهارسيا في إحدى المحافظات أخذت عينة مكونة من ٥٠٠ شخص فوجد أن بينهم ٢٠ شخصاً مصابين بهذا المرض. أوجد فترة الثقة ٩٩٪ لنسبة المصابين بمرض البلهارسيا في المحافظة ككل.

الحـــل:

نسبة المصابين في العينة $(\hat{ل}) = \frac{٢٠}{٥٠٠} = ٠{,}٠٤$

الخطأ المعياري $= \sqrt{\frac{\hat{ل}\hat{ك}}{ن}} = \sqrt{\frac{٠{,}٩٦ \times ٠{,}٠٤}{٥٠٠}} = ٠{,}٠٠٨٨$

α = ٠,٠١ ، $ي_{\frac{\alpha}{٢}}$ = ٢,٥٨

∴ الحد الأدنى لفترة الثقة = $\hat{ل} - ي_{\frac{\alpha}{٢}} \sqrt{\frac{\hat{ل}\hat{ك}}{ن}}$

= ٠,٠٤ – ٢,٥٨ × ٠,٠٠٨٨ = ٠,٠١٧

الحد الأعلى لفترة الثقة = ٠,٠٤ + ٢,٥٨ × ٠,٠٠٨٨ = ٠,٠٦٣

لذلك يمكن القول بأن نسبة المصابين بمرض البلهارسيا في هذه المحافظة تتراوح ما بين ١,٧٪ و ٦,٣٪ وذلك بدرجة ثقة ٩٩٪.

التطبيق (٢٧)

أُخذت عينة عشوائية حجمها ١٠٠ طالب من كل من كليتي التجارة والآداب بإحدى الجامعات فوجد أن نسبة الطلاب الحاصلين على تقدير عام جيد فأقل في العام الماضي هي ٤٥٪، ٦٥٪ في كليتي التجارة والآداب على الترتيب. أوجد بفترة ثقة ٩٩٪ الفرق بين نسبة الطلاب الحاصلين على تقدير عام جيد فأقل في مجتمع الكليتيْن:

الحـــل:

α = ٠,٠١ ، $ي_{\frac{\alpha}{٢}}$ = ٢,٥٨

الخطأ المعياري للفرق بين نسبتيْن = $\sqrt{\frac{\hat{ل}_١\hat{ك}_١}{ن_١} + \frac{\hat{ل}_٢\hat{ك}_٢}{ن_٢}}$

= $\sqrt{\frac{٠,٤٥ \times ٠,٥٥}{١٠٠} + \frac{٠,٦٥ \times ٠,٣٥}{١٠٠}}$ = ٠,٠٦٩

∴ الحد الأدنى لفترة الثقة = $(\hat{ل}_١ - \hat{ل}_٢) - ي_{\frac{\alpha}{٢}} ع_{(\hat{ل}_١ - \hat{ل}_٢)}$

= (٠,٤٥ – ٠,٦٥) – ٢,٥٨ × ٠,٠٦٩ = - ٠,٣٧٨

الحد الأعلى لفترة الثقة = (٠,٤٥ – ٠,٦٥) + ٢,٥٨ × ٠,٠٦٩ = - ٠,٠٢٢

بما يعني أن الفرق بين نسبة الحاصلين على تقدير عام جيد فأقل في كليتي التجارة والآداب يتراوح ما بين – ٣١,٤، – ٨,٦٪ وذلك بدرجة ثقة ٩٩٪.

التطبيق (٢٨)

يرغب أحد الباحثين في اختيار عينة لتقدير متوسط الأجر الشهري للعاملين في إحدى شركات القطاع الخاص وذلك في حدود ١٢ جنيهاً (زيادة أو نقصاناً) عن المتوسط الحقيقي للأجور الشهرية للعاملين في هذه الشركة. فلو افترضنا أن هذا الباحث يعلم من خبرته الشخصية أن توزيع الأجور في الشركة يتبع توزيعاً معتاداً انحرافه ٥٠ جنيهاً، أوجد الحد الأدنى لحجم العينة الواجب اختياره بدرجة ثقة ٩٥٪.

وبافتراض عدم توافر معلومات عن الانحراف المعياري للأجور في الشركة، وأن البيانات المتاحة كانت كما يلي:

- أعلى أجر = ٣٥٠ جنيهاً
- أصغر أجر = ١١٠ جنيهاً

وضِّح كيف يمكن استخدام هذه البيانات في تحديد حجم العينة في هذه الحالة.

الحـــل:

حيث أن:

$$\text{خ} = ١٢ \quad ، \quad \text{ي}_{\frac{\alpha}{٢}} = ١,٩٦ \quad ، \quad \sigma = ٥٠$$

ونظراً لأنه لم يرد صراحة عدد الموظفين في الشركة ككل فإننا سوف نعتبر المجتمع هنا غير محدود، لذلك حجم العينة يتحدد كما يلي:

$$\text{ن} = \frac{\text{ي}^٢\sigma^٢}{\text{خ}^٢} = \frac{(١,٩٦)^٢(٥٠)^٢}{(١٢)^٢} = ٦٧$$

أي أن حجم العينة الواجب أخذه هو ٦٧ عاملاً.

حيث أنه من المعروف أن:

$$\textbf{المدى} = ٤\sigma$$

أي أن: $٣٥٠ - ١١٠ = ٤\sigma \qquad \therefore \sigma = ٦٠$

$$\therefore \text{ن} = \frac{\text{ي}^٢\sigma^٢}{\text{خ}^٢} = \frac{(١,٩٦)^٢(٦٠)^٢}{(١٢)^٢} = ٩٧$$

أي أن حجم العينة الواجب اختياره في هذه الحالة هو ٩٧ عاملاً على الأقل.

التطبيق (٢٩)

يرغب صاحب مزرعة كبيرة لتربية الدواجن في اختيار عينة لتقدير متوسط الدجاجة في مزرعته وذلك في حدود ١٢٥ جم (زيادة أو نقصاناً) عن المتوسط الحقيقي لوزن الدجاجة في المزرعة وذلك بدرجة ثقة ٩٩٪. فإذا افترضنا أنه معلومُ لدى صاحب المزرعة - من واقع خبرته السابقة - أن الانحراف المعياري لوزن الدجاجة في مزرعته يبلغ تقريبا ٠,٤ كجم. أوجد الحد الأدنى لحجم العينة الواجب اختياره حتى يتحقق هذا الغرض.

الحــــل:

نظراً لأنه لم يرد صراحة إجمالي عدد الدجاج في المزرعة فإننا سنعتبر أن المجتمع هنا غير محدود.

وحيث أن:

$$خ = ٠,١٢٥ \text{ كجم} \quad ، \quad ي_{\frac{\alpha}{٢}} = ٢,٥٨ \quad ، \quad \sigma = ٠,٤ \text{ كجم}$$

$$\therefore \; ن = \frac{ي^٢\sigma^٢}{خ^٢} = \frac{(٢,٥٨)^٢(٠,٤)^٢}{(٠,١٢٥)^٢} = ٦٩$$

أي أن حجم العينة الواجب اختياره هو ٦٩ دجاجة على الأقل.

التطبيق (٣٠)

يرغب مركز الدراسات السياسية والاقتصادية بمؤسسة الأهرام في اختيار عينة لتقدير نسبة الناخبين الذين يتوقع أن يعطوا اصواتهم لصالح أحد الأحزاب في إحدى المحافظات على ألا يتجاوز الخطأ ٠,٠٢ (زيادة او نقصانا) عن النسبة الحقيقة في المحافظة. ما هو الحد الأدنى لحجم العينـة الواجب اختياره إذا علمت أنه معلوم من دراسات أولية أن النسبة المئوية لتأييد هذا الحزب تبلغ ٢٥٪ (استخدم درجة ثقة ٩٥٪).

الحــــل:

حيث أن:

$$خ = ٠,٠٢ \quad ، \quad ن \text{ غير محدودة} \quad ، \quad ل = ٠,٢٥ \quad ، \quad ي_{\frac{\alpha}{٢}} = ١,٩٦$$

$$\therefore \; ن = \frac{ي^٢ ل ك}{خ^٢} = \frac{(١,٩٦)^٢(٠,٢٥)(٠,٧٥)}{(٠,٠٢)^٢} = ١٨٠١$$

أي أن حجم العينة الواجب اختياره هو ١٨٠١ ناخباً على الأقل.

التطبيق (٣١)

يرغب صاحب مصنع للملابس الجاهزة في تقدير نسبة الوحدات المعيبة في إنتاج مصنعه وذلك في حدود ± ٣٪ وبدرجة ثقة ٩٥٪. أوجد الحد الأدنى لحجم العينة الواجب اختياره لتحقيق هذا الغرض في الحالات التالية:

أ- إذا كانت الخبرة السابقة تشير إلى أن نسبة المعيب في إنتاج هذا المصنع تبلغ تقريباً ١٢٪.

ب- إذا لم تتوافر اية معلومات عن نسبة المعيب في إنتاج هذا المصنع.

قارن بين نتائجك في الحالتيْن مُبدياً ما تراه مناسباً من تعليق.

الحـــل:

أ- حيث أن:

$$\text{خ} = ٠,٠٣ \quad ، \quad \text{ي}_{\frac{\alpha}{٢}} = ١,٩٦ \quad ، \quad \text{ل} = ٠,١٢ \quad ، \quad \text{ن غير محدودة}$$

$$\therefore \text{ن} = \frac{\text{ي}^٢\text{ل}\text{ك}}{\text{خ}^٢} = \frac{(١,٩٦)^٢(٠,١٢)(٠,٨٨)}{(٠,٠٣)^٢} = ٤٥١$$

بما يعني أن حجم العينة الواجب اختياره هو ٤٥١ وحدة أو أكثر.

ب- حيث أنه لا تتوافر أية معلومات عن نسبة المعيب في إنتاج هذا المصنع فإنه – أخذاً بجانب الحذر – سوف نعتبر أن هذه النسبة هي ٠,٥ وذلك حتى نحصل على الحد الأعلى لحجم العينة الواجب اختياره في ظل الظروف المتاحة. أي أن:

$$\therefore \text{ن} = \frac{\text{ي}^٢\text{ل}\text{ك}}{\text{خ}^٢} = \frac{(١,٩٦)^٢(٠,٥)(٠,٥)}{(٠,٠٣)^٢} = ١٠٦٨$$

لذلك فإن حجم العينة الواجب اختياره في هذه الحالة هو ١٠٦٨ وحدة. وبمقارنة النتيجة في الحالتيْن نجد أن حجم العينة في الحالة الثانية يبلغ أكثر من ضعف حجمها في الحالة الأولى. وهذا أمر طبيعي ومتوقَّع نظراً لاعتبارنا ل = ك = ٠,٥ في الحالة الثانية وحيث يبلغ المقدار (ل ك) أكبر قيمة له.

وفي حالة اعتبار ل = ك = ٠,٥ يمكننا استخدام الصورة التالية في تحديد حجم العينة (ن):

$$\text{ن} = \left(\frac{\text{ي}_{\frac{\alpha}{٢}}}{٢\text{خ}}\right)^٢$$

والتي تعطى نفس النتيجة وذلك على النحو التالي:

$$ن = \left(\frac{١,٩٦}{٢ \times ٠,٠٣}\right)^{٢} = ١٠٦٨ \text{ وحدة}$$

وهي نفس النتيجة التي حصلنا عليها من قبل.

ويسمى حجم العينة في هذه الحالة بـ **" حجم العينة المتحفِّظ"** Conservative Sample Size، هذا إن صحت هذه التسمية.

تطبيقات
الفصل الثالث
(اختبار الفرضيات الإحصائية)

التطبيق (٣٢)

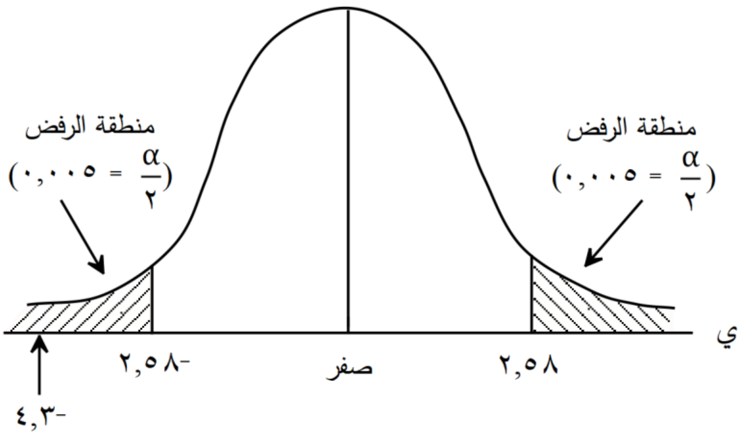
إذا كان متوسط عدد البيض لسلالة معينة من الدجاج في الشهر هو ٢٠ بيضة للدجاجة الواحدة. فإذا تم اختيار عينة مكونة من ١٠٠ دجاجة حيث وُجد أن متوسط عدد البيض للعينة ١٩ بيضة والانحراف المعياري ٢,٣ بيضة.
هل يمكن القول بأن هذه العينة تنتمي إلى هذه السلالة من الدجاج؟
استخدم مستوى معنوية ٠,٠١.

الحــل:

$\bar{س} = ١٩$ ، $ع = ٢,٣$ ، $ن = ١٠٠$ ، $م_{.} = ٢٠$

$ف_{.}: م = ٢٠$ ، $ف_{١}: م \neq ٢٠$

$$ي = \frac{\bar{س} - م_{.}}{\frac{ع}{\sqrt{ن}}} = \frac{١٩ - ٢٠}{\frac{٢,٣}{\sqrt{١٠٠}}} = \frac{-١}{٠,٧٣} = -٤,٣$$



وحيث أن الاختبار هنا ذو ذيليْن، وأن قيمة ي المحسوبة (-٤,٣) تقع في منطقة الرفض فإننا نرفض فرضية العدم ونقبل الفرضية البديلة (القيمة الحرجة = ± ٢,٥٨). أي أنه يمكن القول بأن هذه العينة لا تنتمي إلى هذه السلالة من الدجاج.

التطبيق (٣٣)

تتلقى جهة حكومية شكاوى كثيرة من المستهلِكين فحواها أن عبوات مسحوق الصابون التي تبيعها إحدى الشركات تحتوي على كمية أقل مــن ٢٥٠ جم من المسحوق وهي الكمية المُعلَن عنها. وللتحقق من صحة شكاوى المستهلِكين اشترت الجهة الحكومية ٩ عبوات من المسحوق

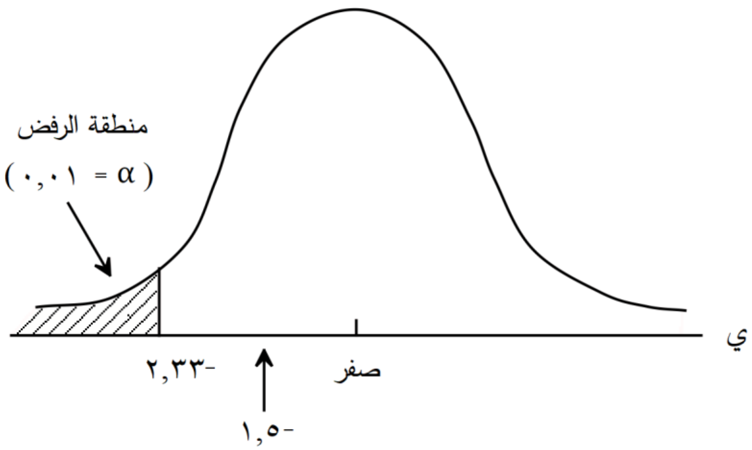
حيث وُجِد أن متوسط وزن المسحوق في العبوة = ٢٤٥ جم والانحراف المعياري =١٠جم. كيف يمكن للجهة الحكومية التحقق من صحة شكاوى المستهلِكين عند مستوى معنوية ٠,٠١ وذلك إذا عُلم أن وزن المسحوق في العبوات يتبع توزيعاً معتاداً؟

الحـــل:

$\bar{\text{س}} = ٢٤٥$ ، $\text{ع} = ١٠$ ، $\text{ن} = ٩$ ، $\text{م}_{٠} = ٢٥٠$

$\text{ف}_{٠}: \text{م} = ٢٥٠$ ، $\text{ف}_{١}: \text{م} < ٢٥٠$

$$\text{ي} = \frac{\bar{\text{س}} - \text{م}_{٠}}{\frac{\text{ع}}{\sqrt{\text{ن}}}} = \frac{٢٤٥ - ٢٥٠}{\frac{١٠}{\sqrt{٩}}} = -١,٥$$



الاختبار هنا ذو ذيل واحد وتتحدد منطقة الرفض بالمنطقة ($\text{ي} < -\text{ي}_{٠,٠١}$)، كما أن القية الحرجة تساوي -٢,٣٣ . وحيث أن قيمة **ي** المحسوبة تقع في منطقة عدم الرفض، فإننا لا نرفض فرضية العدم القائلة بأن متوسط وزن العبوة هو ٢٥٠ جراماً، وأن ادعاء مستهلِكي هذا النوع من الصابون ليس هناك ما يدعمه.

التطبيق (٣٤)

سُحبت عينة عشوائية من ٨٠ وحدة من إنتاج إحدى الآلات فوجد بها ٥ وحدات معيبة. المطلوب اختبار الفرضية القائلة بأن نسبة المعيب في إنتاج هذه الآلة هو ٦,٥٪ في مقابل الفرضية البديلة بأن هذه النسبة لا تساوى ٦,٥٪ وذلك عند مستوى معنوية ٠,٠٥.

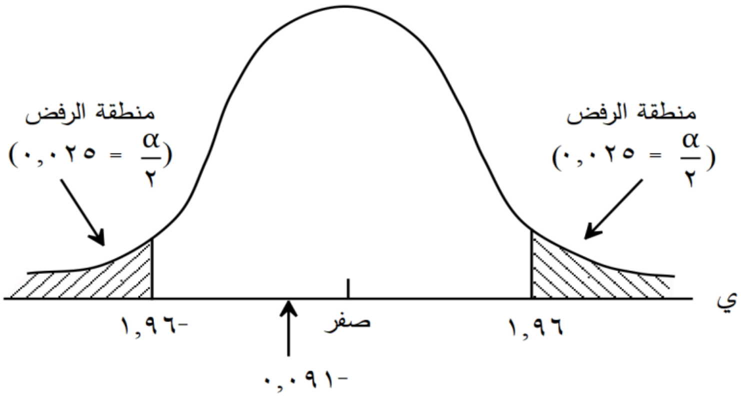
الحـــل:

نسبة الوحدات المعيبة في العينة ($\hat{ل}$) = $\frac{٥}{٨٠}$ = ٠,٠٦٢٥

∴ $\hat{ل}$ = ٠,٠٦٢٥ ، ں = ٨٠ ، $ل_{.}$ = ٠,٠٦٥

$ف_{.}$: ل = ٠,٠٦٢٥ ، $ف_{١}$: ل ≠ ٠,٠٦٥

$$ي = \frac{\hat{ل} - ل_{.}}{\sqrt{\frac{ل_{.}(١ - ل_{.})}{\text{ں}}}} = \frac{٠,٠٦٢٥ - ٠,٠٦٥}{\sqrt{\frac{٠,٠٦٥ \times ٠,٩٣٥}{٨٠}}} = ٠,٠٩١$$



الاختبار هنا ذو ذيليْن، وتتحـــــــدد منطقـــــة الرفـــــض بالمنطقتيْن (**ي** < – **ي**$_{٠,٠٢٥}$ أو **ي** > **ي**$_{٠,٠٢٥}$)، بينما القيمة الحرجة تســـاوي ± ١,٩٦. وحيث أن قيمة **ي** المحســـوبة (–٠,٠٩١) تقع في منطقة عدم الرفض فإنه يمكننا عدم رفض فرضـــية العدم. أي أنه يمكن القول بأن نسبة الوحدات المعيبة في إنتاج هذه الآلة هو ٦,٥٪.

التطبيق (٣٥)

في استطلاع لآراء الآباء حول الفصل بين الذكور والإناث في المرحلة الابتدائية، أجاب ٢٥٥ من بين ٥٠٠ بالموافقة. هل يمكن القول – وبمستوى معنوية ٠,٠٥ – بأن أغلبية الآباء يؤيدون هذا الأمر؟

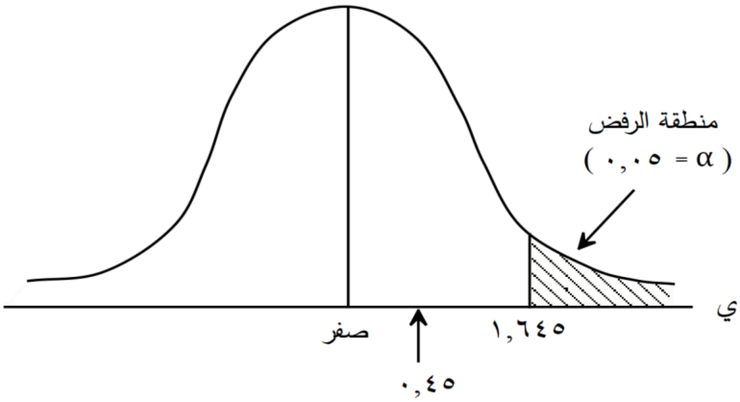
الحـــل:

نسبة أولياء الأمور الموافقين ($\hat{ل}$) = $\frac{٢٥٥}{٥٠٠}$ = ٠,٥١

$\hat{\text{ل}} = ٠,٥١$ ، $\text{ن} = ٥٠٠$ ، $\text{ل}_{٠} = ٠,٥$

$\text{ف}_{٠}: \text{ل} = ٠,٥$ ، $\text{ف}_{١}: \text{ل} > ٠,٥$

$$\text{ي} = \frac{\hat{\text{ل}} - \text{ل}_{٠}}{\sqrt{\frac{\text{ل}_{٠}(١ - \text{ل}_{٠})}{\text{ن}}}} = \frac{٠,٥١ - ٠,٥٠}{\sqrt{\frac{٠,٥ \times ٠,٥}{٥٠٠}}} = ٠,٤٥$$



وحيث أن الاختبار هنا ذو ذيل واحد، وأن منطقة الرفض تتحدد بالمنطقة ($\text{ي} > \text{ي}_{٠,٠٥}$)، كما أن القيمة الحرجة في هذه الحالة تساوي ١,٦٤٥. وبما أن قيمة **ي** المحسوبة تقع في منطقة عدم الرفض، فإننا لا نرفض فرضية العدم التي تفيد بأن نسبة أولياء الأمور الموافقين على قضية فصل الذكور والإناث في المرحلة الابتدائية هي ٥٠٪. وعلى ذلك فإن القول بأن أغلبية الآباء يؤيدون هذا الأمر هو قول لا تتوافر الأدلة الكافية على إثباته.

التطبيق (٣٦)

فيما يلي أطوال ٧ طلاب ذكور و ٩ طالبات إناث بالسنتيمتر تم اختيارهم عشوائياً من بين طلاب إحدى الكليات:

الطلاب (الذكور):	١٧٣	١٦٤	١٦٩	١٦٧	١٦٨	١٨٥	١٦٥
الطالبات (الإناث):	١٦٣	١٦٤	١٦٥	١٦٧	١٦٨	١٦٤	١٧١
	١٧٣	١٦٤					

باستخدام مستوى المعنوية ٥٪، هل تؤيد هذه البيانات صحة القول بأن الذكور أكثر طولاً من الإناث؟ وذلك بافتراض:

- أن طلاب هذه الكلية (ذكوراً و إناثاً) يمثلون عينة عشوائية من المجتمع عموماً وذلك فيما يتعلق بظاهرة الطول.
- أن أطوال الذكور والإناث في المجتمع تتبع توزيعاً معتاداً.

الحـــل:

بإيجاد الوسط الحسابي والانحراف المعياري لكل من عينتي الطلاب الذكور والطالبات الإناث نجد أن:

$$\bar{\text{س}}_{١} = ١٧٠,١ \quad ، \quad \text{ع}_{١} = ٧,٢ \quad ، \quad \text{ن}_{١} = ٧$$

$$\bar{\text{س}}_{٢} = ١٦٦,٦ \quad ، \quad \text{ع}_{٢} = ٣,٥ \quad ، \quad \text{ن}_{٢} = ٩$$

$$\text{ف}_{.}: (\text{م}_{١} - \text{م}_{٢}) = \text{صفر} \quad ، \quad \text{ف}_{١}: (\text{م}_{١} - \text{م}_{٢}) > \text{صفر}$$

وبافتراض تساوي التباينيْن في المجتمعيْن اللذيْن سحبت منهما العينتان، وبإيجاد التباين المشترك نجد أن:

$$\text{ع}_{\text{م}}^{٢} = \frac{(\text{ن}_{١} - ١)\text{ع}_{١}^{٢} + (\text{ن}_{٢} - ١)\text{ع}_{٢}^{٢}}{\text{ن}_{١} + \text{ن}_{٢} - ٢}$$

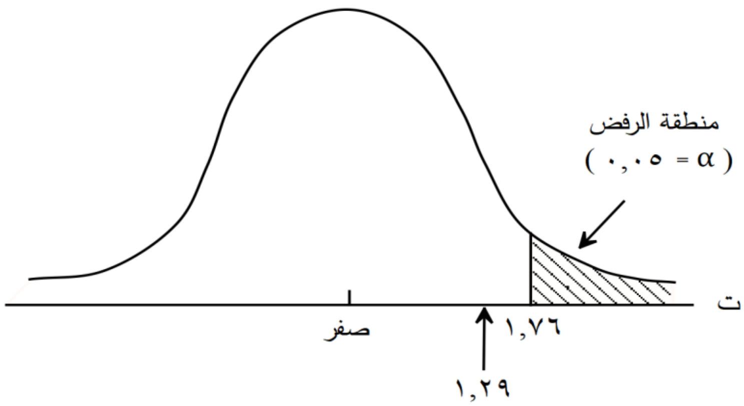
$$= \frac{(٧ - ١)(٧,٢)^{٢} + (٩ - ١)(٣,٥)^{٢}}{٧ + ٩ - ٢} = \frac{٢١٣,٠٤}{١٤} = ١٥,٢٢$$

وحيث أن حجم كل من العينتيْن صغير ($\text{ن}_{١}$، $\text{ن}_{٢}$ < ٣٠)، وأن أطوال الذكور والإناث في المجتمع تتبع التوزيع المعتاد، فإن المختبِر الإحصائي الملائم هو:

$$\text{ت} = \frac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\sqrt{\text{ع}_{\text{م}}^{٢}\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}}$$

$$= \frac{١٧٠,١ - ١٦٦,٦}{\sqrt{١٥,٢٢\left(\frac{١}{٧} + \frac{١}{٩}\right)}} = ١,٧٨$$

وذلك بدرجات حرية ١٤.



الاختبار هنا ذو ذيل واحد، وتتحدد منطقة الرفض بالمنطقة (ت > ت٠,٠٥)، بافتراض مستوى معنوية ٠,٠٥)، كما أن القيمة الحرجة في هذه الحالة هي ت ٠,٠٥،١٤ = ١,٧٦.

وحيث أن قيمة ت المحسوبة (١,٢٩) تقع في منطقة عدم الرفض، فإنه يمكننا عدم رفض فرضية العدم التي تعني أن الطلاب الذكور والطالبات الإناث لا يختلفون من حيث الطول. وأن القول بأن الذكور أكثر طولاً من الإناث هو محض ادعاء ليس هناك ما يؤكده.

التطبيق (٣٧)

قيس الزمن الذي يستغرقه عاملان فنيان في إصلاح عطل معين في نوع من السيارات في ٦٤ حالة. فإذا افترضنا أن قياساتهما تخضع لتوزيع معتاد، وأن الوسط الحسابي لقياسات العامل الأول كان ١٥٠ دقيقة وللعامل الثاني ١٤٥ دقيقة. وكان التباين في قياساتهما هو ٣٦، ٢٥ على التوالي:

هل يمكن القول بأن كفاءة العاملَيْن متعادلة في إصلاح ذلك العطل؟

استخدم مستوى معنوية ٠,٠٥.

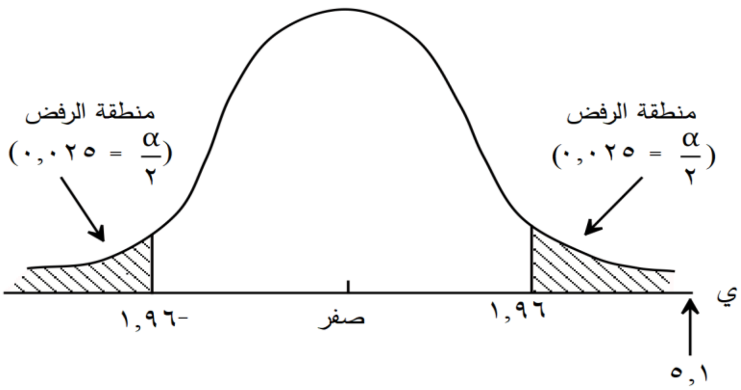
الحـــل:

$\bar{\text{س}}_{١} = ١٥٠$ ، $\text{ع}_{١}^{٢} = ٣٦$ ، $\text{ن}_{١} = ٦٤$

$\bar{\text{س}}_{٢} = ١٤٥$ ، $\text{ع}_{٢}^{٢} = ٢٥$ ، $\text{ن}_{٢} = ٦٤$

ف٠: $(\text{م}_{١} - \text{م}_{٢}) = \text{صفر}$ ، ف١: $(\text{م}_{١} - \text{م}_{٢}) \neq \text{صفر}$

$$\text{ي} = \frac{\bar{\text{س}}_{٢} - \bar{\text{س}}_{٢}}{\sqrt{\frac{\text{ع}_{١}^{٢}}{\text{ن}_{١}} + \frac{\text{ع}_{٢}^{٢}}{\text{ن}_{٢}}}} = \frac{١٥٠ - ١٤٥}{\sqrt{\frac{٣٦}{٦٤} + \frac{٢٥}{٦٤}}} = ٥,١$$



الاختبار في هذا الحالة هو اختبار ذو ذيليْن، ومنطقة الرفض تتحدد بالمنطقة ($ي < - ي_{٠,٠٢٥}$ أو $ي > ي_{٠,٠٢٥}$)، بينما القيمة الحرجة في هذه الحالة هي $ي_{٠,٠٢٥}$ وتساوي ± ١,٩٦.

وحيث أن قيمة ي المحسوبة (٥,١) تقع في منطقة الرفض فإننا نرفض فرضية العدم ونقبل الفرضية البديلة. أي أن القول بأن كفاءة العامليْن متعادلة في إصلاح ذلك العطل في السيارة هو أمر لا تؤكده الأدلة المتاحة.

التطبيق (٣٨)

أخذت عينتان من مصنعيْن ينتجان المصابيح الكهربائية وذلك لقياس العمر الاستهلاكي لهذه المصابيح (بالساعة) فكانت البيانات كما يلي:

العمر في المصنع الاول (أ) $ن_١ = ٢٠$				العمر في المصنع الثاني (ب) $ن_٢ = ١٥$		
٥٠	٩٥	٢١٠	١٢٠	٢١٠	٨٠	١٠٠
١٧٠	٩٠	٢٠٠	١٦٠	١٣٠	١٢٥	١٤٠
١٨٥	١١٠	١٠٠	١٣٠	٢٠٥	١٩٠	١٤٥
١٩٥	١٢٠	١١٠	١٥٠	١٦٠	١٨٥	١٢٠
١١٢	٨٠	٩٥	٧٠	١٠٠	٩٥	١١٠

بافتراض أن أعمار المصابيح الكهربائية في كل من المصنعيْن تتبع التوزيع المعتاد.

المطلوب:

أولاً: هل يمكن القول بأن متوسط عمر المصابيح في المصنع (ب) يبلغ ١٢٥ ساعة؟ استخدم α = ٠,٠١

ثانياً: ما مدى صحة ادعاء مدير المصنع (ب) والقائل بأن إنتاج مصنعه يفوق إنتاج المصنع (أ) من حيث كون مصابيحه أطول عمراً؟ استخدم α = ٠,٠٥

ثالثاً: هل يمكن القول بأن أعمار المصابيح في المصنعيْن واحدة؟ استخدم مستوى المعنوية ٠,٠١.

الحـــل:

بإيجاد الوسط الحسابي والانحراف المعياري لأعمار المصابيح الكهربائية في كل من العينتيْن المأخوذتيْن من إنتاج المصنعيْن أ و ب على الترتيب نجد أن:

$\bar{س}_١ = ١٢٧,٦$ ، $ع_١ = ٤٥,٩٥$ ، $ن_١ = ٢٠$

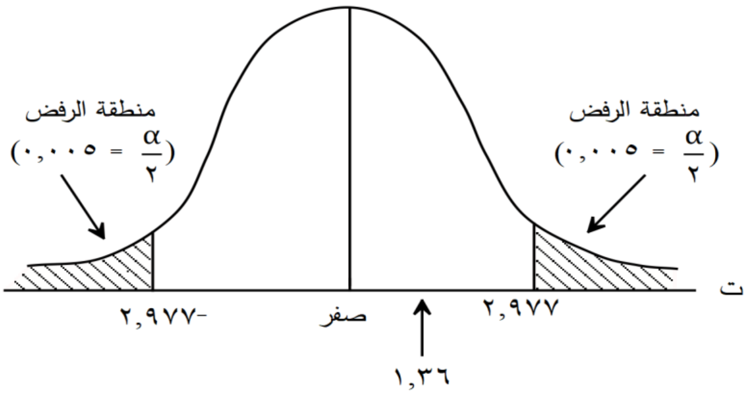
$\bar{س}_٢ = ١٣٩,٧$ ، $ع_٢ = ٤١,٨٥$ ، $ن_٢ = ١٥$

أولاً: $ف_٠ : م_٢ = ١٢٥$ ، $ف_١ : م_٢ \neq ١٢٥$

وحيث أن أعمار المصابيح في المصنع (ب) تتبع توزيعاً معتاداً كما أن ($ن_٢ < ٣٠$) فإن المختبر الإحصائي المناسب هو:

$$ت = \frac{\bar{س}_٢ - م_٠}{\frac{ع_٢}{\sqrt{ن_٢}}} = \frac{١٣٩,٧ - ١٢٥}{\frac{٤١,٨٥}{\sqrt{١٥}}} = ١,٣٦$$

وذلك بدرجات حرية ($ن_٢ - ١$) أي ١٤ درجة حرية.



منطقة الرفض في هذه الحالة تتحدد بالمنطقتيْن (ت < ت٠,٠٠٥ و ت > ت٠,٠٠٥) حيث أن الاختبار هنا ذي ذيليْن وأن منطقة الرفض تقع عند كل من الذيليْن الأيمن والأيسر. هذا والقيمة الحرجة هي ت ٠,٠٠٥،٣٣ = ٢,٧٣٣.

وبما أن قيمة ت المحسوبة (١,٣٦) تقع في منطقة عدم الرفض فإنه يمكننا عدم رفض فرضية العدم. أي أنه يمكن القول بأن متوسط أعمار المصابيح في المصنع ب يبلغ ١٢٥ ساعة.

ثانياً: $ف_٠ : (م_١ - م_٢) = صفر$ ، $ف_١ : (م_١ - م_٢) < صفر$

وبافتراض تساوي تباين أعمار المصابيح في المصنعيْن (أ و ب) فإن التباين المشترك هو:

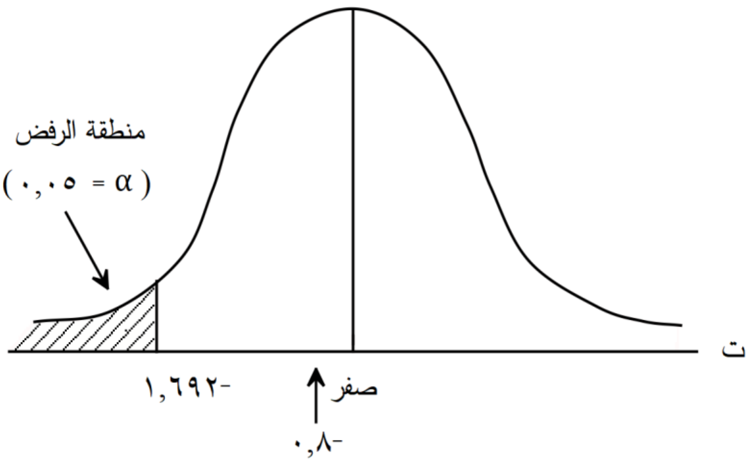
$$ع_م^٢ = \frac{(ن_١ - ١)ع_١^٢ + (ن_٢ - ١)ع_٢^٢}{ن_١ + ن_٢ - ٢}$$

$$= \frac{(٢٠ - ١)(٤٥,٩٥)^٢ + (١٥ - ١)(٤١,٨٥)^٢}{٢٠ + ١٥ - ٢} = ١٩٥٨,٦٨$$

وبذلك نجد أن المختبر الإحصائي المناسب هو:

$$\text{ت} = \frac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\sqrt{\text{ع}_{\text{م}}^{٢}\left(\frac{١}{\text{ن}_{١}} + \frac{١}{\text{ن}_{٢}}\right)}}$$

وذلك لأن أعمار المصابيح في كل من المصنعيْن تتبع توزيعاً معتاداً، كما أن حجم كل من العينتين صغير ($\text{ن}_{١}$، $\text{ن}_{٢}$ < ٣٠).



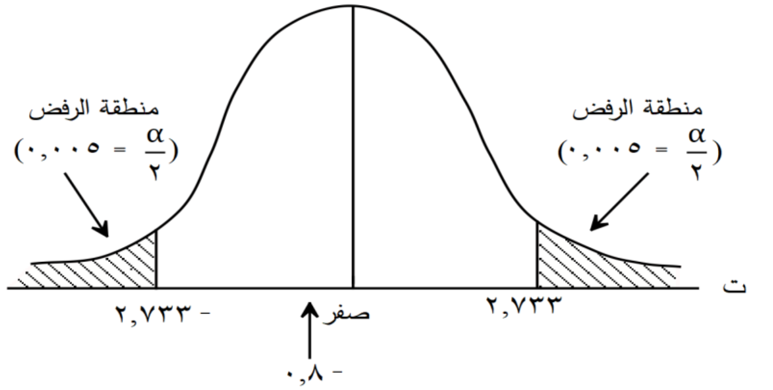
أي أن

$$\text{ت} = \frac{١٢٧,٦ - ١٣٩,٧}{\sqrt{١٩٥٨,٦٨\left(\frac{١}{٢٠} + \frac{١}{١٥}\right)}} = \frac{-١٢,١}{١٥,١٢} = -٠,٨$$

الاختبار في هذه الحالة هو اختيار ذو ذيل واحد، كما أن منطقة الرفض تقع عند الذيل الأيسر ($\text{ت} < -\text{ت}_{٠,٠٥}$)، وهنا قيمة ت الحرجة هي ت ٠,٠٥،٣٣ وتبلغ ١,٦٩٢. وحيث أن قيمة ت المحسوبة (–٠,٨) تقع في منطقة عدم الرفض، فإننا لا نرفض فرضية العدم. أي أنه يمكن القول بأن متوسط أعمار المصابيح في المصنع (أ) لا يختلف عن نظيره في المصنع (ب). وأن ادعاء مدير المصنع (ب) بأن إنتاج مصنعه يفوق إنتاج المصنع (أ) من حيث كون مصابيحه أطول عمراً هو ادعاء تنقصه الأدلة.

ثالثاً: $\text{ف}_{٠}: (\text{م}_{١} - \text{م}_{٢}) = \text{صفر}$ ، $\text{ف}_{١}: (\text{م}_{١} - \text{م}_{٢}) \neq \text{صفر}$

هذا وقد تم حساب قيمة المختبر الإحصائي (ت) في ثانياً.



وحيث أن الاختبار هنا ذو ذيليْن وأن منطقة الرفض تقع عند كل من الذيليْن الأيمن والأيسر (ت < – ت٠,٠٠٥، ت > ت٠,٠٠٥) فإن قيمة ت المحسوبة (– ٠,٨) تقع في منطقة عدم الرفض (القيمة الحرجة هي ت ٠,٠٠٥،٣٣ وتبلغ ٢,٧٣٣ . وعليه يمكن القول بان متوسط أعمار المصابيح واحد في المصنعيْن وذلك عند مستوى معنوية ٠,٠١.

التطبيق (٣٩)

في عينة تحتوي على ١٠٠ قارئ للصحف اليومية في مدينة القاهرة وُجد أن منهم ٦٥ قارئاً يحرصون على قراءة صحيفة الأخبار. وفي عينـة أخرى من مدينة الزقازيق حجمها ٨٠ قارئاً وُجد أن منهم ٥٠ قارئاً يحرصون على قراءة نفس الصحيفة.

هل يمكن القول بأن صحيفة الأخبار هي أكثر تفضيلاً لدى قراء الصحف في مدينة القاهرة عنها في مدينة الزقازيق؟ استخدم مستوى معنوية ٠,٠٥

الحـــل:

من بيانات العينتيْن نجد ما يلي:

النسبة في العينة الأولى $(\hat{ل}_{١}) = \frac{٦٥}{١٠٠} =$ ٠,٦٥

النسبة في العينة الثانية $(\hat{ل}_{٢}) = \frac{٥٠}{٨٠} =$ ٠,٦٢٥

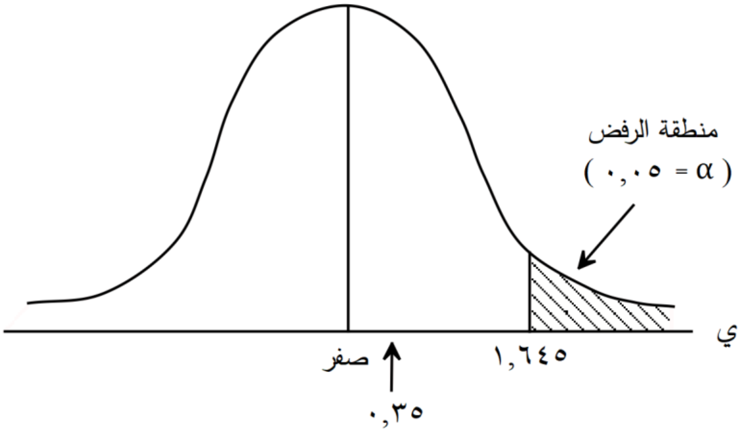
ف٠: $(ل_{١} - ل_{٢})$ = صفر ، ف١: $(ل_{١} - ل_{٢})$ > صفر

المختبِر الإحصائي في هذه الحالة يأخذ الصورة:

$$ي = \frac{\hat{ل}_{١} - \hat{ل}_{٢}}{\sqrt{\hat{ل}\left(١ - \hat{ل}\right)\left(\frac{١}{ن_{١}} + \frac{١}{ن_{٢}}\right)}}$$

حيث: $\hat{\text{ل}} = \dfrac{\text{س}_{١} + \text{س}_{٢}}{\text{ن}_{١} + \text{ن}_{٢}} = \dfrac{٦٥ + ٥٠}{١٠٠ + ٨٠} = \dfrac{١١٥}{١٨٠} = ٠,٦٤$

$$\therefore \text{ي} = \dfrac{٠,٦٥ - ٠,٦٢٥}{\sqrt{٠,٦٤ \times ٠,٣٦ \left(\dfrac{١}{١٠٠} + \dfrac{١}{٨٠}\right)}} = \dfrac{٠,٠٢٥}{٠,٠٧٢} = ٠,٣٥$$



الاختبار هنا ذو ذيل واحد حيث تتحدد منطقة الرفض بالمنطقة ($\text{ي} > \text{ي}_{٠,٠٥}$)، القيمة الحرجة هنا هي $\text{ي}_{٠,٠٥}$ وتساوي ١,٦٤٥. وبما أن قيمة **ي** المحسوبة (٠,٣٥) تقع في منطقة عدم الرفض فإننا لا نرفض فرضية العدم. أي أن القول بأن صحيفة الأخبار أكثر تفضيلاً لدى قراء الصحف في مدينة القاهرة عنها في مدينة الزقازيق هو قول لا تؤيده الأدلة المتاحة.

التطبيق (٤٠)

لنفترض أن لدينا عينتيْن الأولى حجمها ١٢٠ شخصاً والثانية حجمها ١٨٠ شخصاً. وأن عدد المدخنين في العينة الأولى هو ١٢ شخصاً وفي العينة الثانية ٢٥ شخصاً.

المطلوب:

أولاً: اختبر صحة الفرضية القائلة بأن نسبة المدخنين في كل من المجتمعيْن الذيْن سحبت منهما العينتان واحدة. استخدم مستوى معنوية ٠,٠١

ثانياً: هل يمكن القول بأن ظاهرة التدخين هي أكثر انتشاراً في المجتمع الثاني عنها في المجتمع الأول؟ استخدم معنوية ٠,٠٥

الحــل:

من بيانات العينتيْن نجد أن:

نسبة المدخنين في العينة الأولى ($\hat{\text{ل}}_{١}$) = $\dfrac{١٢}{١٢٠} = ٠,١$

نسبة المدخنين في العينة الثانية ($\hat{ل}_٢$) = $\frac{٢٥}{١٨٠} = ٠,١٤$

أولاً: $ف_٠$: ($ل_١ - ل_٢$) = صفر ، $ف_١$: ($ل_١ - ل_٢$) $\neq$ صفر

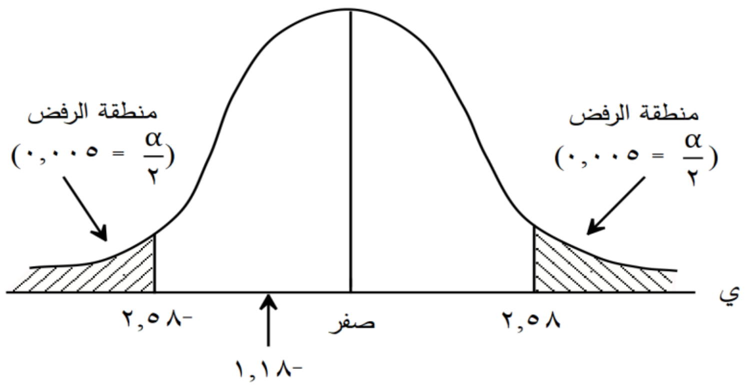
الاختبار في هذا الحالة ذو ذيلين حيث تقع منطقة الرفض عند كل من الذيلين الأيمن والأيسر ($ي < -ي_{٠,٠٥}$ و $ي > ي_{٠,٠٥}$).

وحيث أن:

$$ي = \frac{\hat{ل}_١ - \hat{ل}_٢}{\sqrt{\hat{ل}\left(١ - \hat{ل}\right)\left(\frac{١}{ن_١} + \frac{١}{ن_٢}\right)}}$$

حيث: $\hat{ل} = \frac{س_١ + س_٢}{ن_١ + ن_٢} = \frac{١٢ + ٢٥}{١٢٠ + ١٨٠} = \frac{٢٧}{٣٠٠} = ٠,٠٩$

$$\therefore ي = \frac{٠,١ - ٠,١٤}{\sqrt{٠,٠٩ \times ٠,٩١\left(\frac{١}{١٢٠} + \frac{١}{١٨٠}\right)}} = \frac{-٠,٠٤}{٠,٠٣٤} = -١,١٨$$



وبما أن قيمة **ي** المحسوبة (– ١,١٨) تقع في منطقة عدم الرفض فإننا لا نرفض فرضية العدم. وبذلك يمكننا القول بان نسبة المدخنين في المجتمعيْن واحدة.

ثانياً: فرضية العدم ($ف_٠$) والفرضية البديلة ($ف_١$) في هذه الحالة تأخذان الصورة:

$ف_٠$: ($ل_١ - ل_٢$) = صفر ، $ف_١$: ($ل_١ - ل_٢$) < صفر

أي أن الاختبار هنا هو اختبار ذو ذيل واحد حيث تقع منطقة الرفض عند الذيل الأيسر ($ي < -ي_{٠,٠٥}$)، كما أن القيمة الحرجة هي $ي_{٠,٠٢٥}$ وتساوي $\pm$ ٢,٥٨.

وبما أن قيمة **ي** المحسوبة (– ١,١٨) – والتي تم إيجادها في أولاً – تقع في منطقة عدم الرفض فإننا لا نرفض فرضية العدم.

أي أن القول بأن ظاهرة التدخين هي أكثر انتشاراً في المجتمع الثاني عنها في المجتمع الأول هو محض ادعاء ليس له ما يبرره.

التطبيق (٤١)

في دراسة لتقييم فعالية نوع جديد من الأسبرين، تم اختيار عينة عشوائية مكونة من ٢٠٠ شخص يعانون من الصداع حيث وُجِد أن ١٦٠ شخصاً منهم قد زالت عنهم آثار الصداع.
ناقش ادعاء المصنع المنتِج لهذا النوع من الأسبرين بأنه يستطيع أن يشفي ٩٠٪ ممن يتعاطونه من الصداع. استخدم مستوى معنوية ٠,٠٥.

الحـــل:

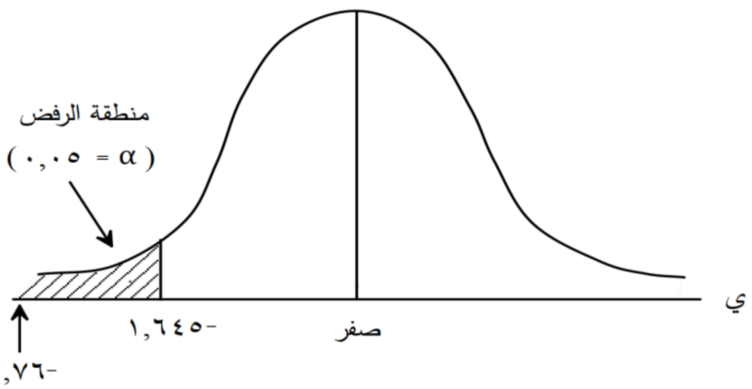
من بيانات العينة نجد أن:

$$\hat{\text{ل}} = \frac{١٦٠}{٢٠٠} = ٠,٨$$

وفي هذه الحالة تأخذ فرضية العدم والفرضية البديلة الصورة:

$\text{ف}_{.}: \text{ل} = ٠,٩$ ، $\text{ف}_{١}: \text{ل} < ٠,٩$

$$\text{ي} = \frac{\hat{\text{ل}} - \text{ل}_{.}}{\sqrt{\frac{\text{ل}_{.}(١ - \text{ل}_{.})}{\text{ن}}}} = \frac{٠,٨ - ٠,٩}{\sqrt{\frac{٠,٩ \times ٠,١}{٢٠٠}}} = -٤,٧٩$$



ولكون الاختبار ذي ذيل واحد في هذه الحالة وأن منطقة الرفض تقع عنـــد الذيل الأيسر ($\text{ي} < -\text{ي}_{٠,٠٥}$)، كما أن القيمة الحرجة في حالتنا هذه هي – ١,٦٤٥. وحيـث أن قيمـــة ي المحسوبـــة تقع في منطقة الرفض، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة. وهذا يعني أن ادعاء المصنع المنتِج لهذا النوع من الأسبرين بأنه يستطيع أن يشفي ٩٠٪ ممن يتعاطونه من الصداع هو مجرد ادعاء تنقصه الأدلة.

تطبيقات
الفصل الرابع
(اختبارات مربع كاي)

التطبيق (٤٢)

لدراسة العلاقة ما بين وقت العمل وإنتاجية العامل (معبَّراً عنها بعدد الوحدات المنتَجة) خلال فترة العمل، أُخذت عينة عشوائية مكونة من ٥٠٠ عامل من أحد المصانع الكبيرة فكانت النتائج التالية:

وقت العمل / الإنتاجية	صباحاً	ظهراً	مساءً	المجموع
٢٠ ــ ٢٤	٣٠	١٠٠	٧٠	٢٠٠
٢٥ ــ ٢٩	٤٠	٢٥	٤٥	١١٠
٣٠ ــ ٣٤	٥٠	٧٥	٦٥	١٩٠
المجموع	١٢٠	٢٠٠	١٨٠	٥٠٠

هل تعطي بيانات العينة دليلاً كافياً للقول بأن إنتاجية العامل في هذا المصنع تتأثر بوقت العمل؟ استخدم مستوى المعنوية ٠,٠٥.

الحــــل:

١- صياغة فرضية العدم والفرضية البديلة:

ف$_{٠}$: ليست هناك علاقة بين إنتاجية العامل ووقت العمل.

ف$_{١}$: إنتاجية العامل تتأثر بوقت العمل.

٢- اختيار التوزيع الملائم:

التوزيع الملائم هنا هو توزيع مربع كاي وذلك لأننا نبحث استقلالية صفتيْن (وقت العمل وإنتاجية العامل). ومن ثم فإن الاختبار الملائم هو اختبار مربع كاي للاستقلال.

٣- تحديد منطقتي الرفض وعدم الرفض:

في هذا التطبيق نجد أن:

ص = ٣ و د = ٣

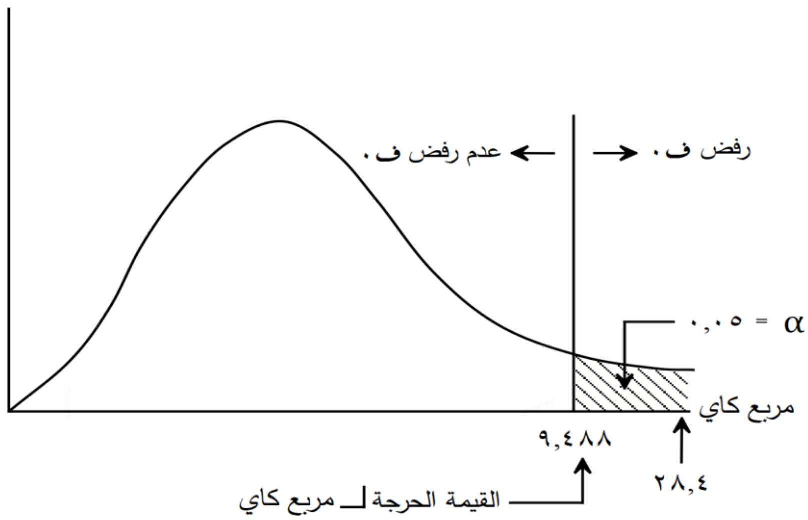
أي أن درجات الحرية = (ص ــ ١)(د ــ ١)

= (٣ ــ ١)(٣ ــ ١) = ٢ × ٢ = ٤

ونظراً لأن مستوى المعنوية المطلوب هو α = ٠,٠٥، فإننا نبحث في جدول المساحة تحت منحنى توزيع مربع كاي (الملحق ٣) عن القيمة الحرجة لمربع كاي عند درجات حرية ٤ ومستوى معنوية (مساحة على الطرف الأيسر) ٠,٠٥ حيث نجد أن:

القيمة الحرجة لمربع كاي = ٩,٤٨٨

ويوضح الشكل التالي منطقتي الرفض وعدم الرفض لهذا الاختبار.



٤- إيجاد قيمة إحصاء الاختبار:

نوجد أولاً التكرار المتوقَّع لكل خلية من خلايا الجدول وذلك بافتراض صحة فرضية العدم.

الصف الأول:

التكرار المتوقَّع للخلية الأولى = $\frac{٢٠٠ \times ١٢٠}{٥٠٠}$ = ٤٨

التكرار المتوقَّع للخلية الثانية = $\frac{٢٠٠ \times ٢٠٠}{٥٠٠}$ = ٨٠

التكرار المتوقَّع للخلية الثالثة = $\frac{٢٠٠ \times ١٨٠}{٥٠٠}$ = ٧٢

الصف الثاني:

التكرار المتوقَّع للخلية الرابعة = $\frac{١١٠ \times ١٢٠}{٥٠٠}$ = ٢٦,٤

التكرار المتوقَّع للخلية الخامسة = $\frac{١١٠ \times ٢٠٠}{٥٠٠}$ = ٤٤

التكرار المتوقَّع للخلية السادسة = $\frac{١١٠ \times ١٨٠}{٥٠٠}$ = ٣٩,٦

الصف الثالث:

التكرار المتوقَّع للخلية السابعة = $\frac{١٩٠ \times ١٢٠}{٥٠٠}$ = ٤٥,٦

التكرار المتوقَّع للخلية الثامنة = $\frac{١٩٠ \times ٢٠٠}{٥٠٠}$ = ٧٦

التكرار المتوقَّع للخلية التاسعة = $\frac{١٩٠ \times ١٨٠}{٥٠٠}$ = ٦٨,٤

وهنا، دعنا نحسب قيمة مربع كاي باستخدام الصيغة (٤ – ٤) وهي:

$$\text{مربع كاي} = \sum \frac{\text{ش}^{٢}}{\text{ت}} - \text{ن}$$

أي أن:

$$\text{مربع كاي} = [\frac{(٣٠)^{٢}}{٤٨} + \frac{(١٠٠)^{٢}}{٨٠} + \frac{(٧٠)^{٢}}{٧٢} + \frac{(٤٠)^{٢}}{٢٦,٤} + \frac{(٢٥)^{٢}}{٤٤} + \frac{(٤٥)^{٢}}{٣٩,٦}$$

$$+ \frac{(٥٠)^{٢}}{٤٥,٦} + \frac{(٧٥)^{٢}}{٧٦} + \frac{(٦٥)^{٢}}{٦٨,٤}] - ٥٠٠$$

= (١٨,٨ + ١٢٥ + ٦٨,١ + ٦٠,٦ + ١٤,٢ + ٥١,١ + ٥٤,٨ + ٧٤,٠
+ ٦١,٨) – ٥٠٠ = ٥٢٨,٤ – ٥٠٠ = ٢٨,٤

والجدول التالي يوضح التكرار المشاهَد والتكرار المتوقَّع لكل خلية من خلايا جدول التوافق.

ملحوظة: أن التكرار المتوقَّع موضوع بين قوسيْن في هذا الجدول والجداول التالية.

وقت العمل / الإنتاجية	صباحاً	ظهراً	مساءً	المجموع
٢٠ – ٢٤	٣٠ (٤٨)	١٠٠ (٨٠)	٧٠ (٧٢)	٢٠٠
٢٥ – ٢٩	٤٠ (٢٦,٤)	٢٥ (٤٤)	٤٥ (٣٩,٦)	١١٠
٣٠ – ٣٤	٥٠ (٤٥,٦)	٧٥ (٧٦)	٦٥ (٦٨,٤)	١٩٠
المجموع	١٢٠	٢٠٠	١٨٠	٥٠٠

٥- اتخاذ القرار:

حيث أن قيمة مربع كاي المحسوبة (٢٨,٤) أكبر من القيمة الحرجة لمربع كاي (٩,٤٨٨) وتقع في منطقة الرفض، فإننا نرفض فرضية العدم ونقبل الفرضية البديلة.

٦- رفض فرضية العدم وقبول الفرضية البديلة يفيد بأنه يمكننا القول بأن إنتاجية العامل تتأثر بوقت العمل. أي أن الصفتيْن (إنتاجية العامل ووقت العمل) غير مستقلتيْن.

التطبيق (٤٣)

لاحَظ مدير أحد فروع شركة عمر أفندي أن العملاء يفضلون بعض البائعين على البعض الآخر. وللتحقق من ذلك، تم تسجيل العملاء الذين دخلوا الفرع في أحد الأيام والبائعين الذين قاموا بخدمتهم فكانت النتائج التالية:

رقم البائع	١	٢	٣	٤	٥	المجموع
عدد العملاء	١٦٠	١٣٠	١٥٠	١٩٠	١٧٠	٨٠٠

هل تؤكد هذه البيانات صحة ملاحظة المدير؟ استخدم مستوى المعنوية ٠,٠١.

الحـــل:

١- صياغة فرضية العدم والفرضية البديلة:

ف. : العملاء يوزعون بالتساوي على البائعين.

ف١: العملاء يفضلون بعض البائعين على البعض الآخر.

٢- اختيار التوزيع الملائم:

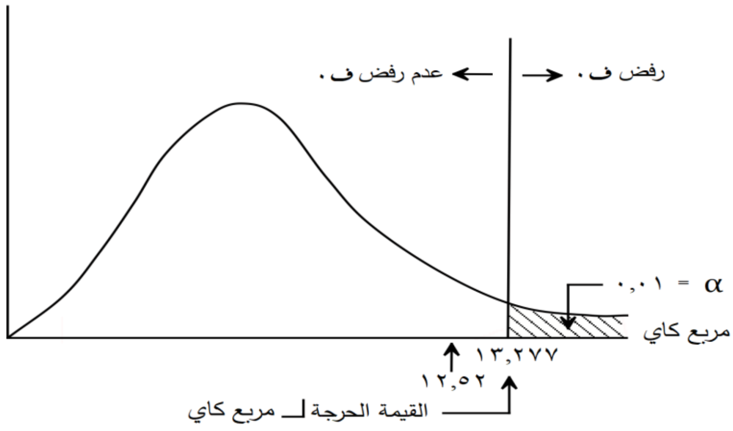
الجدول المعطى لا يمثل جدول توافق ونحن هنا بصدد اختبار مربع كاي لجودة التوفيق.

٣- تحديد منطقتي الرفض وعدم الرفض:

حيث أن: عدد البائعين = ٥ ، أي أن: ك = ٥

فإن درجات الحرية = ك – ١ = ٥ – ١ = ٤

ونظراً لأن مستوى المعنوية المطلوب ٠,٠١ فإنه بالبحث عن القيمة الحرجة لـ مربع كاي في جدول المساحة تحت منحنى توزيع مربع كاي عند درجات حرية ٤ ومستوى معنوية (المساحة على الطرف الأيمن للمنحنى) ٠,٠١ فإننا نجد أن: القيمة الحرجة لـ مربع كاي = ١٣,٢٧٧ هذا والشكل التالي يوضح منطقتي الرفض وعدم الرفض لهذا الاختيار



٤- إيجاد قيمة إحصاء الاختبار:

الجدول التالي يوضح الخطوات اللازمة لحساب التكرارات المتوقَّعة وذلك بافتراض توزيع العملاء بالتساوي على البائعين الخمسة، أي بافتراض صحة فرضية العدم. هذا وصحة فرضية العدم تعني أن احتمال أن يذهب العميل إلى أي بائع من البائعين الخمسة هو $\frac{١}{٥}$.

رقم البائع	التكرار المشاهَد (ش)	الاحتمال (ع)	التكرار المتوقَّع ت = ن ع	ش – ت	$\frac{(ش – ت)^{2}}{ت}$
١	١٦٠	$\frac{١}{٥}$	٨٠٠$(\frac{١}{٥})$ = ١٦٠	صفر	صفر
٢	١٣٠	$\frac{١}{٥}$	٨٠٠$(\frac{١}{٥})$ = ١٦٠	-٣٠	٥,٦٣
٣	١٥٠	$\frac{١}{٥}$	٨٠٠$(\frac{١}{٥})$ = ١٦٠	-١٠	٠,٦٣
٤	١٩٠	$\frac{١}{٥}$	٨٠٠$(\frac{١}{٥})$ = ١٦٠	٣٠	٥,٦٣
٥	١٧٠	$\frac{١}{٥}$	٨٠٠$(\frac{١}{٥})$ = ١٦٠	١٠	٠,٦٣
المجموع	٨٠٠	١,٠	٨٠٠	صفر	١٢,٥٢

أي أن قيمة إحصاء الاختبار المحسوبة = ١٢,٥٢

٥- اتخاذ القرار:

نظراً لأن قيمة إحصاء الاختبار المحسوبة (١٢,٥٢) أصغر من القيمة الحرجة لـ مربع كاي (١٣,٢٧٧)، فإن قيمة إحصاء الاختبار تقع في منطقة عدم الرفض. أي أننا لا نرفض فرضية العدم,

٦- عدم رفض فرضية العدم يفيد بأن العملاء يتوزعون بالتساوي على البائعين الخمسة وأن ادعاء مدير الفرع ليس هناك ما يؤكده.

التطبيق (٤٤)

في دراسة لتقييم فعالية حملة تطعيم ضد مرض شلل الأطفال، كانت النتائج التالية فيما يتعلق بعينة مكونة من ١٠٠ طفل.

الحالة	أُصيب بالمرض	لم يُصب بالمرض	المجموع
تم التطعيم	١٠	٥٠	٦٠
لم يُطعَّم	٣٥	٥	٤٠
المجموع	٤٥	٥٥	١٠٠

على ضـــوء نتائج العينة، هل يمكن القول بأن حملة التطعيم كانت فعَّالة؟ اســتخدم مســتوى المعنوية ٠,٠١.

الحـــل:

١- صياغة فرضية العدم والفرضية البديلة:

ف$_{٠}$: حملة التطعيم غير فعّالة، أي أن التطعيم ضد المرض ليس له علاقة بالإصابة بالمرض. وبعبارة أخرى، الإصابة بهذا المرض ليس لها علاقة بكون الطفل تم تطعيمه أم لا.

ف$_{١}$: حملة التطعيم فعّالة، أي أن الإصابة بالمرض من عدمها تتوقف على كون الطفل تم تطعيمه أم لا.

٢- اختيار التوزيع الملائم:

الاختبار هنا معني بوجود علاقة بين الإصابة بالمرض وبين التطعيم ضد المرض. لذلك فإن الاختبار الملائم هو اختبار مربع كاي للاستقلال.

٣- تحديد منطقتي الرفض وعدم الرفض:

في هذا المثال، نجد أن: ص = ٢ و د = ٢

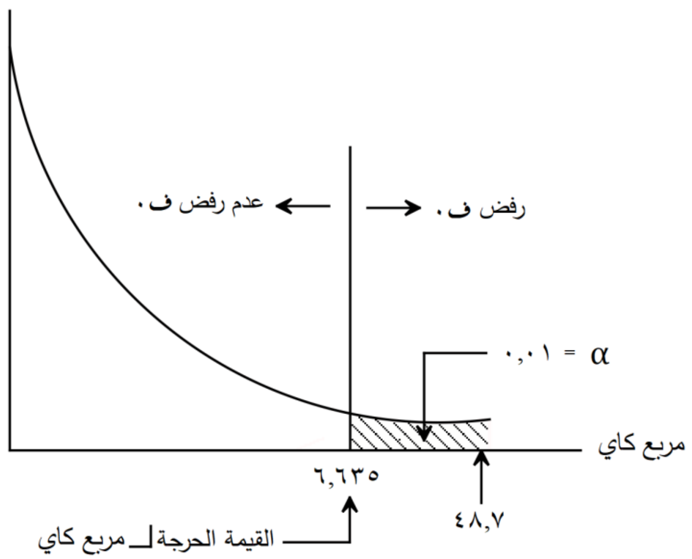
أي أن: درجات الحرية = (ص – ١)(د – ١)

= (٢ – ١)(٢ – ١) = درجة واحدة

وحيث أن مستوى المعنوية المطلوب هو α = ٠,٠١، فإنه بالكشف عن القيمة الحرجة لـ مربع كاي في جدول المساحة تحت منحنى توزيع مربع كاي بدرجات حرية ١ ومستوى معنوية (مساحة على الطرف الأيمن للمنحنى) ٠,٠١ نجد أن:

القيمة الحرجة لـ مربع كاي = ٦,٦٣٥

لذلك تتحدد منطقتي الرفض وعدم الرفض كما هو موضح في الشكل التالي:



٤- إيجاد قيمة إحصاء الاختبار:

نوجد أولاً التكرار المتوقَّع لكل خلية من خلايا جدول التوافق وذلك بافتراض صحة فرضية العدم.

الصف الأول:

التكرار المتوقَّع للخلية الأولى = $\frac{٦٠ \times ٤٥}{١٠٠} = ٢٧$

التكرار المتوقَّع للخلية الثانية = $\frac{٦٠ \times ٥٥}{١٠٠} = ٣٣$

الصف الثاني:

التكرار المتوقَّع للخلية الثالثة = $\frac{٤٠ \times ٤٥}{١٠٠} = ١٨$

التكرار المتوقَّع للخلية الرابعة = $\frac{٤٠ \times ٥٥}{١٠٠} = ٢٢$

والجدول التالي يوضح التكرار المشاهَد والتكرار المتوقَّع لكل خلية من خلايا الجدول.

الحالة	أصيب بالمرض	لم يصب بالمرض	المجموع
تم التطعيم	١٠ **(٢٧)**	٥٠ **(٣٣)**	٦٠
لم يُطعّم	٣٥ **(١٨)**	٥ **(٢٢)**	٤٠
المجمـــــوع	٤٥	٥٥	١٠٠

وبإيجاد قيمة مربع كاي باستخدام الصيغة (٤ – ١) نحصل على:

$$\text{مربع كاي} = \frac{(١٠ - ٢٧)^٢}{٢٧} + \frac{(٥٠ - ٣٣)^٢}{٣٣} + \frac{(٣٥ - ١٨)^٢}{١٨} + \frac{(٥ - ٢٢)^٢}{٢٢}$$

$$= ١٠,٧ + ٨,٨ + ١٦,١ + ١٣,١ = ٤٨,٧$$

كما يمكن إيجاد قيمة مربع كاي باستخدام الصيغة (٤ – ٣) وذلك على النحو التالي:

$$\text{مربع كاي} = \frac{(١٠ \times ٤٥ - ٣٥ \times ٥٠)^٢ (١٠٠)}{٦٠ \times ٤٠ \times ٤٥ \times ٥٥} = ٤٨,٧$$

وهي نفس النتيجة السابقة.

وهنا تجدر الإشارة إلى أن استخدام الصيغة (٤ – ٤) في هذه الحالة سوف يؤدي أيضاً إلى نفس النتيجة.

٥- اتخاذ القرار:

حيث أن قيمة إحصاء الاختبار (٤٨,٧) أكبر من القيمة الحرجة لـ مربع كاي (٦,٦٣٥)، فإن ذلك يشير إلى وقوع قيمة إحصاء الاختبار في منطقة الرفض. وهذا يعني رفض فرضية العدم وقبول الفرضية البديلة.

٦- رفض فرضية العدم وقبول الفرضية البديلة يمكِّننا من القول بأن حملة تطعيم الأطفال ضد مرض الشلل هي حملة فعالة. أي أن إصابة الطفل بهذا المرض تتوقف على كون هذا الطفل تم تطعيمه ضد الإصابة بهذا المرض أم لا. ومن جدول التوافق يمكننا استنتاج أن تطعيم الطفل ضد هذا المرض يُقلِّل من الإصابة به.

التطبيق (٤٥)

الجدول التالي يوضح تقديرات نجاح عينة من طلاب إحدى كليات التجارة حسب نظام الدراسة وذلك في أحد الأعوام الدراسية.

نظام الدراسة / تقدير النجاح	انتظام	انتساب موجه	المجموع
جيد فأقل	٦٥	١٣٠	١٩٥
جيد جداً	٢٠	١٨	٣٨
ممتـــــاز	١٥	٢	١٧
المجموع	١٠٠	١٥٠	٢٥٠

هل تدل هذه البيانات على أن توزيع الطلاب المنتظمين وتوزيع الطلاب المنتسبين على تقديرات النجاح المختلفة متشابهان؟ استخدم مستوى المعنوية ٠,٠٥. وبعبارة أخرى، هل يمكن القول بأن التقدير الذي يحصل عليه الطالب يتوقف على كونه طالباً منتظماً أو منتسباً؟

الحـــل:

١- صياغة فرضية العدم والفرضية البديلة:

$ف_٠$: توزيع الطلاب على تقديرات النجاح واحد في حالة الانتظام والانتساب الموجه.

$ف_١$: توزيع الطلاب على تقديرات النجاح يختلف في حالة الانتظام عنه في حالة الانتساب الموجه.

٢- اختيار التوزيع الملائم:

نحن هنا بصدد بحث تجانس توزيعين للطلاب. وبالتالي فإن الاختبار الملائم هو اختبار مربع كاي للتجانس.

٣- تحديد منطقتي الرفض وعدم الرفض:

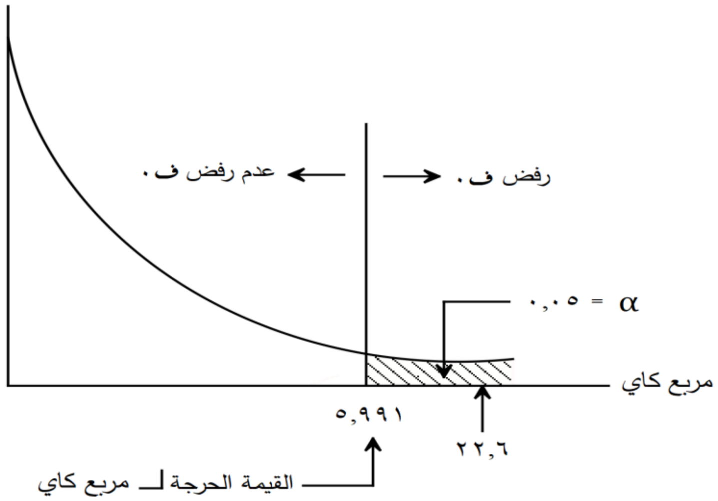
في هذا المثال: ص = ٣ و د = ٢

أي أن: درجات الحرية = (ص – ١)(د – ١)

= (٣ – ١)(٢ – ١) = ٢

وبالبحث في جدول المساحة تحت منحنى توزيع مربع كاي (الملحق ٣) نجد أن القيمة الحرجة لـ مربع كاي عند درجات حرية ٢ ومستوى معنوية (مساحة على الطرف الأيمن للمنحنى) ٠,٠٥ هي ٥,٩٩١.

لذلك فإن منطقتي الرفض وعدم الرفض تتحددان كما هو موضح في الشكل التالي:



٤- إيجاد قيمة إحصاء الاختبار:

بحساب التكرار المتوقَّع لكل خلية من خلايا جدول التوافق بافتراض صحة فرضية العدم نحصل على ما يلي:

الصف الأول:

التكرار المتوقَّع للخلية الأولى = $\frac{١٩٥ \times ١٠٠}{٢٥٠}$ = ٧٨

التكرار المتوقَّع للخلية الثانية = $\frac{١٩٥ \times ١٥٠}{٢٥٠}$ = ١١٧

الصف الثاني:

التكرار المتوقَّع للخلية الثالثة = $\frac{٣٨ \times ١٠٠}{٢٥٠}$ = ١٥,٢

التكرار المتوقَّع للخلية الرابعة = $\frac{٣٨ \times ١٥٠}{٢٥٠}$ = ٢٢,٨

الصف الثالث:

$$\text{التكرار المتوقَّع للخلية الخامسة} = \frac{١٧ \times ١٠٠}{٢٥٠} = ٦,٨$$

$$\text{التكرار المتوقَّع للخلية السادسة} = \frac{١٧ \times ١٥٠}{٢٥٠} = ١٠,٢$$

والجدول التالي يوضح التكرار المشاهَد والتكرار المتوقَّع لكل خلية من خلايا جدول التوافق.

تقدير النجاح / نظام الدراسة	انتظام	انتساب	مجموع
جيد فأقل	٦٥ **(٧٨)**	١٣٠ **(١١٧)**	١٩٥
جيد جداً	٢٠ **(١٥,٢)**	١٨ **(٢٢,٨)**	٣٨
امتيـــاز	١٥ **(٦,٨)**	٢ **(١٠,٢)**	١٧
المجمــــوع	١٠٠	١٥٠	٢٥٠

وبإيجاد قيمة إحصاء الاختبار باستخدام الصيغة (٤ ــ ١) نحصل على:

$$\text{مربع كاي} = \frac{(٦٥ - ٧٨)^{٢}}{٧٨} + \frac{(١٣٠ - ١١٧)^{٢}}{١١٧} + \frac{(٢٠ - ١٥,٢)^{٢}}{١٥,٢}$$

$$+ \frac{(١٨ - ٢٢,٨)^{٢}}{٢٢,٨} + \frac{(١٥ - ٦,٨)^{٢}}{٦,٨} + \frac{(٢ - ١٠,٢)^{٢}}{١٠,٢}$$

$$= ٢,٢ + ١,٤ + ١,٥ + ١,٠ + ٩,٩ + ٦,٦ = ٢٢,٦$$

وباستخدام الصيغة (٤ ــ ٤) في حساب قيمة إحصاء الاختيار نجد أن:

$$\text{مربع كاي} = \left[\frac{(٦٥)^{٢}}{٧٨} + \frac{(١٣٠)^{٢}}{١١٧} + \frac{(٢٠)^{٢}}{١٥,٢} + \frac{(١٨)^{٢}}{٢٢,٨} + \frac{(١٥)^{٢}}{٦,٨} + \frac{(٥٠)^{٢}}{٤٥,٦}\right] - ٢٥٠$$

$$= (٥٤,٦ + ١٤٤,٤ + ٢٦,٣ + ١٤,٢ + ٣٣,١ + ٠,٤) - ٢٥٠$$

$$= ٢٧٢,٦ - ٢٥٠ = ٢٢,٦$$

وهي نفس النتيجة التي حصلنا عليها باستخدام الصيغة (٤ ــ ١) من قبل.

٥- اتخاذ القرار:

نظراً لأن قيمة إحصاء الاختبار المحسوبة (٢٢,٦) أكبر من القيمة الحرجة لـ مربع كاي (٥,٩٩١) فهي تقع في منطقة الرفض. أي أننا نرفض فرضية العدم ونقبل الفرضية البديلة.

٦- رفض فرضية العدم وقبول الفرضية البديلة يدل على أن توزيع الطلاب على تقديرات النجاح في حالة الطلاب المنتظمين يختلف عن نظيره في حالة الطلاب المنتسبين. أي أن توزيع الطلاب على تقديرات النجاح المختلفة ليس واحداً في الحالتيْن.

التطبيق (٤٦)

الجدول التالي يوضح التوزيع التكراري لعينة مكونة من ١٥٠ طالباً متغيباً عن محاضرات إحدى المواد في إحدى الكليات موزعة حسب أيام الأسبوع.

اليوم	السبت	الأحد	الإثنين	الثلاثاء	الأربعاء	الخميس	المجموع
عدد الغائبين	٣٠	٢٥	١٨	١٦	٢٦	٣٥	١٥٠

هل يمكن القول بأن غياب الطلاب موزع بالتساوي على أيام الأسبوع؟

استخدم مستوى المعنوية ٠,٠١.

الحـــل:

١- صياغة فرضية العدم والفرضية البديلة:

ف.: غياب الطلاب موزع بالتساوي على أيام الأسبوع. أي أن غياب الطلاب لا يختلف من يوم لآخر.

ف١: غياب الطلاب يختلف من يوم لآخر.

٢- اختيار التوزيع الملائم:

هذا الاختبار يبحث في تجانس الأيام من حيث غياب الطلاب عن الالتحاق بإحدى المحاضرات أو غيابهم عنها. لذلك فإن الاختبار الملائم هو اختبار مربع كاي للتجانس.

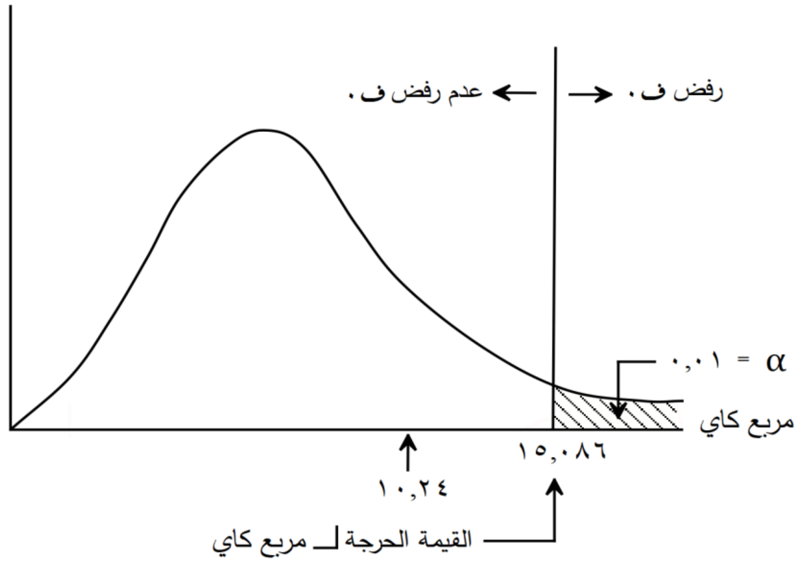
٣- تحديد منطقتي الرفض وعدم الرفض:

عدد الأيام = ٦ أيام

أي أن: ك = ٦

وبالتالي نجد أن: درجات الحرية = ك – ١ = ٦ – ١ = ٥

وبالبحث في جدول المساحة تحت منحنى توزيع مربع كاي عن القيمة الحرجة لـ مربع كاي بدرجات حرية ٥ ومستوى معنوية (مساحة على الطرف الأيمن للمنحنى) ٠,٠١ نجد أن: القيمة الحرجة لـ مربع كاي = ١٥,٠٨٦
لذلك فإن منطقتي الرفض وعدم الرفض تتحدان كما هو موضح في الشكل التالي:



٤- إيجاد قيمة إحصاء الاختبار:

يمكن إيجاد قيمة إحصاء الاختبار بافتراض صحة فرضية العدم وذلك على النحو التالي:
حيث أن فرضية العدم تنص على أن غياب الطلاب موزع بالتساوي على أيام الأسبوع، فإن ذلك يعني أن:

$$\text{احتمال غياب الطالب في أي يوم من أيام الأسبوع} = \frac{١}{\text{عدد أيام المحاضرات}} = \frac{١}{٦}$$

والجدول في الصفحة التالية يوضح الحسابات اللازمة لإيجاد التكرار المتوقَّع لكل يوم. ومن هذا الجدول نجد أن قيمة إحصاء الاختبار ١٠,٢٤.

٥- اتخاذ القرار:

حيث أن قيمة إحصاء الاختبار المحسوبة (١٠,٢٤) تقل عن القيمة الحرجة لـ مربع كاي (١٥,٠٨٦) فإنها تقع في منطقة عدم الرفض. أي أن البيانات المتوافرة لا تمكننا من رفض فرضية العدم.

٦- عدم رفض فرضـــية العدم يعني بأنه يمكننا القول بأن الغياب موزع بالتســـاوي على أيام الأسبوع.

حساب قيمة مربع كاي لبيانات التطبيق (٤٦)

اليوم	التكرار المشاهَد (ش)	الاحتمال (ع)	التكرار المتوقَّع ت = ن ع	ش – ت	$\frac{(ش - ت)^{٢}}{ت}$
السبت	٣٠	$\frac{١}{٦}$	$١٥٠(\frac{١}{٦}) = ٢٥$	٥	١,٠
الأحد	٢٥	$\frac{١}{٦}$	$١٥٠(\frac{١}{٦}) = ٢٥$	صفر	صفر
الإثنين	١٨	$\frac{١}{٦}$	$١٥٠(\frac{١}{٦}) = ٢٥$	٧-	١,٩٦
الثلاثاء	١٦	$\frac{١}{٦}$	$١٥٠(\frac{١}{٦}) = ٢٥$	٩-	٣,٢٤
الأربعاء	٢٦	$\frac{١}{٦}$	$١٥٠(\frac{١}{٦}) = ٢٥$	١	٠,٠٤
الخميس	٣٥	$\frac{١}{٦}$	$١٥٠(\frac{١}{٦}) = ٢٥$	١٠	٤,٠٠
المجموع	١٥٠	١,٠	١٥٠	صفر	١٠,٢٤

تطبيقات
الفصل الخامس
(الإحصاءات الحيوية والصحية)

التطبيق (٤٧)

إذا توافرت البيانات التالية عن نتائج تعداد تم إجراؤه في عام ٢٠١٦ في إحدى الدول:

- عدد السكان = ٤٠ مليون
- عدد المواليد أحياء خلال عام ٢٠١٦ = ١,١ مليون مولود حي.
- نسبة النساء اللاتي في سن الحمل = ٣٢٪ من عدد السكان.
- نسبة المتزوجات اللاتي في سن الحمل = ٧٥٪ من إجمالي عدد النساء اللاتي في سن الحمل.

المطلوب: حساب المعدلات التالية:

أ- معدل المواليد الخام **ب-** معدل الخصوبة العام **ج-** معدل التوالد العام

مع شرح ما يعنيه كل معدل مستعيناً بالنتائج التي حصلت عليها.

الحـــل:

١- معدل المواليد الخام = $\frac{\text{عدد المواليد الأحياء خلال العام}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$

$= \frac{١١٠٠٠٠٠}{٤٠٠٠٠٠٠٠} \times ١٠٠٠ = ٢٧,٥$ في الألف

أي أن هناك ٢٧,٥ مولود حي في المتوسط لكل ألف من السكان.

٢- معدل الخصوبة العام = $\frac{\text{عدد المواليد الأحياء خلال السنة}}{\text{عدد النساء في سن الحمل في منتصف العام}} \times ١٠٠٠$

وحيث أن:

عدد النساء في سن الحمل = $٤٠٠٠٠٠٠٠ \times \frac{٣٢}{١٠٠} = ١٢٨٠٠٠٠٠$ أنثى

$\therefore$ معدل الخصوبة العام = $\frac{١١٠٠٠٠٠}{١٢٨٠٠٠٠٠} \times ١٠٠٠ \simeq ٨٦$ في الألف

أي أن كل ألف امرأة في سن الحمل تضيف ٨٦ مولوداً جديداً في عام ٢٠١٦.

٣- معدل التوالد العام = $\frac{\text{عدد المواليد الأحياء خلال العام}}{\text{عدد المتزوجات في سن الحمل في منتصف العام}} \times ١٠٠٠$

وحيث أن:

عدد المتزوجات في سن الحمل = $١٢٨٠٠٠٠٠ \times \frac{٧٥}{١٠٠} = ٩٦٠٠٠٠٠$ امرأة

∴ معدل التوالد العام = $\frac{١١٠٠٠٠٠}{٩٦٠٠٠٠٠}$ × ١٠٠٠ ≃ ١١٥ في الألف

أي أن هناك ١١٥ مولوداً جديداً لكل ألف أنثى متزوجة وفي سن الحمل في عام ٢٠١٦.

التطبيق (٤٨)

فيما يلي البيانات الخاصة بالتوزيع العمري لسكان إحدى المدن (ذكوراً وإناثاً) وكذلك عدد المواليد وعدد الوفيات حسب الجنس وذلك في سنة ما:

فئات الأعمار	عدد السكان(بالألف)		عدد المواليد أحياء حسب عمر الأم		عدد الوفيات	
	ذكور	إناث	ذكور	إناث	ذكور	إناث
صفر –	٦	٩	–	–	٢٠٠	٢٥٠
١ –	٢٦	٢٥	–	–	٥٠٠	٥٥٠
٥ –	٢٢	٢٦	–	–	١٢٠	١٣٠
١٠ –	٢٦	٢٨	–	–	٧٠	٨٠
١٥ –	٢٤	٢٥	٨٠	٧٠	١٢٥	١٢٥
٢٠ –	٣٢	٣٤	٣٠٥٠	٣٠٨٠	١٨٠	٢٢٠
٣٠ –	٢٦	٢٨	٢١٢٠	٢١٠٠	٢٠٠	٢٥٠
٤٠ –	٢٢	٢٧	١٣٥٠	١٣٠٠	٢٣٠	٢٧٠
٥٠ – فأكثر	١٦	١٨	–	–	٥٠٠	٤٠٠
المجموع	٢٠٠	٢٢٠	٦٦٠٠	٦٥٥٠	٢١٢٥	٢٢٧٥

المطلــوب: حساب المعدلات التالية:

١- معدل المواليد الخام

٢- معدل الخصوبة العام

٣- معدل الخصوبة لفئه السن (٢٠–٣٠) في المتوسط

٤- معدل الخصوبة الأنثوية لفئة السن (١٠ – ١٥)

٥- معدل الخصوبة الكلى

٦- معدل التوالد العام: بافتراض أن المتزوجات يمثلن ٩٥٪ من إجمالي عدد النساء في كل فئة من فئات سن الحمل.

٧- معدل التكاثر الإجمالي

٨- معدل الوفيات الخام

٩- معدل الوفيات في فئة العمر (٣٠ – ٤٠).
١٠- معدل وفيات كل من الذكور والإناث في فئة العمر (١٥ –٢٠) مع المقارنة بينهما.
١١- معدل وفيات الأطفال الرُضَّع.
١٢- معدل وفيات الأمومة، وذلك بافتراض أن عدد وفيات النساء أثناء الحمل والولادة كان ٤٠ حالة وفاة.

الحـــل:

١- معدل المواليد الخام:

$$\text{معدل المواليد الخام} = \frac{\text{عدد المواليد الأحياء خلال العام}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$$

$$= \frac{٦٦٠٠ + ٦٥٥٠}{٢٠٠٠٠٠ + ٢٢٠٠٠٠} \times ١٠٠٠$$

$$= \frac{١٣١٥٠}{٤٢٠٠٠٠} \times ١٠٠٠ \simeq ٣١ \text{ في الألف}$$

بما يعني أن هناك ٣١ مولوداً حياً في المتوسط لكل ألف من السكان.

٢- معدل الخصوبة العام:

$$\text{معدل الخصوبة العام} = \frac{\text{عدد المواليد الأحياء خلال السنة}}{\text{عدد النساء في سن الحمل في منتصف العام}} \times ١٠٠٠$$

ومن الجدول نجد أن:

عدد النساء في سن الحمل = (٢٥ + ٣٤ + ٢٨ + ٢٧) × ١٠٠٠ = ١١٤٠٠٠ أنثى

$$\therefore \text{معدل الخصوبة العام} = \frac{١٣١٥٠}{١١٤٠٠٠} \times ١٠٠٠ \simeq ١١٥ \text{ في الألف}$$

أي أن هناك ١١٥ مولوداً حياً لكل ألف أنثى في سن الحمل في هذا العام.

٣- معدل الخصوبة لفئة السن (٢٠ – ٣٠) في المتوسط:

معدل الخصوبة لفئة السن (٢٠ – ٣٠) في المتوسط:

$$= \frac{\text{عدد المواليد الأحياء في هذه الفئة}}{\text{عدد النساء في هذه الفئة}} \times ١٠٠٠$$

$$= \frac{٣٠٥٠ + ٣٠٨٠}{٣٤٠٠٠} \times ١٠٠٠ \simeq ١٨٠ \text{ في الألف}$$

وهذا يدل على أن هناك ١٨٠ مولوداً حياً لكل ألف أنثى يتراوح عمرها ما بين ٢٠ إلى أقل من ٣٠.

٤- معدل الخصوبة الأنثوية لفئة السن (١٠ – ١٥):

معدل الخصوبة الأنثوية لفئة السن (١٠ –١٥)

$$= \frac{\text{عدد المواليد الإناث الأحياء لهذه الفئة}}{\text{عدد النساء في هذه الفئة}} \times ١٠٠٠$$

$$= \frac{\text{صفر}}{٢٨٠٠٠} = \text{صفر}$$

أي أن الإناث اللاتي تتراوح أعمارهم ما بين ١٠ إلى أقل من ١٥ سنة لا يُضِفْن أي مواليد جديدة إلى المجتمع في هذا العام. وهذا أمر طبيعي إذ أنه من المُفترض أن الإناث في هذه الفئة من السن لم يتزوجن بعد.

٥- معدل الخصوبة الكلى:

$$= \sum\left(\frac{\text{عدد المواليد الأحياء للفئة}}{\text{عدد النساء في هذه الفئة}} \times \text{طول الفئة}\right) \times ١٠٠٠$$

$$= \frac{٧٠ + ٨٠}{٢٥٠٠٠} \times ٥ \times ١٠٠٠ + \frac{٣٠٨٠ + ٣٠٥٠}{٣٤٠٠٠} \times ١٠ \times ١٠٠٠ +$$

$$\frac{٢١٠٠ + ٢١٢٠}{٢٨٠٠٠} \times ١٠ \times ١٠٠٠ + \frac{١٣٠٠ + ١٣٥٠}{٢٧٠٠٠} \times ١٠ \times ١٠٠٠$$

= ٣٠ + ١٨٠٢,٩ + ١٥٠٧,١ + ٩٨١,٥ ≃ ٤٣٢٢ في الألف

$\therefore$ معدل الخصوبة الكلي = $\frac{٤٣٢٢}{١٠٠٠}$ = ٤,٣ ≃ ٤ لكل أم

وهذا يعني أن كل أم متزوجة تظل على قيد الحياة وتمارس نفس معدلات الإنجاب الواردة في الجدول سوف تنجب في المتوسط ٤ أطفال أحياء خلال فترة العمر (١٥ إلى أقل من ٥٠ سنة).

٦- معدل التوالد العام:

$$\text{معدل التوالد العام} = \frac{\text{عدد المواليد الأحياء خلال العام}}{\text{عدد المتزوجات في سن الحمل في منتصف العام}} \times ١٠٠٠$$

وحيث أن:

عدد المتزوجات في سن الحمل

$$= (٢٥٠٠٠ + ٣٤٠٠٠ + ٢٨٠٠٠ + ٢٧٠٠٠) \times \frac{٩٥}{١٠٠} = ١٠٨٣٠٠ \text{ امرأة}$$

$\therefore$ معدل التوالد العام = $\frac{٦٥٥٠ + ٦٦٠٠}{١٠٨٣٠٠} \times ١٠٠٠$ ≃ ١٢١ في الألف

أي أن كل ألف أنثى متزوجة في سن الحمل تضيف ١٢١ مولوداً حياً في هذا العام.

٧- معدل التكاثر الإجمالي:

$$= \sum \left(\frac{\text{عدد المواليد الأحياء من الإناث في الفئة}}{\text{عدد المتزوجات في هذه الفئة}} \times \text{طول الفئة} \right) \times \text{١٠٠٠}$$

$$= \frac{\text{٧٠}}{\text{٢٥٠٠٠} \times \text{٠,٩٥}} \times \text{٥} \times \text{١٠٠٠} + \frac{\text{٣٠٨٠}}{\text{٣٤٠٠٠} \times \text{٠,٩٥}} \times \text{١٠} \times \text{١٠٠٠} +$$

$$\frac{\text{٢١٠٠}}{\text{٢٨٠٠٠} \times \text{٠,٩٥}} \times \text{١٠} \times \text{١٠٠٠} + \frac{\text{١٣٠٠}}{\text{٢٧٠٠٠} \times \text{٠,٩٥}} \times \text{١٠} \times \text{١٠٠٠}$$

$$= \text{١٤,٧} + \text{٩٥٣,٦} + \text{٧٨٩,٥} + \text{٥٠٦,٨} \simeq \text{٢٢٦٥ في الألف}$$

أي أن عدد الأطفال الإناث الذي يمكن أن تنجبه ألف سيده تبقين على قيد الحياة هو ٢٢٦٥ طفلاً أنثى. أي بمعدل طفلان تقريباً لكل أنثى.

٨- معدل الوفيات الخام:

$$\text{معدل الوفيات الخام} = \frac{\text{عدد الوفيات خلال العام}}{\text{عدد السكان في منتصف العام}} \times \text{١٠٠٠}$$

$$= \frac{\text{٢١٢٥} + \text{٢٢٧٥}}{\text{٢٠٠٠٠٠} + \text{٢٢٠٠٠٠}} \times \text{١٠٠٠}$$

$$= \frac{\text{٤٤٠٠}}{\text{٤٢٠٠٠٠}} \times \text{١٠٠٠} \simeq \text{١٠ في الألف}$$

أي أن هناك ١٠ حالات وفاة لكل ألف من السكان في هذا العام.

٩- معدل الوفيات في فئة العمر (٣٠ – ٤٠):

$$= \frac{\text{عدد الوفيات في هذه الفئة}}{\text{عدد السكان في هذه الفئة}} \times \text{١٠٠٠}$$

$$= \frac{\text{٢٠٠} + \text{٢٥٠}}{\text{٢٦٠٠٠} + \text{٢٨٠٠٠}} \times \text{١٠٠٠} \simeq \text{٨ في الألف}$$

أي أن هناك ٨ حالات وفاة بين كل ألف من السكان الذين تتراوح أعمارهم ما بين ٣٠ سنة إلى أقل من ٤٠ سنة.

١٠- معدل وفيات الذكور والإناث في فئة العمر (١٥ – ٢٠)

معدل وفيات الذكور في فئة العمر (١٥ – ٢٠):

$$= \frac{\text{عدد وفيات الذكور في هذه الفئة}}{\text{عدد الذكور في هذه الفئة}} \times \text{١٠٠٠}$$

$$= \frac{١٢٥}{٢٤٠٠} \times ١٠٠٠ \simeq ٥,٢ \text{ في الألف}$$

بما يعني أن هناك ٥,٢ وفيات ذكور في المتوسط لكل ألف من الذكور في هذه الفئة.

معدل وفيات الإناث في فئة العمر (١٥ ــ ٢٠):

$$= \frac{\text{عدد وفيات الإناث في هذه الفئة}}{\text{عدد الإناث في هذه الفئة}} \times ١٠٠٠$$

$$= \frac{١٢٥}{٢٥٠٠٠} \times ١٠٠٠ \simeq ٥ \text{ في الألف}$$

وبمقارنة المعدلين نجد أن الذكور في فئة العمر (١٥ ــ٢٠) هم أكثر قليلاً من الإناث من حيث التعرض للوفاة.

١١- معدل وفيات الأطفال الرضع:

$$\text{معدل وفيات الأطفال الرضع} = \frac{\text{عدد وفيات الأطفال دون العام}}{\text{عدد المواليد أحياء}} \times ١٠٠٠$$

$$= \frac{٢٥٠ + ٢٠٠}{٦٥٥٠ + ٦٦٠٠} \times ١٠٠٠$$

$$= \frac{٤٥٠}{١٣١٥٠} \times ١٠٠٠ \simeq ٣٤ \text{ في الألف}$$

أي أن هناك ٣٤ حالة وفاة للأطفال قبل بلوغهم العام الأول وذلك من بين كل ألف مولود حي في هذا العام.

١٢- معدل وفيات الأمومة:

$$= \frac{\text{عدد وفيات النساء أثناء الحمل و الولادة}}{\text{عدد المواليد أحياء خلال العام}} \times ١٠٠٠$$

$$= \frac{٤٠}{٦٥٥٠ + ٦٦٠٠} \times ١٠٠٠ \simeq ٣ \text{ في الألف}$$

وهذا يدل على أنه من بين كل ألف امرأة في حالة حمل أو ولادة هناك ٣ حالات وفاة خلال هذا العام.

التطبيق (٤٩)

إذا توافرت بيانات الزواج والطلاق التالية عن إحدى المدن في سنة ما:

- عدد حالات الزواج التي تمت خلال السنة = ١٢٠٠٠ حالة زواج.
- تقدير عدد السكان في منتصف السنة = ٢ مليون نسمة.
- عدد السكان في سن الزواج = ٨٠٠٠٠٠ شخص.
- أن المتزوجين يمثلون ٦٥٪ من عدد السكان في سن الزواج.
- عدد حالات الطلاق التي تمت خلال السنة = ٢٠٠٠ حالة طلاق.
- عدد حالات المتزوجين خلال السنة يمثل ٥٪ من إجمالي عدد المتزوجين.
- عدد المطلقين خلال السنة = ٤٠٠٠ شخص.

المطلـــوب: احسب المعدلات التالية

أ- معدل الزواج الخام

ب- معدل الزواج المعدل

ج- معدل الطلاق الخام

د- معدل الطلاق المُعَدل

هـ - معدل الطلاق الخاص

مع شرح النتائج التي تحصل عليها.

الحـــل:

أ- معدل الزواج الخام = $\frac{\text{عدد الزيجات خلال السنة}}{\text{عدد السكان في منتصف السنة}} \times ١٠٠٠$

$= \frac{١٢٠٠٠}{٢٠٠٠٠٠٠} \times ١٠٠٠ = ٦$ في الألف

أي أن هناك ٦ حالات زواج بين كل ألف من السكان في هذه السنة.

ب- معدل الزواج المعدل = $\frac{\text{عدد الزيجات خلال السنة}}{\text{عدد السكان في سن الزواج و غير متزوجين منتصف السنة}} \times ١٠٠٠$

وحيث أن:

عدد السكان في سن الزواج وغير متزوجين = $٨٠٠٠٠٠ \times \frac{٣٥}{١٠٠} = ٢٨٠٠٠٠$

∴ معدل الزواج المعدل = $\frac{١٢٠٠٠}{٢٨٠٠٠٠} \times ١٠٠٠ \simeq ٤٣$ في الألف

أي أن هناك ٤٣ حالة زواج بين كل ألف من السكان الذين هم في سن الزواج وغير متزوجين.

ج- معدل الطلاق الخام = $\frac{\text{عدد حالات الطلاق خلال العام}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$

$= \frac{٢٠٠٠}{٢٠٠٠٠٠٠} \times ١٠٠٠ = ١$ في الألف

أي أن هناك حالة طلاق واحدة بين كل ألف من السكان.

د- معدل الطلاق المعدَّل = $\frac{\text{عدد حالات الطلاق أثناء السنة}}{\text{عدد غير المتزوجين في منتصف السنة}} \times ١٠٠٠$

وحيث أن:

عدد المتزوجين = $٨٠٠٠٠٠ \times \frac{٦٥}{١٠٠} = ٥٢٠٠٠٠$ شخص

∴ معدل الطلاق المعدل = $\frac{٢٠٠٠}{٥٢٠٠٠٠} \times ١٠٠٠ \simeq ٤$ في الألف

أي أن هناك ٤ حالات طلاق بين كل ألف من المتزوجين خلال هذه السنة.

هـ - معدل الطلاق الخاص = $\frac{\text{عدد الذين طلقوا خلال السنة}}{\text{عدد المتزوجين فعلاً خلال السنة}} \times ١٠٠٠$

وحيث أن:

عدد المتزوجين خلال السنة = $٥٢٠٠٠٠ \times \frac{٥}{١٠٠} = ٢٦٠٠٠$ شخصاً

∴ معدل الطلاق الخاص = $\frac{٤٠٠٠}{٢٦٠٠٠} \times ١٠٠٠ \simeq ١٥٤$ في الألف

أي هنا ١٥٤ شخصاً شملتهم حالات الطلاق وذلك من بين كل ألف تم زواجهم خلال هذه السنة.

التطبيق (٥٠)

بافتراض أن مرض الجذُري قد انتشر في إحدى الدول في سنة ما فظهر أن عدد الأفراد المصابين بالمرض هو ٢٠٠٠ شخص، وأن عدد الذين ماتوا متأثرين بالمرض هو ١٢٠ شخصاً. فإذا كان تقدير عدد السكان في هذه السنة هو ١٢ مليون نسمة، وأنه عند إجراء الفحص الوقائي تَبيَّن أن هناك ٣٠٠ شخص حاملين للمرض وأن عدد المعرضين للخطر هو ٢٥٠٠٠٠ فردا.

المطلـــــوب: حساب المعدلات التالية

أ- معدل الإصابة بالمرض ب- معدل وفيات المرض

ج- معدل حاملي المرض د- معدل المعرضين للخطر

هـ- معدل انتشار المرض، وذلك بافتراض أن هناك ١٠٠ حالة إصابة قديمة بهذا المرض.

الحــــل:

أ- معدل الإصابة بالمرض = $\frac{\text{عدد الأشخاص الذين أصيبوا بالمرض خلال السنة}}{\text{عدد السكان في منتصف السنة}}$ × ١٠٠٠

= $\frac{٢٠٠٠}{١٢٠٠٠٠٠٠٠}$ × ١٠٠٠ = ٠,١٧ في الألف

أي أن هناك ١٧ شخصاً مصاباً بهذا المرض من بين كل مائة ألف من السكان وذلك في هذه السنة.

ب- معدل وفيات المرض = $\frac{\text{عدد المتوفين بسبب المرض}}{\text{عدد المصابين بهذا المرض}}$ × ١٠٠٠

= $\frac{١٢٠}{٢٠٠٠}$ × ١٠٠٠ = ٦٠ في الألف

أي أن هناك ٦٠ حالة وفاة بين كل ألف مصاب بهذا المرض في هذه السنة.

ج- معدل وفيات المرض = $\frac{\text{عدد حاملي المرض}}{\text{عدد السكان في منتصف العام}}$ × ١٠٠٠

= $\frac{٣٠٠}{١٢٠٠٠٠٠٠٠}$ × ١٠٠٠ = ٠,٠٢٥ في الألف

أي أن هناك ٢٥ شخصاً حاملين للمرض من بين كل مليون من السكان.

د- معدل وفيات المرض = $\frac{\text{عدد المعرضين للخطر}}{\text{عدد السكان}}$ × ١٠٠٠

= $\frac{٢٥٠٠٠٠٠}{١٢٠٠٠٠٠٠٠}$ × ١٠٠٠ $\simeq$ ٢١ في الألف

وهذا يعني أن هناك ٢١ شخصاً معرضاً للإصابة بهذا المرض من بين كل ألف من السكان.

هـ- معدل انتشار المرض = $\frac{\text{عدد الإصابات (الجديدة + القديمة)}}{\text{عدد السكان في منتصف العام}}$ × ١٠٠٠

= $\frac{١٠٠ + ٢٠٠٠}{١٢٠٠٠٠٠٠٠}$ × ١٠٠٠ = ٠,١٧٥ في الألف

أي أن هناك ١٧٥ شخصاً مصاباً بالمرض من بين كل مليون من السكان.

التطبيق (٥١)

إذا توافرت البيانات التالية عن أحد المستشفيات في بلد ما في عام ٢٠١٧.

- عدد سكان البلد = ٢٠٠٠٠ نسمة
- عدد الأسِرَّة في المستشفى = ٢٧٠ سريراً
- عدد أيام التمريض = ٣٦٠٠٠ يوم
- عدد حالات الخروج = ٤٨٠٠
- عدد الأطباء = ١٢٠ طبيباً
- عدد أفراد الهيئة التمريضية = ٣٦٠ فرداً

المطلـــوب: إيجاد المؤشرات التالية

– نسبة إشغال الأسِرَّة.
– النسبة المئوية لخلو السرير.
– معدل الأسِرَّة للسكان.
– متوسط مدة البقاء.
– معدل دوران السرير.
– معدل الأطباء لعدد السكان.
– معدل التمريض لعدد السكان.
– فترة خلو السرير.

الحـــل:

– نسبة إشغال الأسِرَّة:

$$\text{نسبة إشغال الأسِرّة} = \frac{\text{عدد أيام التمريض خلال السنة}}{\text{مجموع أيام التمريض خلال السنة}} \times ١٠٠$$

$$= \frac{\text{عدد أيام التمريض خلال عام ٢٠١٧}}{\text{عدد الأسِرّة} \times ٣٦٥} \times ١٠٠$$

$$= \frac{٣٦٠٠٠}{٢٥٠ \times ٣٦٥} \times ١٠٠ = ٣٩,٥٪$$

أي أنه خلال عام ٢٠١٧ تم إشغال ٣٩,٥٪ فقط من إجمالي عدد الأسّرة في هذا المستشفى.

– النسبة المئوية لخلو السرير:

النسبة المئوية لخلو السرير = ١٠٠٪ – نسبة إشغال الأسِرّة

= ١٠٠٪ – ٣٩,٥٪ = ٦٠,٥٪

أي أنه خلال عام ٢٠١٧ كانت نسبة الأسِرّة الشاغرة ٦١,٥٪.

– معدل الأسِرّة للسكان:

$$\text{نسبة الأسِرّة للسكان} = \frac{\text{عدد الأسِرة في المستشفى}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$$

$$= \frac{٢٥٠}{٢٠٠٠٠} \times ١٠٠٠ = ١٢,٥ \text{ في الألف}$$

وهذا يدل على أن هناك ١٢٥ سريراً لكل ١٠ آلاف من السكان في هذا العام ٢٠١٧.

– متوسط مدة البقاء:

$$\text{متوسط مدة البقاء} = \frac{\text{مجموع أيام التمريض خلال السنة}}{\text{عدد المرضى الذين خرجوا خلال السنة}} = \frac{٣٦٠٠٠}{٤٨٠٠} = ٧,٥ \text{ يوم}$$

مما يعني أن كل مريض غادر المستشفى خلال هذا العام تم تنويمه في المستشفى لمدة سبعة أيام ونصف اليوم في المتوسط.

– معدل دوران السرير:

$$\text{معدل دوران السرير} = \frac{\text{مجموع حالات الخروج خلال العام}}{\text{عدد الأسِرّة خلال العام}} \times ١٠٠٠$$

$$= \frac{٤٨٠٠}{٢٥٠} \simeq ١٩$$

ويفيد ذلك بأن كل سرير في المستشفى تم إشغاله خلال العام بواسطة ١٩ مريضاً في المتوسط.

– معدل الأطباء لعدد السكان:

$$\text{معدل الأطباء لعدد السكان} = \frac{\text{عدد الأطباء خلال العام}}{\text{عدد السكان في منتصف العام}} \times ١٠٠٠$$

$$= \frac{١٢٠}{٢٠٠٠٠} \times ١٠٠٠ = ٦ \text{ في الألف}$$

أي أن هناك ٦ أطباء لكل ألف من السكان.

– معدل التمريض لعدد السكان:

$$\text{معدل التمريض لعدد السكان} = \frac{\text{عدد أفراد الهيئة التمريضية}}{\text{عدد السكان}} \times ١٠٠٠$$

$$= \frac{٣٦٠}{٢٠٠٠٠} \times ١٠٠٠ = ١٨ \text{ في الألف}$$

مما يدل على أن هناك ١٨ فرد تمريض لكل ألف من السكان.

– فترة خلو السرير:

$$\text{فترة خلو السرير} = \frac{\text{عدد الأسِرّة} \times \text{عدد الأيام} - \text{مجموع أيام التمريض}}{\text{عدد حالات الخروج}}$$

$$= \frac{٢٥٠ \times ٣٦٥ - ٣٦٠٠٠}{٤٨٠٠} = ١١,٥ \text{ يوماً}$$

وهذا يعني أن المدة بين ترك المريض للسرير الذي يشغله وبين دخول مريض آخر في نفس السرير هي ١١,٥ يوماً في المتوسط.

وفي النهاية، نأمل أ، تعود هذه التطبيقات بالنفع على القارئ خاصة هؤلاء الذين سوف يعملون بعد التخرج – إن شاء الله – في أي من تلك المجالات.

امتحانات

سنوات سابقة

الامتحان رقم (١)

المادة: إحصاء تطبيقي **الفرقة: الثالثة**

الكلية: التجارة بسوهاج **دور: مايو ١٩٩٧م**

طلاب الدراسة باللغة العربية

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: إذا كان احتمال نجاح أحمد في مادة الإحصاء التطبيقي هو ٠,٨، واحتمال نجاح أحمد ومجدي معاً في نفس المادة هو ٠,٦ فإذا تقدم أحمد ومجدي لأداء امتحان في مادة الإحصاء التطبيقي، وبافتراض أن نجاح أحدهما لا يتأثر بنجاح الآخر، احسب الاحتمالات التالية:

(١) نجاح واحد منهما على الأقل، مستخدماً في ذلك طريقتيْن مختلفتيْن.

(٢) رسوب أحمد أو مجدي.

ثانياً: إذا كانت دالة كثافة الاحتمال للمتغير العشوائي س هي:

$$\text{ف}(\text{س}) = \begin{cases} ٢\text{س} & \text{صفر} \le \text{س} \le ٢ \\ \text{صفر} & \text{خلاف ذلك} \end{cases}$$

المطلوب:

(١) حدد التوقُّع والتباين للمتغير س.

(٢) إيجاد دالة التوزيع الاحتمالي التجميعي.

(٣) احسب الاحتمالات التالية.

(أ) ع(س > ٠,٥) **(ب)** ع(−٠,٥ ≤ س ≤ ٠,٨)

السؤال الثاني:

أولاً: إذا كانت الدرجات التي حصل عليها طلاب الفرقة الثالثة في إحدى كليات التجارة في مادة المحاسبة تتبع توزيعاً معتاداً وسطه الحسابي ٧٠ درجة وانحرافه المعياري ٥ درجات.

المطلوب:

(١) إذا تم اختيار أحد الطلاب عشوائياً، احسب الاحتمالات التالية:

(أ) أن تبلغ درجة الطالب ٦٠,٢ على الأكثر.

(ب) أن تتراوح درجة الطالب ما بين ٧٠ درجة و٨٢,٩ درجة.

(٢) ما هي الدرجة التي يُتوقَّع أن يقل عنها أو يساويها ٩٧,٥٪ من إجمالي عدد الطلاب.

ثانياً: إذا كانت نسبة الوحدات المعيبة في إنتاج إحدى الآلات هي ٢٠٪، وأن عدد الوحدات المعيبة يتبع بواسون. احسب الاحتمالات التالية:

(١) الحصول على وحدة معيبة من بين ٥ وحدات تم اختيارها عشوائياً من إنتاج هذه الآلة.

(٢) الحصول على ٨ وحدات معيبة على الأكثر في عينة حجمها ١٠ وحدات مُختارة عشوائياً من إنتاج هذه الآلة.

السؤال الثالث:

أولاً: لنفترض أن مجتمعاً يتكون من أربعة أشخاص هم: أ ، ب ، جـ ، د . فإذا كان معلوماً لدينا أن هناك طالباً واحداً مدخناً للسجائر هو الطالب ج.

المطلوب:

(١) إيجاد توزيع المعاينة لنسبة المدخنين وذلك باستخدام عينة حجمها شخصان.

(٢) احسب الوسط الحسابي والانحراف المعياري لتوزيع المعاينة المحدَّد في (١).

ثانياً: في دراسة إحصائية لمعرفة معدل استهلاك نوع معين من السيارات للوقود تم اختيار تسع سيارات من هذا النوع من السيارات وتم تشغيلها وسُجلت المسافة التي تقطعها كل سيارة لكل لتر من البنزين، حيث وجد أن الوسط الحسابي والانحراف المعياري لتلك المسافات هما ١٢، ٢,٥ كم على الترتيب. بافتراض أن معدل استهلاك البنزين لهذا النوع من السيارات يتبع التوزيع المعتاد.

المطلوب:

(١) إيجاد فترة الثقة ٩٥٪ لمتوسط المسافة التي تقطعها السيارة المنتجة من هذا النوع لكل لتر واحد من البنزين، ثم فسر النتائج التي تحصل عليها.

(٢) إذا كان من المعروف أن الانحراف المعياري للمسافة التي تقطعها السيارة من هذا النوع لكل لتر واحد من البنزين هو ٢,٥ كم. استخدم هذا البيان في إعادة تقدير فترة الثقة المشار إليها في (١).

(٣) قارن بين نتائجك في (١) و (٢) مع إبداء التعليق المناسب.

يمكنك الاستعانة بالقيم الجدولية الملائمة مما يلي:

$\text{ي}_{٠,٠٥} = ١,٦٤٥$ ، $\text{ي}_{٠,٠٢٥} = ١,٩٦$ ، $\text{ت}_{٩، ٠,٠٢٥} = ٢,٢٦٢$

$\text{ت}_{٨، ٠,٠٥} = ١,٨٦$ ، $\text{ت}_{٨، ٠,٠٢٥} = ٢,٣٠٦$ ، $\text{ت}_{٩، ٠,٠٥} = ١,٨٣٣$

السؤال الرابع:

أولاً: عند تقدير فترة الثقة لنسبة ما في مجتمع توافرت البيانات التالية: حجم العينة = ٩٦، النسبة في العينة = ٠,٦، طول فترة الثقة = ٠,٢٥٨. حدِّد درجة الثقة الملائمة في تلك الحالة.

ثانياً: ترغب إدارة التسويق في إحدى الشركات في اختيار عينة لإجراء دراسة بهدف تحديد نسبة الذي يفضلون منتَجاً معيناً من منتَجات هذه الشركة وذلك في حدود خطأ مسموح به لا يتجاوز ± ٠,٠٢ عن النسبة الحقيقية في مجتمع المستهلِكين لهذا المنتَج وبدرجة الثقة ٩٥٪. مع العلم بأنه لا تتوافر أية بيانات عن النسبة في المجتمع.

المطلوب:

(١) كيف يمكن تحديد حجم العينة في حالة كتلك؟ وما هو هذا الحجم؟

(٢) ما هو مقدار التغير في حجم العينة إذا أُريد زيادة درجة الثقة إلى ٩٩٪ مع الإبقاء على حجم الخطأ المسموح به كما هو؟ قارن نتائجك هنا بما حصلت عليه في (١) مع التعليق.

السؤال الخامس:

أولاً: تتلقى جهة حكومية شكاوى كثيرة من مستهلكي نوع معين من مسحوق تنظيف الملابس فحواها أن وزن العبوة لا يصل إلى المقدار المعلن عنه وهو ٢٤٠ جم. وللتحقق من صحة شكاوى المستهلِكين قامت الجهة الحكومية بشراء ٤٠ عبوة حيث وجدت أن متوسط وزن المسحوق في العبوة الواحدة هو ٢٣٩ جم وأن الانحراف المعياري لهذا الوزن هو ٢ جم.
كيف يمكن للجهة الحكومية حسم هذا الأمر؟ استخدم مستوى معنوية ٠,٠١.

ثانياً: في دراسة لبيان أثر اغتراب الطلاب عن أسرهم على نتائجهم الدراسية في إحدى الكليات تم أخذ عينتيْن إحداهما من الطلاب المغتربين الذين يعيشون بمفردهم وبعيداً عن أسرهم، وأما العينة الأخرى فكانت من الطلاب الذين يعيشون مع أسرهم، وبدراسة نتائجهم في نهاية العام الدراسي كانت البيانات التالية:

	طلاب مغتربون	طلاب يعيشون مع أسرهم
حجم العينة	٨٠	٢٥٠
عدد الطلاب الناجحين	٥٦	١٧٠

المطلوب:

(١) قدرة بفترة ثقة ٩٩٪ الفرق بين نسبتي النجاح في هاتيْن الفئتيْن من الطلاب.

(٢) مستخدماً نتائجك في (١)، هل يمكن القول بأن نتيجة الطالب في نهاية العام الدراسي تتوقف على كونه مغترباً أو يعيش مع أسرته.

الامتحان رقم (٢)

المادة: إحصاء تطبيقي وحاسب آلي **الفرقة:** الثالثة
الكلية: التجارة بسوهاج **دور:** مايو ١٩٩٨م

طلاب الدراسة باللغة العربية

ملحوظة: الامتحان مكتوب باللغة الإنجليزية أيضاً لطلاب شعبة الدراسة باللغة الإنجليزية.

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: البيانات التالية توضح توزيع العاملين في أحد المصانع الكبيرة حسب مستوى تعليمهم وكونهم مدخنين أم لا.

حالة التدخين	المستوى التعليمي		
	خريج إعدادي	خريج ثانوي	خريج جامعة
يدخن	٢	٨	١٥
لا يدخن	٨	٣٢	٣٥

فإذا تم اختيار أحد العاملين عشوائياً.

المطلوب:

احسب الاحتمالات التالية

(١) أن لا يكون متخرجاً من الجامعة.
(٢) أن يكون مدخناً ومتخرجاً من الجامعة.
(٣) أن يكون متخرجاً من الجامعة بشرط أن يكون غير مدخن.
(٤) أن يكون مدخناً أو تخرج من المدارس الثانوية.

ثانياً: إذا كان المتغير س يمثل عدد الحوادث التي تحدث خلال أسبوع في إحدى المدن. وإذا كان التوزيع الاحتمالي لهذا المتغير هو كما يلي:

س	صفر	١	٢	٣
ع(س)	٠,٢	٠,٤	٠,٣	٠,١

المطلوب:

(١) احسب احتمال أن يكون عدد الحوادث التي تمت خلال أسبوع معين هو:
(أ) ثلاثة حوادث **(ب)** حادثان على الأقل **(ج)** أقل من حادثيْن
(٢) احسب توقع وتباين س.

السؤال الثاني:

أولاً: إذا كانت نسبة المعيب في إنتاج إحدى الآلات هي ٢٪. احسب احتمال أن عينة حجمها ١٤٠ وحدة تم اختيارها عشوائياً من إنتاج هذه الآلة سوف تحتوي على:

(١) وحدتيْن معيبتيْن.

(٢) وحدتيْن معيبتيْن على الأكثر.

(٣) وحدة واحدة معيبة على الأقل.

(٤) اكتب أوامر MINITAB اللازمة لحساب الاحتمالات المطلوبة في (١) ، (٢) ، (٣).

ثانياً: إذا كان عمر نوع معين من البطاريات المنتَجة في أحد المصانع تتبع توزيع معتاداً متوسطه ٣٠٠ ساعة وانحرافه المعياري ٥٠ ساعة.

المطلوب:

(١) إذا تم اختيار إحدى البطاريات عشوائياً، احسب الاحتمالات التالية:

(أ) أن يتراوح عمر البطارية ما بين ٢٠٢ ، ٣٩٨ ساعة.

(ب) أن يزيد عمر البطارية على ٣٩٨ ساعة.

(٢) ما هو عمر البطارية الذي يزيد عليه ٢,٥٪ من البطاريات.

(٣) ما هي أوامر MINITAB اللازمة للإجابة على كلٍ من (١) و (٢)؟

السؤال الثالث:

أولاً: في دراسة استطلاعية شملت ٢٥٠ ناخباً تبيَّن أن ٤٨٪ من الناخبين سوف يصوتون لصالح مرشح معين (أ).

المطلوب:

(١) ما هو التقدير بنقطة لنسبة الناخبين (ل) المؤيدين للمرشح (أ) وذلك في الدائرة الانتخابية ككل؟

(٢) كوِّن فترة الثقة ٩٥٪ للنسبة المشار إليها في (١)، موضحاً ما تعنيه هذه الفترة.

(٣) ما هو الحد الأقصى لتقدير حجم الخطأ في (٢)؟

(٤) وضِّح كيف يمكنك إنقاص طول فترة الثقة المكوّنة في (٢)؟ وما هي البدائل الممكنة لتحقيق ذلك؟ وما هو البديل الأمثل من وجهة نظرك؟ ولماذا؟

(٥) هل تحتوي فترة الثقة المكوّنة في (٢) على القيمة ل = ٠,٤٥ وماذا يعني ذلك بالنسبة لك؟

(٦) كيف يمكنك التحقق من صحة الادعاء القائل بأن أقل من نصف الناخبين في الدائرة الانتخابية هم الذين سوف يدلون بأصواتهم لصالح المرشح (أ)؟ استخدم مستوى المعنوية ٠,٠١.

ثانياً: الجدول التالي يوضـــح بعض البيانات عن أجور العاملين في عينتيْن من الشـــركتيْن (أ) و (ب):

الشركة	حجم العينة	الوسط الحسابي	الانحراف المعياري
الشركة (أ)	٤٠	٣٠٠ جنيه	٢٠
الشركة (ب)	٥٠	٢٨٠ جنيه	٢٥

المطلوب:

(١) كوِّن فترة الثـقة ٩٩٪ لمتوسط أجور العاملين في الشركة (أ) ككل.

(٢) ما هي أوامر MINITAB اللازمة لتحديد فترة الثـقة المطلوبة في (١)؟ إعرض مخرجات MINITAB لتلك الأوامر مستخدماً في ذلك النتائج التي حصلت عليها في (١).

(٣) ما هو التقدير بنقطة للفرق بين متوسطي أجور العاملين في الشركتيِن (أ) و (ب)؟

(٤) كوِّن فترة الثـقة ٩٩٪ للفرق الُمشار إليه في (٣).

(٥) بافتراض معرفة أجر كل عامل على حدة في العينتيْن، ما هي أوامر MINITAB اللازمة لتحديد فترة الثـقة المطلوبة في (٤)؟

(٦) مستخدماً نتائجك في (٤)، اختبر – بمستوى المعنوية ١٪ - ما إذا كانت أجور العاملين في الشركتيْن مختلفة.

(٧) بدون إجراء أية حسابات، وضِّح هل ستختلف نتائجك في (٦) فيما لو استخدمت مستوى المعنوية ٥٪ بدلاً من مستوى المعنوية ١٪؟ ولماذا؟

(٨) يدّعي مدير العاملين في الشركة (أ) بأن أجور العاملين في شركته تفوق نظيرتها في الشركة (ب). هل تؤيد بيانات العينتيْن ادعاء هذا المدير؟ استخدم مستوى المعنوية ٠,٠٠٥.

(٩) بافتراض معرفة أجر كل عامل على حدة في العينتيْن، اعرض أوامر MINITAB التي تمكنك من إجراء الاختبار المستخدم في (٨).

السؤال الرابع:

أولاً: إذا كان من المفترض أن عمر الرجال يتبع توزيعاً معتاداً انحرافه المعياري ٢,٥ بوصة. ما هو حجم العينة اللازم أخذه حتى نكون واثقين بنسبة ٩٥٪ بأن متوسط العينة لا يختلف عن متوسط المجتمع الذي سحبت منه العينة بأكثر من ٠,٥ بوصة زيادة أو نقصاناً؟

ثانياً: تهدف إحدى الدراسات إلى معرفة نسبة الأرامل بين النساء اللاتي تجاوزن العمر ٥٥ سنة. ما هو حجم العينة الواجب أخذه حتى نكون واثقين بنسبة ٩٩٪ بأن نسبة الأرامل بين

نساء العينة لا تختلف بأكثر من ٠,٠٥ عن نظيرتها بين النساء اللاتي تخطين العمر ٥٥ سنة بشكل عام؟ وذلك بافتراض:

(١) أن دراسة حديثة قد أفادت بأن نسبة الأرامل بين نساء عينة يزيد عمر كل منهن عن ٥٥ سنة هو ٢٠٪.

(٢) أنه ليست هناك معلومات متاحة عن نسبة الأرامل بين النساء اللاتي تزيد أعمارهن عن ٥٥ سنة.

(٣) قارن بين نتائجك في كل من (١) و (٢) مبدياً ما تراه مناسباً من تعليق.

الامتحان رقم (٣)

المادة: إحصاء تطبيقي وحاسب آلي **الفرقة:** الثالثة
الكلية: التجارة بسوهاج **دور:** مايو ١٩٩٩م

طلاب شعبة الدراسة باللغة الإنجليزية

ملحوظة: الامتحان تمت ترجمته هنا إلى اللغة العربية

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: يحتوي صندوق على ٢ كرة سوداء و٢ كرة بيضاء. إذا تم سحب كرة ثم وُضِع في الصندوق كرة من اللون المخالف. بعد ذلك، تم سحب كرة أخرى من الصندوق. احسب احتمال أن الكرة الأولى المسحوبة كانت بيضاء إذا علمت أن الكرة الثانية كانت سوداء.

ثانياً: إذا كان المتغير س يمثل عدد الأخطاء المطبعية التي تحتوي عليها صفحة تم اختيارها عشوائياً من كتاب. والجدول التالي يوضح التوزيع الاحتمالي للمتغير س.

س	صفر	١	٢	٣	٤
ع(س)	٠,١	٠,٢	٠,٤	٠,٢	٠,١

المطلوب:

(١) إيجاد الوسط الحسابي والتباين للمتغير ص، حيث: ص = ٢س.

(٢) علِّق على التوزيع من حيث تماثله أو التوائه.

(٣) احسب الاحتمالات التالية:

(أ) ع(س ≥ ٣) **(ب)** ع(س = ٢) **(ج)** ع(١ ≤ س ≤ ٤)

السؤال الثاني:

أولاً: يحتوي اختبار على خمسة أسئلة، وكل سؤال له أربعة اختيارات للإجابة عليه، ومن بين الاختيارات الأربعة هناك إجابة واحدة صحيحة. فإذا كانت الإجابة على كل سؤال مجرد تخمين لاختيار إجابة واحدة من بين الأربعة.

المطلوب:

(١) إيجاد الاحتمالات التالية:

(أ) سؤال واحد على الأقل إجابته صحيحة.

(ب) الإجابة صحيحة على أقل من سؤالَيْن.

(ج) الإجابة غير صحيحة على ثلاثة أسئلة.

(٢) استخدم MINITAB للإجابة على كل من (أ) ، (ب) ، (ج).

ثانياً: إذا كانت معدلات الذكاء (.I.Q) لمجموعة من الأشخاص تتبع توزيعاً معتاداً وسطه الحسابي ١٠٠ وانحرافه المعياري ١٦.

المطلوب:

(١) ما هو احتمال أن يكون معدل الذكاء لشخص تم اختياره عشوائياً:

(أ) أقل من أو يساوي ٨٠؟

(ب) ما بين ٩٠ و ١١٠؟

(٢) ما هو معدل الذكاء الذي يقل عنه ٥٪ من إجمالي عدد الأشخاص ويكون منتصفها المعدل ١٠٠.

(٣) إذا تم اختيار عينة عشوائية حجمها ٦٤ شخصاً من هذه المجموعة، احسب احتمال أن يبلغ متوسط العينة ١٠٤ على الأكثر.

(٤) استخدم MINITAB للإجابة على كل من (١) ، (٢) ، (٣) ، (٤).

السؤال الثالث:

أولاً: أُجريت دراسة لمقارنة نسبة الراضين Satisfied عن سياراتهم الجديدة المحلية والأجنبية. وفي عينة حجمها ٦٠٠ مشتري لسيارات محلية جديدة تم اختيارها عشوائياً وُجد أن هناك ٤٨٠ مشتري سعداء بسياراتهم. وفي عينة عشوائية أخرى حجمها ٤٥٠ مشتري لسيارات أجنبية جديدة وُجد أن هناك ٣٦٩ مشتري سعداء بشراء تلك السيارات.

المطلوب:

(١) إيجاد التقدير بنقطة لـ (ل١ – ل٢).

(٢) تكوين فترة الثقة ٩٥٪ للفرق ما بين نسبتي المجتمعيْن (ل١ – ل٢). فسِّر ما تحصل عليه من نتائج.

(٣) باستخدام نتائجك في (٢)، هل تعتقد أن هناك فرقاً جوهرياً بين النسبتيْن؟ اشرح نتائجك.

(٤) إذا ادعى مشترو السيارات الأجنبية أنهم أكثر اقتناعاً بسياراتهم من الذين اشتروا السيارات المحلية، هل يمكنك التحقق من صحة هذا الادعاء؟ استخدم مستوى المعنوية ٠,٠١.

ثانياً: أُجريت تجربة لمقارنة فعالية نظاميْن غذائييْن أ و ب في إنقاص الوزن. حيث تم اختيار مجموعتيْن من الأشخاص زائدي الوزن تضم كل منهما ٥٠ شخصاً. وقد طُبقّ النظام الغذائي أ على المجموعة الأولى، بينما طُبقّ النظام الغذائي ب على المجموعة الثانية وذلك لمدة ٣٠ يوماً، سُجِّل في نهايتها مقدار النقصان في الوزن لكل شخص. فإذا كان الوسط الحسابي والانحراف المعياري لمقدار النقصان في الوزن للنظاميْن الغذائييْن كما يلي:

	النظام الغذائي (أ)	النظام الغذائي (ب)
الوسط الحسابي	١٤,٥	١٣,٤
الانحراف المعياري	٢,٦	١,٩

المطلوب:

(١) إيجاد فترة الثقة ٩٩٪ للوسط الحسابي م$_{أ}$، حيث م$_{أ}$ هو متوسط الوزن المفقود نتيجة اتباع النظام الغذائي أ. فسِّر نتائجك.

(٢) إجراء الاختبارات التالية:

(أ) ف. : م$_{ب}$ = ١٤ ، ف$_{١}$: م$_{ب}$ > ١٤ ، α = ٠,٠١

(ب) ف. : م$_{ب}$ = ١٤ ، ف$_{١}$: م$_{ب}$ ≠ ١٤ ، α = ٠,٠٥

م$_{ب}$ هو متوسط الوزن المفقود نتيجة اتباع النظام الغذائي ب.

(٣) إيجاد فترة الثقة ٩٩٪ للفرق ما بين المتوسطيْن (م$_{أ}$ − م$_{ب}$).

(٤) هل فترة الثقة المكوَّنة في (٣) تحتوي على القيمة صفر؟ ولماذا يكون ذلك موضع اهتمام الباحث.

(٥) هل يمكننا استنتاج أن النظام الغذائي أ هو أكثر فعالية من النظام الغذائي ب في إنقاص الوزن؟ استخدم مستوى المعنوية ٠,٠٥.

(٦) بافتراض توافر مشاهدات العينتيْن، استخدم MINITAB للإجابة على (١) ، (٣).

السؤال الرابع:

أولاً: يرغب مدير التسويق بإحدى الشركات في أن يطرح منتَجاً جديداً في الأسواق. ولتحقيق ذلك يرى هذا المدير أن يقوم بدراسة استطلاعية على عينة من الأفراد حيث يقوم بسؤالهم عن ما إذا كانوا سيقومون بشراء منتجَه أم لا. ما هو حجم العينة الواجب أخذه إذا أراد المدير أن لا يتجاوز الخطأ في تقديره لنسبة من سيشترون هذا المنتَج في المجتمع ككل ٠,٠٣ ويكون واثقاً في ذلك ٩٩٪.

ثانياً: يود مدير العمليات في مصنع كبير للإنتاج في أن يقدر متوسط الوقت الذي يستغرقه العامل في تجميع منتَج إلكتروني جديد. ما هو حجم العينة الواجب أخذه حتى يضمن المدير أن لا يتجاوز الخطأ في تقديره للمتوسط ٢٠ ثانية وبثقة ٩٩٪ وذلك في الحالتيْن التاليتيْن:

(١) يعرف المدير من دراسات سابقة بأن الانحراف المعياري للوقت المستغرَق في تجميع هذا المنتَج بواسطة العمال يبلغ ٢,٥ دقيقة.

(٢) بعد ملاحظة عدد كبير من عمال يقومون بتجميع منتَج مماثل، كان أقل وقت استغرقته عملية التجميع هو ١٠ دقائق وأطول وقت هو ٢٢ دقيقة.

الامتحان رقم (٤)

المادة: إحصاء تطبيقي وحاسب آلي   **الفرقة:** الثالثة

الكلية: التجارة بسوهاج   **دور:** مايو ٢٠٠٠م

طلاب شعبة الدراسة باللغة الإنجليزية

ملحوظة: الامتحان تمت ترجمته هنا إلى اللغة العربية.

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: لعب فريقان أ و ب مباراة في كرة القدم. والجدول التالي يوضح احتمالات عدم تسجيل أهداف وتسجيل هدف أو اثنين بواسطة كل فريق.

عدد الأهداف	صفر		١		٢	
الفريق	أ	ب	أ	ب	أ	ب
الاحتمال	٠,٣	٠,٢	٠,٥	٠,٧	٠,٢	٠,١

المطلوب:

احسب الاحتمالات التالية:

١- فوز الفريق أ   **٢-** تعادل الفريقيْن   **٣-** فوز الفريق ب

ثانياً: متغير عشوائي س يأخذ القيم الزوجية الصحيحة فقط ما بين الرقميْن ١ و ٩. فإذا كان

$$ع(س) = \frac{س}{\sum س}$$

المطلوب:

(١) تحديد التوزيع الاحتمالي للمتغير س   **(٢)** ع(س ≤ ٦)   **(٣)** توقع المتغير س

السؤال الثاني:

أولاً: إذا كانت الحوادث الخطيرة تحدث بمعدل ١,٥ حادثة في الأسبوع وذلك في مصنع كبير للإنتاج.

المطلوب:

(١) إيجاد احتمال أن تحدث حادثة واحدة خلال أسبوع معين.

(٢) إيجاد احتمال أن تحدث أقل من حادثتيْن خلال فترة مدتها أربعة أسابيع.

(٣) استخدم MINITAB للإجابة على (١) ، (٢).

ثانياً: تتْبع الأجور الأسبوعية للعمال غير المهرة في مصنع كبير توزيعاً معتاداً وسطه الحسابي ٦٠ جنيهاً وانحراف معياري ٥ جنيهات.

المطلوب:

(١) حساب احتمال أن أجر عامل – أُختير عشوائياً – سوف يتراوح ما بين ٦٥ و ٧٥ جنيهاً.

(٢) في مجموعة مكوَّنة من ٤٠٠ عامل غير ماهر، ما هو العدد المتوقَّع للعمال الذين تزيد أجورهم الأسبوعية على ٦٠ جنيهاً.

(٣) إذا كان ٢٠٪ من العمال تقل أجورهم الأسبوعية عن الحد الأدنى المقبول للأجور، ما هو هذا الحد الأدنى؟

(٤) استخدم MINITAB للإجابة على (١) ، (٢) ، (٣).

السؤال الثالث:

أولاً: يرغب باحث طبي في معرفة ما إذا كان معدل نبضات القلب للأشخاص المدخنين يزيد على نظيره لغير المدخنين. ولتحقيق ذلك، تم اختيار عينة حجمها ١٠٠ شخص من كل من المدخنين وغير المدخنين فكانت النتائج التالية:

مدخنون	غير مدخنين
$\bar{س}_{١}$ = ٩٠	$\bar{س}_{٢}$ = ٨٨
$ع_{١}$ = ٥	$ع_{٢}$ = ٦
$ن_{١}$ = ١٠٠	$ن_{٢}$ = ١٠٠

المطلوب:

(١) إيجاد التقدير بنقطة لـ $م_{١}$ ، حيث $م_{١}$ هي متوسط معدل النبضات للمدخنين عموماً.

(٢) تحديد فترة الثقة ٩٩٪ لمتوسط معدل النبضات للمدخنين عموماً. اشرح نتائجك.

(٣) باستخدام النتيجة في (٢)، هل يمكن القول بأن متوسط معدل النبضات للمدخنين عموماً يساوي ٩١؟

(٤) ما هو الحد الأقصى لخطأ التقدير في (٢)؟

(٥) إيجاد فترة الثقة ٩٥٪ للفرق بين المتوسطيْن ($م_{١}$ − $م_{٢}$)، حيث $م_{١}$ كما سبق تعريفها، $م_{٢}$ هي متوسط معدل النبضات لغير المدخنين عموماً. اشرح نتائجك.

(٦) هل يمكن للباحث أن يستنتج – باستخدام مستوى المعنوية ٠,٠٥ – بأن معدل نبضات المدخنين يفوق نظيره عند غير المدخنين؟ ما هو حجم الخطأ من النوع الأول في هذه الحالة؟

(٧) ماذا سيصبح عليه قرارك في (٦) لو أن حجم الخطأ من النوع الأول أصبح صفراً؟ اشرح.

(٨) بافتراض توافر مشاهدات العينتيْن، استخدم MINITAB للإجابة على (٢) و (٥) و (٦).

ثانياً: تريد شركة في أن تعرف ما إذا كان الالتحاق بالبرنامج التدريبي "كيف تكون رجل بيع ناجح" يُزِيد من متوسط المبيعات لموظفيها. ولهذا الغرض، أرسلت الشركة ٦ موظفين للالتحاق بهذا البرنامج. والجدول التالي يوضح مبيعات الموظفين الستة خلال أسبوع قبل وبعد الالتحاق بالبرنامج التدريبي.

قبل	١٢	١٨	٢٥	٩	١٤	١٦
بعد	١٨	٢٤	٢٤	١٤	١٩	٢٠

المطلوب:

(١) كوّن فترة الثقة ٩٩٪ لمتوسط الفرق (ف) بين المبيعات بعد الالتحاق بالبرنامج والمبيعات قبل الالتحاق بالبرنامج.

(٢) باستخدام مستوى المعنوية ٠,٠١، هل يمكنك استنتاج أن الالتحاق بالبرنامج التدريبي يؤدي إلى زيادة متوسط المبيعات الأسبوعية للموظفين؟

(٣) استخدم MINITAB للإجابة على (١).

السؤال الرابع:

لنفترض أن رئيس جمهورية مصر العربية يريد أن يقدر بدرجة ثقة ٩٥٪ نسبة المواطنين الذين يؤيدون سياسته تجاه إسرائيل. وهو يرغب في أن لا يتجاوز التقدير ٠,٠٤ زيادة أو نقصاناً عن النسبة الحقيقة في المجتمع المصري ككل. وقد أظهرت دراسة حديثة بأن نسبة المؤيدين للسياسة هي ٠,٠٨.

المطلوب:

(١) ما هو حجم العينة الواجب أخذه؟

(٢) ما هو حجم العينة إذا لم تتوافر بيانات الدراسة الحديثة.

(٣) كيف تفسِّر الفرق بين حجم العينة المحسوب في (١) ، (٢)؟

الامتحان رقم (٥)

المادة: إحصاء تطبيقي وحاسب آلي　　　**الفرقة:** الثالثة
الكلية: التجارة بسوهاج　　　**دور:** مايو ٢٠٠١م

طلاب الدراسة باللغة العربية

ملحوظة: الامتحان مكتوب باللغة الإنجليزية أيضاً لطلاب شعبة الدراسة باللغة الإنجليزية.

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: ناقِش باختصار الفرق بين أمريْ MINITAB في كل مما يلي:

(أ) الأمران: DELETE و REASE　　**(ب)** الأمران: SET و READ

ثانياً: فيما يلي درجات مجموعة مكوَّنة من خمسة طلاب وذلك في مادة الإحصاء:

٨٥　　٦٧　　٦٣　　٩٢　　٧٤

المطلوب: كتابة أوامر MINITAB اللازمة لتنفيذ كل مما يلي:

ملحوظة: - يجب التعامل مع كل أمر على حدة.

- يُكتفي بكتابة أوامر MINITAB لكل حالة ولا يجب إجراء أية حسابات سواء باليد أو باستخدام الآلة الحاسبة.

(١) إدخال البيانات باستخدام الأمر SET وذلك في العمود C1.

(٢) تسمية العمود C1 بالاسم STAT.

(٣) التعرف على أكبر درجة.

(٤) ترتيب درجات الطلاب تصاعدياً.

(٥) إيجاد المقاييس الإحصائية التالية لدرجات الإحصاء:

(أ) الوسط الحسابي　**(ب)** الوسيط　**(ج)** الانحراف المعياري　**(د)** المدى

السؤال الثاني:

أولاً: فيما يلي التوزيع الاحتمالي لمتغير عشوائي س هي:

$$ع(س) = \begin{cases} \frac{١}{٨} & س = صفر ، ٣ \\ \frac{٣}{٨} & س = ١ ، ٢ \end{cases}$$

المطلــوب:

(١) إيجاد الاحتمالات التالية:　**(أ)** ع(س ≤ ١)　**(ب)** ع(س = ٣)

(٢) احسب الوسط الحسابي والتباين للمتغير س.
(٣) علِّق على التوزيع الاحتمالي للمتغير س من حيث التماثل أو الالتواء.

ثانياً: تتبع درجات الطلاب في مادة الإحصاء توزيعاً معتاداً متوسطه ١٤ درجة وانحرافه المعياري درجتان. فإذا تم اختيار أحد الطلاب عشوائياً، احسب الاحتمالات التالية:
(أ) أن تزيد درجة هذا الطالب عن ١٥ درجة.
(ب) **أن** تتراوح درجة هذا الطالب ما بين ١٢ و ١٦ درجة.
(ج) استخدم MINITAB للإجابة على المطلوب في كل من (أ) و (ب).

ثالثاً: تتبع أوزان طلاب إحدى الكليات توزيعاً معتاداً متوسطه ٦٥ كجم وانحرافه المعياري σ. فإذا كانت أوزان ٦٧٪ من الطلاب تقل عن ٧٠ كجم، احسب قيمة σ.

السؤال الثالث:

أولاً: في عينة مكونة من ٢٠٠ رجل ممن لهم حق التصويت في الانتخابات تبيَّن أن ٨٠ منهم قد أدلوْا بأصواتهم في الانتخاب الأخير لمجلس الشعب. وفي عينة أخرى مكوَّنة من ١٠٠ امرأة ممن لهن حق التصويت في الانتخابات كان من بينهن ٣٨ امرأة قد أدليْن بأصواتهن في نفس الانتخاب.

المطلوب:

(١) إيجاد فترة الثقة ٩٥٪ للفرق بين نسبتيْ من أدلوْا بأصواتهم في الانتخابات الأخيرة لمجلس الشعب من الرجال و النساء في الدائرة الانتخابية ككل.
(٢) هل يمكن القول بأن نسبة النساء الناخبات في الدائرة الانتخابية واللاتي أدليْن بأصواتهن في الانتخابات الأخيرة لمجلس الشعب تزيد على ٣٦٪ من إجمالي عدد من لهن حق التصويت في الدائرة الانتخابية؟ استخدم مستوى المعنوية ٠,٠١.

ثانيا: يرغب مسئول في أحد المصانع الكبيرة في تقدير متوسط الوقت الذي يأخذه العامل في عملية تجميع منتَج معين. وبملاحظة عدد من العمال يقومون بتجميع منتَج مشابه وجد هذا المسئول أن أقل وقت يحتاجه العامل هو ١٠ دقائق، في حين أن أطول وقت هو ٢٢ دقيقة.
ما هو حجم العينة الواجب أخذه حتى يتمكن هذا المسئول من إنجاز هدفه وذلك بدرجة ثقة ٩٥٪؟

السؤال الرابع:

أُجريت تجربة لمقارنة متوسط الوقت (بالأيام) والذي يحتاجه الشخص للشفاء من نزلات البرد العادية وذلك لمجموعة من الأشخاص يتعاطوْن جرعة يومية مقدارها ٤ ميليجرام من فيتامين C ومجموعة أخرى لا تتعاطى تلك الجرعات. فإذا افترضنا أن ٥٠ شخصاً تم اختيارهم عشوائياً من كل مجموعة حيث كانت نتائجهم كما يلي:

	لا يتعاطون فيتامين C	يتعاطون فيتامين C
حجم العينة	٥٠	٥٠
متوسط العينة	٥	٤
الانحراف المعياري للعينة	٣	١
متوسط المجتمع	μ_1	μ_2

فإذا افترضنا أن هدف الدراسة هو معرفة ما إذا كان تناول جرعات فيتامين C يؤدى إلى سرعة الشفاء من نزلات البرد.

المطلوب:

(١) صياغة فرضية العدم والفرضية البديلة لهذا الشأن.

(٢) إجراء الاختبار الإحصائي الذي يمَكِّن من تحقيق الهدف من هذه الدراسة.

(٣) تكوين فترة الثقة ٩٩٪ لـ μ_1.

(٤) بافتراض أن مشاهدات العينة الأولى متاحة، استخدم MINITAB لتكوين فترة الثقة المطلوبة في (٣).

الامتحان رقم (٦)

المادة: إحصاء تطبيقي وحاسب آلي　　**الفرقة:** الثالثة
الكلية: التجارة بسوهاج　　**دور:** مايو ٢٠٠٢م

طلاب شعبة الدراسة باللغة الإنجليزية

ملحوظة: الامتحان تمت ترجمته هنا إلى اللغة العربية.

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: لنفترض أنه في مدينة ما، لا يسقط المطر إطلاقاً في أيام الجمعة أو السبت أو الأحد أو الإثنين. واحتمال أن يسقط المطر في يوم الثلاثاء هو $\frac{١}{٥}$. وفيما يتعلق بيومي الأربعاء والخميس، فإن احتمال سقوط المطر في أي منهما بشرط سقوط المطر في اليوم السابق له مباشرةً هو α، واحتمال سقوط المطر في أي منهما بشرط عدم سقوطه في اليوم السابق له مباشرةً هو β.

المطلوب:

(١) اثبت أن احتمال سقوط المطر في يوم أربعاء هو $\frac{\alpha + ٤\beta}{٥}$، ثم أوجد احتمال سقوط المطر في يوم الخميس.

(٢) إشرح الحالة التي يكون فيها $\alpha = \beta$.

ثانياً: إذا كان المتغير س يمثل عدد الأبناء الذين تبلغ أعمارهم أقل من ١٨ سنة في الأسرة المصرية. وكان التوزيع الاحتمالي للمتغير س كما هو موضح في الجدول التالي.

س	صفر	١	٢	٣	٤	٥
ع(س)	٠,٥	٠,٢١	٠,١٩	٠,٠٧	٠,٠٢	٠,٠١

المطلوب:

(١) علِّق على تماثل أو التواء هذا التوزيع.

(٢) ما هو عدد الأبناء الأكثر شيوعاً في الأسرة المصرية.

(٣) احسب ما يلي:

(أ) ع(س = ٣)　　**(ب)** ع(س ≤ ٤)

(ج) ع(٢ < س ≤ ٥)　　**(د)** توقع (٢س − ١)　　**(هـ)** تباين (٢س + ١)

السؤال الثاني:

أولاً: إذا كان معروفاً من خلال خبرة سابقة أن ٦٠٪ من المتقدمين لوظيفة معينة يجتازون اختباراً معيناً لتقييم من يتقدم إلى هذه الوظيفة. فإذا تقدم ٥ أشخاص لهذا الاختبار، احسب الاحتمالات التالية:

(١) أن جميع المتقدمين سوف يجتازون الاختبار.

(٢) أن واحداً على الأقل سوف يفشل في اجتياز الاختبار.

ثم استخدم MINITAB للإجابة على (١) و (٢).

ثانياً: إذا كان أجر الساعة الواحدة للعمال غير المهرة في مصنع كبير يتبع توزيعاً معتاداً متوسط ٥ جنيهات وانحرافه المعياري ٠,٨ جنيه. فإذا تم اختيار أحد العمال عشوائياً:

(١) احسب احتمال أن هذا العامل يتقاضى أجراً مقداره:

(أ) ما بين ٣ و ٧ جنيهات في الساعة.

(ب) أكبر من ٦ جنيهات في الساعة.

(٢) إذا كان ٨٠٪ من إجمالي عدد العمال يتقاضى أجراً أكبر من الحد الأدنى المقبول، ما هو هذا الحد الادنى؟

(٣) استخدم MINITAB للإجابة على (١) و (٢).

السؤال الثالث:

أولاً: استعانت شركة تأمين كبرى بمكتب استشاري لبحث ما إذا كان البائعون الحاصلون على درجة علمية في الإدارة هم الأفضل. وفي هذا الشأن، قام المكتب الاستشاري باختيار عينة مكونة من ٤٠ بائعاً حاصلين على درجة علمية في الإدارة حيث وُجد أن الوسط الحسابي والانحراف المعياري لعدد بوالص التأمين المُباعة أسبوعياً هما ١٠ و ١,٨ بوليصة تأمين على الترتيب. ثم قام المكتب باختيار عينة عشوائية مكونة من ٤٥ بائعاً لم يحصلوا على درجة علمية في الإدارة حيث وُجد أن الوسط الحسابي والانحراف المعياري لعدد بوالص التأمين المُباعة أسبوعياً هما ٨,٥ و ١,٣٥ بوليصة تأمين على الترتيب.

المطلوب:

(١) تكوين فترة الثقة ٩٩٪ للفرق ما بين متوسطي المجتمعيْن ($م_١ - م_٢$).

(٢) إيجاد فترة الثقة ٩٥٪ للمتوسط $م_٢$. فسِّر نتائجك.

(٣) باستخدام مستوى المعنوية ٠,٠٥، هل يمكن استنتاج أن البائعين الحاصلين على درجة علمية في الإدارة هم أفضل من أولئك غير الحاصلين على درجة علمية في الإدارة.

(٤) بافتراض توافر مشاهدات العينتيْن، استخدم MINITAB في الإجابة على كلٍ من (١) و (٢) و (٣) و (٤).

ثانياً: يُراد معرفة نسبة المستهلِكين الذين يفضلون شراء نوع معين من الزبد. ما هو حجم العينة الواجب أخذه بحيث نكون واثقين بدرجة ٩٥٪ وأن لا يتجاوز خطأ التقدير ٠,٠٢ زيادة أو نقصاناً عن النسبة الحقيقية بين المستهلِكين ككل، وذلك في الحالتيْن التاليتيْن:

(١) من المعروف أن النسبة في المجتمع تبلغ ٠,٦.

(٢) ليست هناك بيانات متوافرة عن النسبة في المجتمع.

علِّق على نتائجك في (١) و (٢).

الامتحان رقم (٧)

المادة: إحصاء تطبيقي **الفرقة:** الثانية
الكلية: التجارة بسوهاج **دور:** مايو ٢٠٠٢م

طلاب شعبة الدراسة باللغة الإنجليزية

ملحوظة: الامتحان تمت ترجمته هنا إلى اللغة العربية.

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: ثلاثة صناديق متشابهة يحتوي كل منها على كرتيْن. الصندوق الأول يحتوي على ٢ كرة بيضاء، بينما يحتوي الثاني على ٢ كرة سوداء ويحتوي الثالث على كرة واحدة سوداء وأخرى بيضاء. تم اختيار أحد الصناديق الثلاثة عشوائياً وأُخذت منه كرة. إذا كانت هذه الكرة بيضاء، ما هو احتمال أن تكون الكرة الأخرى التي لازالت في الصندوق بيضاء أيضاً.

ثانياً: الجدول التالي يوضح توزيع ١٠٠ موظف تم اختيارهم عشوائياً من إحدى الشركات الكبيرة وذلك حسب الجنس والوقت الذي يستغرقه الموظف في الانتقال من بيته إلى عمله.

البيان	الوقت المستغرق		
	أقل من ٣٠ دقيقة	٣٠ دقيقة إلى ساعة	أكثر من ساعة
رجال	٢٥	٢٠	٥
نساء	١٨	٢٠	١٢

المطلوب:

(١) إذا تم اختيار موظف بطريقة عشوائية من بين الـ ١٠٠ موظف، احسب الاحتمالات التالية:
أ- ع(أن يستغرق وقتاً أكثر من ساعة وأن يكون رجلاً).
ب- ع(أن يكون امرأة بشرط أن الوقت المستغرق أقل من ٣٠ دقيقة.
ج- ع(أن يكون رجلاً أو امرأة).

(٢) هل الحدثان "رجل" و "٣٠ دقيقة إلى ساعة" مستقلان؟ ولماذا؟

(٣) هل الحدثان "امرأة" و "أكثر من ساعة" متنافيان؟ لماذا؟

السؤال الثاني:

أولاً: إذا كان ١٠٪ من الوحدات المنتَجة في أحد المصانع معيبة. فإذا تم اختيار عينة عشوائية مكونة من ١٠ وحدات من إنتاج هذا المصنع، احسب احتمالات أن العينة سوف تحتوي على:

(١) وحدات جميعها سليمة.

(٢) وحدة واحدة معيبة. **(٣)** وحدات سليمة على الأكثر.

ثانياً: إذا كان الوقت المُستغرق لتعلُّم تغليف منتَج معين بواسطة موظفين جدد في إحدى المؤسسات يتبع توزيعاً معتاداً متوسطه ٢٤ ساعة وانحرافه المعياري ٢,٥ ساعة.

المطلوب:

(١) ما هي النسبة المئوية للموظفين الجدد الذين يمكنهم التعلُّم في أكثر من ٢٠ ساعة؟

(٢) إذا تم اختيار موظف جديد بطريقة عشوائية، ما هو احتمال أن يتراوح وقت التعلم ما بين ٢٦ و ٣٠ ساعة؟

(٣) ما هو وقت التعلُّم الذي يزيد عليه ٢٠٪ من إجمالي عدد الموظفين الجدد؟

السؤال الثالث:

أولاً: يدعَّي إحصائي بأن متوسط درجات الطلاب في تخصص المحاسبة يزيد على متوسط درجات الطلاب في تخصص الاقتصاد وذلك في اختبار قدرات موحَّد أُعطي لطلاب التخصصيْن. ويوضح الجدول التالي نتائج ٥٠ طالباً في كل تخصص في هذا الاختبار:

المحاسبة	الاقتصاد
$\bar{س}_١ = ١١٨$	$\bar{س}_٢ = ١١٥$
$ع_١ = ١٥$	$ع_٢ = ١٢$
$ن_١ = ٥٠$	$ن_٢ = ٥٠$

المطلوب:

(١) ما هو التقدير بنقطة للفرق بين المتوسطيْن $(م_١ - م_٢)$؟

(٢) كوِّن فترة الثـقة ٩٥٪ للفرق بين المتوسطيْن $(م_١ - م_٢)$. ثم اشرح نتائجك.

(٣) ما هو الحد الأقصى لخطأ التقدير في (٢)؟

(٤) إيجاد فترة الثـقة ٩٩٪ لـ $م_٢$.

(٥) باستخدام مستوى المعنوية ٠,٠١، هل يمكن استنتاج أن متوسط درجات الطلاب عموماً في تخصص المحاسبة أعلى من متوسط درجات الطلاب عموماً في تخصص الاقتصاد؟

(٦) هل تعطي نتائج العينتيْن دليلاً كافياً على صحة ادعاء الإحصائي؟ استخدم مستوى المعنوية ٠,٠٥.

ثانياً: في عينة استطلاعية حجمها ٤٠ تلميذاً أُخذت من إحدى المدارس، تبيَّن أن هناك ٥ طلاب يستعملون أيديهم اليسرى في الكتابة.

المطلوب:

(١) تكوين فترة الثـقة ٩٥٪ لنسبة التلاميذ الذين يستعملون أيديهم اليسرى في الكتابة وذلك في المدرسة ككل.

(٢) ما هو حجم العينة اللازم لإنقاص طول فترة الثـقة المكوَّنة في (١) إلى ٠,١؟

الامتحان رقم (٨)

المادة: إحصاء تطبيقي **الفرقة:** الثالثة

الكلية: التجارة بقنا **دور:** مايو ٢٠٠٣م

طلاب الدراسة باللغة العربية

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: إذا كان احتمال سقوط المطر في أحد أيام شهر يناير هو ٠,٢٥ وإذا سقط المطر في يوم ما من أيام شهر يناير فإن احتمال وقوع حادث تصادم سيارات في هذا اليوم هو ٠,٠٤ وإذا لم يسقط المطر في يوم ما في هذا الشهر فإن احتمال وقوع حادث تصادم سيارات في هذا اليوم هو ٠,٠١ ، احسب الاحتمالات التالية:

المطلوب:

(١) سقوط المطر في يوم ١٢ يناير وأن حادث تصادم سيارات سوف يقع في هذا اليوم؟

(٢) وقوع حادث تصادم سيارات في يوم ١٢ يناير.

ثانياً: متغير عشوائي س له التوزيع الاحتمالي التالي:

س	صفر	١	٢
ع(س)	٠,٢٥	٠,٢٥	٠,٥٠

المطلوب:

(١) علِّق على شكل التوزيع من حيث التماثل أو الالتواء.

(٢) احسب الاحتمال: ع(صفر ≤ س ≤ ٢).

(٣) إيجاد القيمة المتوقَّعة والتباين للمتغير: ص = ٢س – ٣.

السؤال الثاني:

أولاً: إذا كـــان عـــدد حـــالات دخـــول الطوارئ في أحد المستشفيات يتبع توزيع بواسون بمتوسط ٢.

المطلوب:

(١) إيجاد احتمال أنه في يوم معين لن تكون هناك حالات دخول طارئة.

(٢) في بداية أحد الأيـام كان هناك سريران متاحان غير مشغولَيْن في المستشفي، احسب احتمال أن هذا العدد من الأسِرّة لن يكون كافياً لهذا اليوم.

(٣) ما هو احتمال دخول ٣ حالات طارئة خلال يومَيْن متتاليَيْن؟

ثانياً: في أحد الامتحانات النهائية في مادة الإحصاء التطبيقي، كانت درجات الطلاب تتبع توزيعاً معتاداً متوسطه ٧٠ درجة وانحرافه المعياري ١٠ درجات.

المطلوب:

(١) إيجاد احتمال أن درجة طالب تم اختياره عشوائيا من بين الطلاب الذين أدَّوْا هذا الامتحان هي:

(أ) ٧٠ على الأكثر **(ب)** تتراوح ما بين ٥٠ و ٩٠.

(٢) ما هي الدرجة التي يحصل على أكبر منها ٥٪ من إجمالي عدد الطلاب؟

(٣) إذا كانت الدرجة التي حصلت عليها هي ٨٥، عبِّر عن الفرق بين هذه الدرجة وبين الوسط الحسابي بدلالة الانحراف المعياري. ماهي النسبة المئوية للطلاب الذين حصلوا على درجات في هذا الامتحان تزيد على هذه الدرجة؟

السؤال الثالث:

أولاً: يرغب عميد إحدى الكليات في معرفة ما إذا كان هناك فرق معنوي (جوهري) في المهارات الرياضية (الحسابية) بين الطلاب الذكور وبين الطالبات. ولتحقيق ذلك تم اختيار عينة عشوائية مكونة من ١٥٠ طالباً وعينة عشوائية أخرى مكونة من ١٠٠ طالبة حيث أدَّوْا اختبارا لقياس تلك المهارات، فكانت النتائج التالية فيما يتعلق بالدرجات التي حصلوا عليها:

	الطلاب الذكور	الطالبات
الوسط الحسابي	٨٥	٨٠
الانحراف المعياري	١٦	١٢

المطلوب:

(١) ما هو التقدير بنقطة للفرق بين متوسطي المجتمعيْن؟

(٢) كوِّن فترة الثقة ٩٥٪ لمتوسط درجات الطلاب الذكور ككل.

(٣) ما هي قيمة الحد الأقصى للخطأ للتقدير المحسوب في (٢)؟

(٤) إذا افترضنا أن طول فترة الثقة في (٢) كبير جداً، ما هي البدائل الممكنة لإنقاص هذا الطول؟ وما هو أفضل هذه البدائل؟

(٥) إيجاد فترة الثقة ٩٩٪ للفرق ما بين متوسطي المجتمعيْن.

(٦) باستخدام مستوى المعنوية ٥٪ ، هل يمكنك استنتاج أن المهارات الرياضية (الحسابية) عند الطلاب الذكور أفضل منها عند الإناث؟

ثانياً: لاحظ مدير الحسابات في أحد البنوك الرئيسية زيادة نسبة العملاء المتأخرين عن سداد القروض في موعدها. ولإيجاد تقدير لهذه النسبة تمت دراسة سجلات السداد لعينة من هؤلاء العملاء. ويود المدير بأن لا تختلف النسبة في العينة عن النسبة بين جميع العملاء المتأخرين

عن السداد بأكثر من ٣٪ وبدرجة ثقة ٩٥٪. ما هو حجم العينة التي يجب أخذها وذلك في الحالتيْن التاليتيْن:

(١) في دراسة أُجريت منذ عدة سنوات تبيَّن أن نسبة العملاء المتأخرين عن سداد قروضهم تبلغ ٢٠٪.

(٢) ليست هناك أية معلومات متاحة عن هذه النسبة.

الامتحان رقم (٩)

المادة: إحصاء تطبيقي — **الفرقة:** الثالثة

الكلية: التجارة بسوهاج — **دور:** يونيو ٢٠٠٣م

طلاب الدراسة باللغة العربية

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: إذا كـان من الممكن التحكم في رمى زهرة نرد بحيث يكون احتمال الحصول على الرقم ١ هو $\frac{١}{٤}$ س ، واحتمال الحصول على الرقم ٢ هو $\frac{١}{٤}$ ، واحتمال الحصول على الرقم ٦ هو $\frac{١}{٤}$(١ – س)، أخيراً احتمال الحصول على أي من الأرقام ٣ ، ٤ ، أو ٥ هو $\frac{١}{٦}$.

المطلوب:

(١) إثبات أن احتمال الحصول على رقميْن مجموعهما ٧ هو $\frac{\text{٩س - ٩س + ١٠}}{٧٢}$.

(٢) إيجاد قيمة س التي تجعل قيمة هذا الاحتمال أكبر ما يمكن.

ثانيا: يحتوي صندوق على ٣ كرات حمراء و٤ كرات بيضاء و٥ كرات سوداء. فإذا تم اختيار كرتيْن من الصندوق، احسب احتمال أن تكونا من نفس اللون وذلك بافتراض:

(١) إعادة الكرة الأولى المسحوبة من الصندوق قبل سحب الكرة الثانية.

(٢) عدم إعادة الكرة الأولى إلى الصندوق قبل سحب الكرة الثانية.

السؤال الثاني:

أولاً: إذا كان عدد العيوب في متر واحد من القماش المصنوع بواسطة إحدى الآلات يتبع توزيع بواسون بمتوسط ٠,١.

المطلوب:

(١) إيجاد الاحتمالات التالية:

أ- أن قطعة قماش طولها أربعة أمتار تكون خالية تماماً من العيوب.

ب- أن هناك عيبيْن في قطعة قماش طولها متر واحد

(٢) ما هو الطول الذي يجب أن تكون عليه قطعة قماش حتى يكون احتمال أن لا تحتوي على أي عيوب أقل من ٠,٩٥؟

ثانياً: إذا كان عمر نوع معين من المصابيح الكهربائية المنتَجة في أحد المصانع يتبع توزيعاً معتاداً متوسطه ١٠٠ ساعة وانحرافه المعياري ٨ ساعات.

المطلوب:

(١) إذا تم اختيار أحد المصابيح عشوائياً، احسب الاحتمالات التالية:

أ- أن يبلغ عمر المصباح ٩٠ ساعة على الأكثر.

ب- أن يزيد عمر المصباح عن ١٢ ساعة.

(٢) ما هي النسبة المئوية المتوقَّعة للمصابيح الكهربائية التي يزيد عمرها عن ١٠٠ ساعة؟ وما هي القيمة التي يجب أن يكون عليها الانحراف المعياري وذلك حتى تقل هذه النسبة إلى النصف؟

(٣) ما هو عمر المصباح الذي يزيد عليه ٥٪ من إجمالي عدد المصابيح؟

السؤال الثالث:

أولاً: أظهرت عينة مكونة من ١٥٠ شركة صناعية من النوع (أ) بأن ٢٧٪ من هذه الشركات تنفق أكثر من ٣٪ من إجمالي مبيعاتها على الدعاية والإعلان. بينما أفادت عينة أخرى مكونة من ١٠٠ شركة صناعية من النوع (ب) بأن ٢٠٪ من الشركات تنفق أكثر من ٣٪ من إجمالي مبيعاتها على الدعاية والإعلان.

المطلوب:

(١) ما هو التقدير بنقطة للفرق بين نسبتي المجتمعيْن ($ل_{أ} - ل_{ب}$)؟

(٢) كوّن فترة الثقة ٩٩٪ لنسبة الشركات من النوع (ب) ككل والتي تنفق أكثر من ٣٪ من إجمالي مبيعاتها على الدعاية والإعلان ($ل_{ب}$).

(٣) بافتراض ثبات النسبة في العينة ودرجة الثقة على ما هما عليه في (٢)، ما هو حجم العينة اللازم لإنقاص طول فترة الثقة المحسوبة في (٢) بمقدار ٠,٠٢؟

(٤) إيجاد فترة الثقة ٩٩٪ للفرق بين نسبتي المجتمعيْن ($ل_{أ} - ل_{ب}$).

(٥) باستخدام مستوى المعنوية ٥٪، هل يمكنك استنتاج أن نسبة الشركات التي تنفق أكثر من ٣٪ من إجمالي مبيعاتها على الدعاية والإعلان هي واحدة في النوعيْن (أ) و(ب)؟

(٦) باستخدام نتائجك في (٤)، كيف يمكنك التحقق من النتيجة التي وصلت إليها في (٥)؟

ثانياً: يرغب مدير شركة كبرى لتجميع أجهزة الحاسب الآلي في تقدير متوسط الوقت الذي يستغرقه العاملون في تجميع مكونات جهاز الحاسب.

المطلوب:

(١) ما هـو حجم العينة الواجب دراستها حتى يكون مدير الشركة واثـقا بنسبة ٩٥٪ بأن متوسط العينة لا يختلف بأكثر من ١٠ دقائق عن المتوسط الحقيقي للعامليـن ككل في عملية التجميع؟ من المعلوم أنه في دراسة مشابهة أُجريت من قبل أفادت بأن الانحراف المعياري للوقت المستغرق في عملية التجميع يساوى ٤٠ دقيقة.

(٢) إيجاد مقدار وطبيعة التغير في حجم العينة إذا ما رغب مدير الشركة في أن يكون واثـقاً في عملية تقديره بنسبة ٩٩٪ بدلاً من ٩٥٪.

(٣) كيف يمكنك تفسير هذا التغير في حجم العينة؟

الامتحان رقم (١٠)

المادة: إحصاء تطبيقي **الفرقة:** الثانية

الكلية: التجارة بقنا **دور:** مايو ٢٠٠٨م

طلاب الدراسة باللغة العربية

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: من بين ١٠ وحدات مُنتَجة بواسطة إحدى الآلات كان هناك ٤ وحدات معيبة. فإذا تم اختيار وحدتيْن، واحدة تِلو الأخرى (بدون إرجاع).

المطلوب:

احسب الاحتمالين التالييْن:

(١) الحصول على وحدة معيبة وأخرى سليمة. **(٢)** الحصول على وحدتيْن سليمتيْن.

ثانياً: ترغب إحدى الشركات في تغيير موعد بدء العمل اليومي بها من الساعة الثامنة صباحاً إلى الساعة السابعة والنصف صباحاً. وباستطلاع آراء ١٢٠٠ موظف إداري وعامل إنتاج حول هذا الأمر، تَبيَّن موافقة ٥٢٥ عامل إنتاج من بين ٧٥٠ هم إجمالي عدد عمال الإنتاج في هذه الشركة. كما تَبيَّن موافقة ٨٤٠ موظف إداري وعامل إنتاج على تغيير موعد بدء العمل. ولاتخاذ قرار نهائي في هذا الصدد، قرر مدير الشركة التحدث مع موظفين وعمال إنتاج يتم اختيارهم بطريقة عشوائية.

المطلوب:

(١) إذا تم اختيار أحد العاملين بطريقة عشوائية، أوجد الاحتمالات التالية:

(أ) أن يكون مؤيداً للتغيير؟ **(ب)** أن يكون معارضاً للتغيير وموظفاً إدارياً.

(٢) إذا كان من تم اختياره مؤيدا للتغيير، ما هو احتمال أن يكون عامل إنتاج؟

(٣) هل الحدثان "موظف إداري" و "معارض للتغيير" مستقلان؟ وضِّح.

ثالثاً: إذا كان الحدثان ب و جـ مستقلين. كما أن

$$ع(ب \text{ و } جـ) = \frac{١}{٥} \qquad \text{و} \qquad ع(ب) + ع(جـ) = \frac{٤}{٥}$$

المطلوب: إثبات أن الحدثين ب و جـ متنافيان.

السؤال الثاني:

أولاً: يتكون أحد الاختبارات من ثلاثة أسئلة ذات الاختيارات المتعددة في إجاباتها. حيث يوجد أربعة اختيارات للإجابة على كل سؤال من بينها اختيار واحد صحيح. فإذا كانت الإجابة على تلك الأسئلة تقوم أساساً على مجرد التخمين في اختيار الإجابة.

المطلوب:

(١) إيجاد احتمال الحصول على الدرجة النهائية في هذا الاختبار.

(٢) ما هو احتمال الرسوب في هذا الامتحان إذا كان النجـاح يتطلب الإجابة الصحيحة على سؤاليْن على الأقل. **ملحوظة:** أجب على هذا السؤال بطريقتين مختلفتين.

(٣) إذا افترضنا أنه يمكن استبعـاد اختياريْن غير صحيحيْن من الاختيارات الأربعة لكل سؤال، أعِد الإجابة على (٢).

ثانياً: إذا كان الوقت المخصص أسبوعياً لاستذكار مادة الإحصاء التطبيقي من قِبل الطلاب الحاصلين على تقدير ممتاز في هذه المادة يتبع توزيعاً طبيعياً (معتاداً) وسطه الحسابي ٧,٥ ساعة وانحرافه المعياري ١,٨ ساعة.

المطلوب:

(١) ما هي نسبة الطلاب الحـاصلين على تقديـر ممتـاز ويقضون وقتاً يزيـد على ٩ ساعات في استذكـارهم لمادة الإحصاء التطبيقي؟

(٢) أوجد احتمال أن طالباً ممتازاً يقضى ما بين ٦ و٩ ساعات في استذكاره لتلك المادة.

(٣) حدِّد الوقت الذي يقل عنه ١٠٪ من إجمالي الطلاب الذين حصلوا على تقدير ممتاز في استذكارهم لمادة الإحصاء التطبيقي.

ثالثاً: تتبع المبيعات السنوية لأحد المنتَجات توزيعاً طبيعياً (معتاداً) غير معروف وسطه الحسابي وانحرافه المعياري. فإذا كان معلوماً أن ٤٠٪ من المبيعات تزيد على ٤٧٠٠٠٠ جنيه، بينما ١٠٪ من المبيعات تزيد عن ٥٠٠٠٠٠ جنيه.

المطلوب:

تحديد قيمة كل من الوسط الحسابي والانحراف المعياري.

السؤال الثالث:

أولاً: في دراسة لمقارنة إنتاجية آلتيْن تُنتجان نفس المنتَج في أحد المصانع، تمت ملاحظة الآلة (١) لمدة ٤٠ ساعة كما تمت ملاحظة الآلة (٢) لمدة ٥٠ ساعة. وبحساب إنتاجية كل آلة مُعبَّراً عنها بعدد الوحدات المنتَجة في الساعة الواحدة في المتوسط، وبإيجاد الانحراف المعياري لإنتاج كل آلة، كانت النتائج التالية:

	الآلة (١)	الآلة (٢)
متوسط الانتاجية	٦١,٤	٥٩,٥
الانحراف المعياري	٣,١	٢,٨

المطلوب:

(١) ما هو التقدير بنقطة للفرق بين المتوسطيْن $م_١ - م_٢$؟ حيث $م_١$ و $م_٢$ هما متوسط الإنتاجيـة في الساعة الواحدة لكل من الآلتيْن (١) و (٢) على الترتيب.

(٢) أوجد فترة الثـقة ٩٥٪ للفرق بين المتوسطيْن ($م_٢ - م_١$).

(٣) ما هو الحد الأقصى للخطأ في (٢)؟

(٤) هـل تعطي بيانـات العينتيْن دليلاً كافياً للقول بأن إنتاجيـة الآلة (١) تفوق إنتاجيـة الآلة (٢)؟ استخدم مستوى المعنوية ٠,٠١.

(٥) هل يمكن أن تتغير نتائجك التي حصلت عليها في (٤) فيما لو استخـدمت مستوى المعنـوية ٠,٠٥ بـدلاً من ٠,٠١؟ وضِّح. **ملحوظة:** أجب بدون إجراء أي حسابات.

ثانياً: يرغب مدير شركة كبرى لتجميع أجهزة الحاسب الآلي في تقدير متوسط الوقت الذي يستغرقه العاملون في تجميع مكونات جهاز الحاسب.

المطلوب:

(١) ما هو حجم العينة الواجب اختياره حتى يكون مدير الشركة واثـقا بنسبة ٩٥٪ بأن متوسط العينة لن يختلف بأكثر من ١٥ دقيقة عن المتوسط الحقيقي للعاملين ككل في عملية التجميع وذلك في كل من الحالتيْن التاليتيْن:

(أ) في دراسة مشابهة تم إجراؤها من قبل تبيَّن أن الانحراف المعياري للوقت المُستغرق في عملية التجميع يساوى ٥٠ دقيقة.

(ب) بعد مشاهدة عدد من العاملين يقومون بتجميع جهاز مشابه، لاحظ مدير الشركة بأن أقل وقت تأخذه عملية التجميع هو ساعتان، بينما أطـول وقت يبلغ ٥ ساعات.

(٢) بماذا تفسِّر الاختلاف بين حجم العينة في (أ) وحجمها في (ب)؟

السؤال الرابع:

أولاً: يدَّعى المسئولون في شركة تأمين بأن مرض القلب ينتشر بشكل أكبر بين المُدخِّنين منه بين غير المدخنين، وذلك لمن تزيد أعمارهم عن ٥٠ سنة. وعليه يمكن أن تمنح شركة التأمين خصومات على بوالص التأمين الممنوحة لغير المُدخِّنين. ومع ذلك، وقبل اتخاذ أي قرار في هذا الشأن، وللتأكد من صحة الادعاء، اختارت الشركة وبطريقة عشوائية عينة مكوَّنة من ٢٠٠ رجل من بينهم ٨٠ مُدخِّناً و١٢٠ غير مُدخِّن. حيث وُجد أن هناك ١٨ مُدخِّناً يعانون من مرض القلب بينما هناك ١٥ شخصا غير مُدخِّن يعانون من هذا المرض. فإذا كانت:

$ل_١$ = نسبة الرجال المُدخِّنين فوق ٥٠ سنة والذين يعانون من مرض القلب،

$ل_٢$ = نسبة الرجال غير المُدخِّنين فوق ٥٠ سنة والذين يعانون من مرض القلب

المطلوب:

(١) باستخدام مستوى المعنوية ٥٪ ، هل تؤكد بيانات العينة صحة الادعاء بأن مرض القلب أقل انتشاراً بين غير المُدخِّنين منه بين المُدخِّنين؟

(٢) أوجد فترة الثقة ٩٩٪ لنسبة الرجال المُدخِّنين فوق ٥٠ سنة والذين يعانون من مرض القلب.

(٣) هل يمكنك القول بأن نسبة الرجال المُدخِّنين فوق الخمسين والذين يعانون من مرض القلب تختلف عن ٠,٢٤؟ استخدم مستوى المعنوية ١٪.

(٤) هل تتفق نتائج إجراء الاختبار في (٣) مع فترة الثقة التي حصلت عليها في (٢)؟ برِّر إجابتك في حالة الإجابة بـ "نعم" أو "لا".

ثانياً: لاحظ مدير الحسابات في أحد البنوك الرئيسية زيادة نسبة العملاء المتأخرين عن سداد القروض في موعدها المحدد. ولإيجاد تقدير لهذه النسبة تمت دراسة سجلات السداد لعينة من هؤلاء العملاء. ويود المدير بأن لا تختلف النسبة في العينة عن النسبة بين جميع العملاء المتأخرين عن السداد بأكثر من ٢٪ وبدرجة ثقة ٩٩٪.

المطلوب:

ما هو حجم العينة الواجب اختياره في كل من الحالتيْن التاليتيْن:

(١) في دراسة أُجريت منذ عدة سنوات تبيَّن أن نسبة العملاء المتأخرين عن سداد القروض في موعدها تبلغ ٢٥٪.

(٢) ليست هناك معلومات متاحة لدى مدير الحسابات وأنه يريد التأكد من أنه سوف يحصل على عينة كبيرة كِبَرا كافيا تحقق له ما يريد الوصول إليه.

الامتحان رقم (١١)

المادة: إحصاء تطبيقي　　　　**الفرقة:** الثانية

الكلية: التجارة بقنا　　　　**دور:** مايو ٢٠٠٩م

طلاب الدراسة باللغة العربية

أجب عن الأسئلة التالية:

السؤال الأول:

أولاً: إذا كان الحدثان أ و ب مستقليْن، كما أن:

ع(أ) = ٠,٢　و　ع(ب) = ٠,٤

المطلوب:

احسب قيمة كل من الاحتمالات التالية:

(١) ع(أ و ب)　　**(٢)** ع(أ أو ب)

(٣) ع(حدوث أ على أن يكون ب قد حدث من قبل).

ثانياً: يحتوي صندوق على ٢٠ كرة، كل منها ذات لونيْن. فإذا كان هناك ٥ كرات تحمل كل منها اللونيْن الأحمر والأبيض، و٧ كرات ذوات اللونيْن الأبيض والأسود، و٨ كرات ذوات اللونيْن الأسود والأحمر. فإذا تم اختيار كرة واحدة بطريقة عشوائية، حيث لوحظ أن أحد اللونيْن على هذه الكرة هو اللون الأبيض.

المطلوب:

ما هو احتمال أن يكون اللون الآخر هو الأحمر؟

ثالثاً: يتعرض منتَج معين لنوعيْن من العيوب. فإذا كان:

ج١ = المنتَج به العيب الأول ،　**ج٢** = المنتَج به العيب الثاني ،

ج٣ = المنتَج جيد وليس به أي عيوب

فإذا علمت أن:

ع(**ج١**) = ٠,٠٦　و　ع(**ج٢**) = ٠,٠٧　و　ع(**ج٣**) = ٠,٩

وأنه تم اختيار وحدة واحدة من هذا المنتَج.

المطلوب:

ما هو احتمال أن يكون بالوحدة المُختارة العيبان الأول والثاني؟

السؤال الثاني:

أولاً: اتفق صاحب مصنع مع زبائنه على أن يوافيهم بطلباتهم خلال ٧ أسابيع من تاريخ طلبهم لمنتَجه. ومع ذلك فقد وجد صاحب المصنع أن المدة التي يستغرقها حتى يمد زبائنه بطلباتهم تتبع توزيعاً معتاداً متوسطه ٦ أسابيع وانحرافه المعياري أسبوعان.

المطلوب:

(١) ما هي نسبة (احتمال) الزبائن الذين يستلمون بضاعتهم في موعد متأخر عن الموعد المتَّفق عليه؟

(٢) ما هي النسبة المئوية للزبائن الذين يستلمون بضاعتهم خلال مدة تتراوح ما بين ٤ إلى ٧ أسابيع؟

(٣) إذا كان مطلوباً بأن تكون نسبة من يستلمون بضاعتهم في موعد يتجاوز المدة المتَّفق عليها ٢٠٪ فقط، ما هو التغيير المتوقَّع أن يطرأ على الفترة المتفق عليها (٧ أسابيع)؟

(٤) احسب احتمال أن يستلم زبون معين بضاعته خلال ٥ أسابيع وذلك إذا رغب صاحب المصنع في أن يُنقِص الانحراف المعياري لمدة تسليم البضاعة إلى أسبوع واحد فقط مع الإبقاء على متوسط المدة على ما هو عليه (٦ أسابيع).

ثانياً: للرقابة على جودة المنتَج في أحد المصانع، يتم أخذ عينة مكونة من ١٠ وحدات من هذا المنتَج كل ساعة. وفى حالة عدم وجود أي وحدة معيبة في العينة فإنه يتم استمرار العملية الإنتاجية، وما عدا ذلك يتم توقف عملية الإنتاج لإجراء التعديلات المناسبة.

المطلوب:

إذا بلغت نسبة المعيب في إنتاج هذا المصنع ١٠٪:

(١) ما هو احتمال توقف عملية الإنتاج بعد فحص إحدى العينات المشار إليها؟

(٢) ما هو حجم العينة الواجب أخذها وذلك حتى نكون متأكدين بأن احتمال توقف عملية الإنتاج هو ٠,٩٩؟

ثالثـا: إذا كان عدد مرات تعطل آلة معينة خلال أسبوع يتبع توزيع بواسون بمتوسط مقداره ٠,٥

المطلوب:

ما هو احتمال تعطل الآلة مرتان أو أكثر خلال أسبوع معين؟

السؤال الثالث:

يشترى صاحب مصنع للسيارات أحد مُكوِّنات موتور السيارة من أحد المورِّدين. وقد تعهد المورِّد لصاحب المصنع بأن نسبة المعيب في بضاعته لن تتجاوز ٥٪. وللتأكد من صحة قول المورد، وقبل أن يدفع صاحب المصنع قيمة البضاعة التي استلمها بالفعل، قام باختيار عينة عشوائية مكونة من ٢٠٠ وحدة من هذا المكوِّن حيث وجد من بينها ١٨ وحدة معيبة.

المطلوب:

(١) ما هو التقدير بنقطة لنسبة المعيب في بضاعة المورِّد ككل (ل)؟

(٢) كون فترة الثـقة ٩٥٪ للنسبة في المجتمع (ل)، ثم فسِّر ما تحصل عليه من نتيجة.

(٣) هل تعطى نتـائج العينة دليلا كافيا للقول بأن المورِّد لم يُنفِّذ ما اتفق عليه مع صاحب المصنع فيما يتعلق بنسبة المعيب في بضاعته؟ استخدم مستوى المعنوية ٠,٠١

(٤) ما هو الحـد الأدنى لحجم العينة اللازم لتقدير نسبة المعيب في بضاعة المورِّد ككل وذلك في حدود خطأ مقداره ٠,٠٣ وبدرجة ثـقة ٩٥٪، وذلك في كل من الحالتيْن التاليتيْن:

أ- استخدام النسبة في العينة ($\hat{ل}$) كمقدِّر للنسبة في المجتمع (ل).

ب- عدم توافر معلومات عن النسبة في المجتمع.

السؤال الرابع:

يعتبر ناديا الأهلي والزمالك أشهر الأندية في الدوري المصري لكرة القدم. ويرى مشجعو النادي الأهلي بأن فريقهم هو أفضل من فريق الزمالك. وللتحقق من صحة ذلك، تم الرجوع إلى سجل الإحصاءات المتعلقة بعينة من المباريات السابقة في كرة القدم لكل من الفريقيْن. حيث تم حصر عدد مرات الهجوم على مرمى الخصم في كل مباراة لكل فريق. كما تم حساب الوسط الحسابي والانحراف المعياري لعدد مرات الهجوم على مرمى الخصم لكل فريق، حيث كانت النتائج التالية **(بيانات افتراضية)**.

فريق كرة القدم	عدد مرات الهجوم على مرمى الخصم		عدد المباريات
	الوسط الحسابي	الانحراف المعياري	
الأهلي	$س_١ = ٣٠$	$ع_١ = ٥$	$ن_١ = ٥٠$
الزمالك	$س_٢ = ٢٤$	$ع_٢ = ٨$	$ن_٢ = ٥٠$

فإذا كان:

$م_١$ متوسط عدد مرات هجوم فريق النادي الأهلي على مرمى الخصم في المباريات عموماً (جميع المباريات)،

$م_٢$ متوسط عدد مرات هجوم فريق نادى الزمالك على مرمى الخصم في المباريات عموماً (جميع المباريات).

المطلوب:

(١) إيجاد فترة الثـقة ٩٩٪ للفرق بين المتوسطيْن ($م_١$ – $م_٢$).

(٢) ما هو الحد الأقصى للخطأ المتعلق بالتقدير في (١)؟

(٣) باستخدام نتيجتك التي حصلت عليها في (١)، هل يمكن القول بأن الفريقيْن لهما نفس المستوى؟ وضِّح.

(٤) هل تعطى بيانات العينة الأولى دليلاً كافياً على أن متوسط عدد مرات هجوم فريق النادي الأهلي على مرمى الخصم ($م_١$) يزيد على ٢٧ هجمة؟ استخدم مستوى المعنوية ٠,٠٥

(٥) هل تؤيد نتائج العينتيْن صحة رأى مشجعي النادي الأهلي بأن فريقهم يتفوق إلى حد بعيد على فريق نادى الزمالك؟ استخدم في ذلك مستوى المعنوية ٠,٠١

الامتحان رقم (١٢)

المادة: إحصاء تطبيقي　　　　**الفرقة:** الثانية

الكلية: التجارة بقنا　　　　**دور:** مايو ٢٠١٣م

طلاب الدراسة باللغة العربية

أجب على الأسئلة التالية:

السؤال الأول:

أولاً: إذا كان الحدثان ب و جـ مستقليْن، كما أن:

ع(ب و جـ) = ٠,٤ ، ع(ب) + ع(جـ) = ٠,٦

أثبت أن الحدثيْن ب و جـ متنافيان.

ثانياً: صندوق يحتوي على ٣ كرات حمراء، ٤ كرات بيضاء، و٥ كرات سوداء. فإذا تم سحب كرتيْن من الصندوق وبدون إرجاع الكرة الأولي، ما هو احتمال أن تكون الكرتيْن لهما نفس اللون؟

ثالثاً: تم سحب ورقتيْن من مجموعة أوراق اللعب بدون إرجاع الورقة الأولي، ما هو احتمال أن تكون ورقة واحدة على الأقل عشرة؟ وإذا تم سحب ورقة واحدة، ما هو احتمال أن تكون خمسة بشرط أن تحمل اللون الأحمر؟

رابعاً: إذا كان لدينا زهرتيْ نرد غير متحيزتيْن (جـ ، هـ)، وكانت الزهرة (جـ) تحمل الأرقام من ١إلي٦، بينما تحمل الزهرة الثانية (هـ) الأرقام ١ ، ١ ، ٢ ، ٢ ، ٣ ، ٣. وأنه إذا قمنا باختيار ورقة من مجموعة أوراق اللعب عشوائياً وكانت تحمل علامة القلب فإنه يتم رمي الزهرة (جـ) وإلَّا فإنه يتم رمي الزهرة الثانية (هـ).ما هو احتمال الحصول علي الرقم ٢؟

السؤال الثاني:

أولاً: فيما يلي التوزيع الاحتمالي للمتغير س والذي يمثل عدد حوادث المرور التي تحدث في إحدى المدن خلال أسبوع.

س	١	٢	٣
ع(س)	٠,٢	٠,٥	٠,٣

المطلوب:

(١) احسب الاحتمالات التالية:

(أ) ع(س = ٢)　　**(ب)** ع(س < ٣)　　**(جـ)** ع(س= ٢ على الأكثر)

(٢) إيجاد الوسط الحسابي والانحراف المعياري للمتغير س.

(٣) إيجاد الوسط الحسابي والانحراف المعياري للمتغير ص، حيث ص = ٢س + ٣.

ثانيـاً: إذا كان لدينا خمسة أشخاص، احسب الاحتمالات التالية:

(١) لا يوجد بينهم أي شخص يوم ميلاده كان يوم سبت.

(٢) هناك شخصان أو أكثر كان يوم ميلاد كل منهم يوم جمعة.

ثالثاً: إذا كان متوسط عدد العيوب في المتر الواحد من القماش الذي تم إنتاجه بواسطة إحدى الآلات هو ٠,٢

(١) احسب الاحتمالات التالية:

أ- وجود عيبيْن في قطعة قماش طولها خمسة أمتار.

ب- عدم وجود أي عيب في قطعة قماش طولها متران.

(٢) ما هو الطول الذي يجب أن تكون عليه قطعة قماش بحيث يكون احتمال عدم احتوائها على أي عيوب أقل من ٠,٨٥؟

رابعاً: تتبع درجات الطلاب في مادة الإحصـــــاء توزيعا طبيعيا متوســـطه ٧٠ وانحرافه المعياري ١٠.

المطلوب:

(١) إذا تم اختيار أحد الطلاب عشوائياً، احسب الاحتمالات التالية:

(أ) أن تكون درجته ٨٥ على الأكثر. **(ب)** أن تكون درجة هذا الطالب أكبر من ٦٠.

(٢) ما هي الدرجة التي يحصل علي أكبر منها ٢٥٪ من إجمالي عدد الطلاب؟

(٣) إيجاد المدي الربيعي لتوزيع الدرجات، المدي الربيعي = الربيع الثالث – الربيع الأول.

السؤال الثالث:

أولاً: يعتبر ناديا الأهلي والزمالك أشهر الأندية المصرية في كرة القدم. هذا ويدَّعي مشجعو النادي الأهلي بأن فريق كرة القدم به هو أفضل امن فيق الزمالك. وللوقوف على مدي صحة هذا الادعاء، تم الرجوع إلى سجل الإحصاءات المتعلقة بعينة عشوائية من المباريات السابقة في كرة القدم لكل من الفريقيْن، حيث تم حصر عدد مرات الهجوم على مرمي الخصم في كل مباراة لكل فريق. والجدول التالي يوضح الوسط الحسابي والانحراف المعياري لعدد مرات الهجوم على مرمي الخصم وذلك لكل من الفريقيْن. **ملحوظة: البيانات افتراضية.**

الفريق	مرات الهجوم على المرمي		عدد المباريات
	الوسط الحسابي	الانحراف المعياري	
الأهلي	$س_١$ = ٢٠	$ع_١$ = ٤	$ن_١$ = ٣٦
الزمالك	$س_٢$ = ١٦	$ع_٢$ = ٦	$ن_٢$ = ٣٦

فإذا كان:

$م_1$ هو الوسط الحسابي لعدد مرات الهجوم على مرمي الخصم وذلك لجميع مباريات النادي الأهلي على الأمد الطويل.

$م_2$ هو الوسط الحسابي لعدد مرات الهجوم على مرمي الخصم وذلك لجميع مباريات نادي الزمالك على الأمد الطويل.

المطلوب:

(١) إيجاد فترة الثقة ٩٥٪ للمتوسط $م_1$. ما هو الحد الأقصى للخطأ؟

(٢) ما هو حجم العينة الواجب أخذه وذلك لإنقاص طول فترة الثقة المحسوبة في (١) بمقدار ٠,٥؟ **ملحوظة: بقية المُعطيات كما هي في (١).**

(٣) كوِّن فترة الثقة ٩٩٪ للفرق بين متوسطي المجتمعيْن ($م_1$ – $م_2$).

(٤) من نتائجك في (٣)، هل يمكنك استنتاج أن فريقي الأهلي والزمالك على نفس المستوي؟

(٥) بيِّن هل من الممكن الوصول إلى نتيجة مختلفة عن تلك التي وصلت إليها في (٤) وذلك في حالة استخدام درجة الثقة ٩٥٪ بدلاً من درجة الثقة ٩٩٪؟

ملحوظة: بدون إجراء أي حسابات.

(٦) هل تعطي العينة الثانية دليلاً كافياً للقول بأن الوسط الحسابي لعدد مرات الهجوم على مرمي الخصم وذلك لجميع مباريات نادي الزمالك على الأمد الطويل يقل عن ١٨هجمة؟ استخدم مستوي المعنوية ٠,٠١

(٧) باستخدام مستوي المعنوية ٠,٠٥، اختبر صحة الادعاء القائل بأن فريق الأهلي يتفوق علي فريق الزمالك بدرجة كبيرة.

ثانياً: يرغب مدير عمليات أحد المصانع الكبيرة في تقدير متوسط الوقت الذي يقوم فيه العامل بتجميع مُكوِّنات أحد الأجهزة الإلكترونية مستخدماً درجة الثقة ٩٥٪، ما هو حجم العينة الذي يجب أن يأخذه المدير إذا كان لا يرغب في أن يتجاوز خطأ التقدير ٢٤ ثانية؟ وذلك في كل من الحالتيْن التاليتيْن:

(١) أنه من دراسة سابقة، يعرف مدير العمليات أن الانحراف المعياري للوقت اللازم لتجميع مُكوِّنات جهاز إلكتروني مشابه هو دقيقتان.

(٢) أنه بعد مشاهدة عدد من العمال يقومون بتجميع مُكوِّنات جهاز مشابه تبيَّن أن أقل وقت استغرقته عملية التجميع كان ٨ دقائق بينما كان أكبر وقت ١٨ دقيقة.

السؤال الرابع:

أولاً: ترغب إدارة أحد المحال التجارية الكبيرة في التحقق من أن النسبة المئوية للرجال تختلف عن النسبة المئوية للنساء وذلك فيما يتعلق بتفضيل شراء المنتجات الوطنية. وفي هذا الشأن،

تم أخذ عينة مكوَّنة من ٤٠٠ رجل حيث وُجِد أن من بينهم ١٦٠ رجلاً يفضلون شراء المنتجات الوطنية. كما تم أخذ عينة مكوَّنة من ٢٥٠ امرأة تبيَّن أن من بينهن ٩٠ امرأة تفضلن شراء المنتجات الوطنية.

فإذا كانت:

ل١ هي نسبة الرجال الذين يفضلون شراء المنتجات الوطنية وذلك في مجتمع الرجال ككل،

ل٢ هي نسبة النساء اللائي تفضلن شراء المنتجات الوطنية وذلك في مجتمع النساء ككل،

المطلوب:

(١) ما هو التقدير بنقطة للفرق بين نسبتي المجتمعين (ل٢ – ل١)؟

(٢) كوِّن فترة الثقة ٩٥٪ للفرق بين نسبة الرجال الذين يفضلون شراء المنتجات الوطنية وبين تلك النسبة المتعلقة بالنساء (ل١ – ل٢) وذلك بشكل عام.

(٣) باستخدام مستوي المعنوية ٠,٠٥، هل يمكن استنتاج أن نسبة الذين يفضلون شراء المنتجات الوطنية بين الرجال ككل تختلف عن نسبة اللائي تفضلن شراء المنتجات الوطنية بين النساء ككل؟

(٤) باستخدام نتائجك في (٢)، بيِّن كيف تؤكد ما توصلت إليه من نتائج في (٣).

ثانياً: تم أخذ عينة عشوائية حجمها ٤٠٠ سمكة من إحدى البحيرات، حيث تم تثبيت علامة مميِّزة علي جسم كل سمكة ثم إرجاعها إلي البحيرة. وبعد فترة مناسبة، أُخذت عينة أخري مكوَّنة من ٤٠٠ سمكة أيضاً حيث وُجد أن من بينها ١٦ سمكة ممن تم تثبيت العلامة المميِّزة لها من قبل.

المطلوب:

(١) ما هو التقدير بنقطة لنسبة السمك المُثبت على جسمه العلامة المميِّزة في البحيرة؟

(٢) إيجاد تقدير لعدد السمك في البحيرة.

(٣) تكوين فترة الثقة ٩٥٪ لعدد السمك في البحيرة.

الامتحان رقم (١٣)

المادة: إحصاء تطبيقي　　　　**الفرقة:** الثانية
الكلية: التجارة بقنا　　　　**دور:** يونيو ٢٠١٤م

طلاب الدراسة باللغة العربية

أجب على الأسئلة التالية:

السؤال الأول:

أولا: لاعبان (ب) و(جـ) يتنافسان في عدد من مباريات تنس الطاولة. فإذا افترضنا أن هذه المباريات مستقلة، وأنه في كل مباراة يكون احتمال فوز اللاعب (ب) هو ٠,٣ بينما احتمال أن يفوز اللاعب (جـ) هو ٠,٢، وفي حالة عدم فوز أي منهما بالمباراة تعتبر النتيجة تعادلاً.

(١) إيجاد الاحتمالات التالية فيما يتعلق بمباراة واحدة أُجريت بينهما.

(أ) نتيجة المباراة هي التعادل.

(ب) نتيجة المباراة هي التعادل أو فوز المتباري (ب) بها.

(٢) إذا لعب (ب) و (جـ) مباراتيْن، ما هو احتمال فوز المتباري (جـ) بكليهما.

ثانياً: صندوق يحتوي على ١٠ كرات كل منها تحمل لونيْن. فإذا كان هناك كرتان تحمل كل منهما اللونيْن الأحمر والأبيض، وثلاث كرات ذات اللونين الأبيض والأسود، وخمس كرات تحمل كل منها اللونيْن الأسود والأحمر. فإذا تم اختيار كرة واحدة بطريقة عشوائية، حيث لوحظ أن أحد اللونيْن على هذه الكرة هو اللون الأحمر. ما هو احتمال أن يكون اللون الآخر على هذه الكرة هو اللون الأبيض؟

ثالثاً: الجدول التالي يوضح الحالة الاجتماعية لعينة مكونة من ٢٠٠ شخص موزعين حسب الجنس.

الجنس	الحالة الاجتماعية				المجموع
	أعزب (ع)	متزوج (ج)	أرمل (ل)	مطلق (ط)	
ذكر (ك)	١٠	٥٠	١٢	٨	٨٠
أنثى (ث)	١٥	٧٥	١٨	١٢	١٢٠
المجموع	٢٥	١٢٥	١٨	١٢	٢٠٠

(١) إذا تم اختيار أحد الأشخاص عشوائياً، ما هو احتمال أن يكون هذا الشخص:

(أ) أنثى　**(ب)** مطلقاً بشرط أن يكون ذكراً　**(جـ)** أنثى أو متزوج

(٢) باستخدام نتائجك في (١ - ب)، هل الحدثان "مطلق" و"ذكر" مستقلان؟

(٣) هل الحدثان "أنثى" و "أعزب" مانعان (متنافيان)؟

السؤال الثاني:

أولاً: إذا كانت الحوادث على أحد الطرق السريعة تقع بمعدل ١,٢ أسبوعياً. احسب الاحتمالات التالية:

(١) وقوع حادث واحد خلال أسبوع.

(٢) وقوع حادث واحد على الأقل خلال مدة مقدارها ٤ أسابيع.

ثانياً: إذا كان المتغير س له توزيع ثنائي الحديْن فيه ٮ = ٤.

فإذا كان ع(س = ٢) = ٠,٣٤٥٦، ما هي قيمة ل؟

ثالثاً: إذا كانت الأجور اليومية (س) لعمال الإنتاج في أحد المصانع تتبع توزيعاً طبيعياً متوسطه ١٠٠ جنيه وانحرافه المعياري ١٠ جنيهات.

المطلوب:

(١) إذا تم اختيار أحد العمال عشوائياً، احسب الاحتمالات التالية:

(أ) أن يكون أجره اليومي ١٢٠ جنيهاً أو اقل.

(ب) أن يكون أجره اليومي أكبر من ٧٢ جنيهاً.

(٢) حدِّد الأجر الذي يجب أن يحصل عليه عامل حتى يكون من بين الـ ١٠٪ الأعلى أجراً بين العمال؟

(٣) ما هما الأجران (ب) و (جـ) اللذان يحصران بينهما أجور ٩٥٪ من العمال ويقعان على بعديْن متساوييْن من الوسط الحسابي. أي أن: ع(س > جـ) = ع(س ≤ ب)؟

السؤال الثالث:

أولاً: ترغب شركة كبرى للخدمات العامة في مقارنة مقدار استهلاك الكهرباء (جنيه مصري) في فصل الصيف وذلك في مدينتيْن (١) و (٢). ولتحقيق ذلك تم اختيار عينة عشوائية من الأسر من كل مدينة حيث قامت الشركة بتسجيل قيمة فاتورة الكهرباء الشهرية لكل أسرة فكانت النتائج التالية:

المدينة	**الوسط الحسابي**	**الانحراف المعياري**	**عدد الأسر (الفواتير)**
المدينة (١)	$س_{١}$ = ٥٠	$ع_{١}$ = ١٨	$ٮ_{١}$ = ٤٠٠
المدينة (٢)	$س_{٢}$ = ٤٨	$ع_{٢}$ = ١٢	$ٮ_{٢}$ = ٣٢٤

فإذا كان:

$م_{١}$ هو الوسط الحسابي للفواتير الشهرية في المدينة (١)،

$م_{٢}$ هو الوسط الحسابي للفواتير الشهرية في المدينة (٢).

المطلوب:

(١) إيجاد فترة الثقة ٩٩٪ للمتوسط $م_٢$. ما هو الحد الأقصى للخطأ؟

(٢) ما هو حجم العينة الواجب أخذه وذلك لإنقاص طول فترة الثقة في (١) لـ ٢,٥؟
ملحوظة: بقية المُعطيات كما هي في (١).

(٣) كوِّن فترة الثقة ٩٥٪ للفرق بين متوسطي المجتمعيْن ($م_٢$ – $م_١$).

(٤) من نتائجك في (٣)، هل يمكنك استنتاج أن المتوسطيْن $م_١$ و $م_٢$ متساويان؟ وضِّح.

(٥) باستخدام مستوي المعنوية ٠,٠٥ ، هل هناك من الأدلة ما يكفي للقول بأن المتوسط الشهري للفواتير في المدينة (١) يزيد علي ٤٧ جنيهاً؟

(٦) من نتائج العينتيْن، هل يمكنك القول بأن المتوسط الشهري للفواتير في المدينة (٢) يقل عن نظيره في المدينة (١)؟ استخدم مستوي المعنوية ٠,٠١

(٧) بافتراض أن هذه الشركة ترغب في تقدير المتوسط الشهري للفواتير في المدينة (١) وذلك في حدود خطأ لا يتجاوز ١,٥ جنيه وبدرجة ثقة ٩٥٪، وأن الشركة لا تتوافر لديها البيانات السابقة. لذلك فقد قدرت الشركة الانحراف المعياري بطريقتها الخاصة ليكون ١٥ جنيهاً. ما هو حجم العينة المطلوب في هذه الحالة؟

ثانياً: تهتم منظمة لقياس الرأي العام بتقدير نسبة الناخبين الذين سيدلون بأصواتهم لصالح المرشح (أ) وذلك في حملة رئاسية. ولتحقيق هذا الغرض، تم استطلاع آراء عينة مكوَّنة من ٥٠٠ ناخب حيث تبين أن ٧٠٪ منهم يؤيدون هذا المرشح.

المطلوب:

(١) تحديد فترة الثقة **٩٩**٪ للنسبة في المجتمع (ل).

(٢) باستخدام مستوي المعنوية ٠,٠٥ ، هل يمكنك استنتاج أن نسبة المؤيدين لهذا المرشح في المجتمع تزيد علي ٦٥٪ ؟

(٣) ما هو حجم العينة الواجب أخذه إذا رغبت هذه المنظمة وبدرجة ثقة ٩٥٪ في أن لا يختلف تقديرها لنسبة المؤيدين لهذا المرشح عن النسبة الحقيقية في المجتمع بأكثر من ٠,٠٣؟ **ملحوظة:** يتم استخدام نسبة المؤيدين في العينة كتقدير لنسبة المؤيدين في المجتمع.

(٤) في حالة عدم إجراء المنظمة لهذه الدراسة الاستطلاعية، ما هو حجم العينة الذي يمكن الحصول عليه في (٣)؟ قارن بين نتائجك في الحالتيْن مع إبداء التعليق المناسب.

الامتحان رقم (١٤)

المادة: إحصاء تطبيقي **الفرقة:** الثانية
الكلية: التجارة بقنا **دور:** مايو ٢٠١٦م

طلاب الدراسة باللغة العربية

أجب عن الأسئلة التالية:

السؤال الأول:

أولا: ثلاثة موظفين (أ) و (ب) و(جـ) يشتركون في مكتب به تليفون واحد. وبشكل عشوائي، يكون ورود المكالمات لهذا المكتب وفقاً للنسب ٠,٤ للموظف (أ) و ٠,٤ للموظف (ب) و ٠,٢ للموظف (جـ). كما أن عملهم يتطلب مغادرة المكتب في أوقات عشوائية، بحيث يكون الموظف (أ) خارج المكتب نصف وقت عمله بينما يكون كل من الموظفيْن (ب) و (جـ) خارج المكتب ربع وقت كل منهما. وفيما يتعلق بالمكالمات الواردة خلال ساعات العمل، احسب الاحتمالات التالية:

(١) عدم الرد على التليفون من قِبَل الموظفين الثلاثة.
(٢) سوف يتم الرد على مكالمة واحدة بواسطة الموظف الواردة له هذه المكالمة.
(٣) ورود ثلاث مكالمات متتالية لنفس الموظف.
(٤) ورود ثلاث مكالمات متتالية لموظفيْن مختلفيْن.
(٥) المتصل الذي يرغب في محادثة الموظف (ب) يُضطَّر الي المحاولة أكثر من ثلاث مرات حتى يرد عليه هذا الموظف.

ثانياً: في مجموعة من ١٠ أشخاص، هناك ٣ منهم يؤيدون الحزب (س)، ٥ يؤيدون الحزب (ص) واثنان لم يدلوا برأيهم. فإذا تم اختيار شخصيْن على التوالي (بدون إرجاع).

المطلوب:

(١) إذا كان الأول مؤيداً للحزب (س)، ما هو احتمال أن يكون الثاني مؤيداً للحزب (ص)؟
(٢) احسب احتمال أن أحدهما يؤيد الحزب (س) بينما يؤيد الآخر الحزب (ص).
(٣) ما هو احتمال أن يكون واحد منهما على الأقل من أنصار الحزب (ص)؟

ثالثاً: وضِّح أن ع(ب أو جـ) يمكن كتابته على الصورة:

ع(ب أو جـ) = ع(ب) + ع(جـ)[١ − ع(ب | جـ)]

رابعاً: إذا كان ٤٠٪ من زبائن مطعم الجامعة يقومون بطلب وجبات ساخنة. كما أن الطلاب يمثلون ٨٠٪ من إجمالي زبائن المطعم. فإذا علمت أن ٣٠٪ من الطلاب زبائن المطعم يطلبون وجبات ساخنة.

المطلوب:

(١) إذا تم اختيار أحد زبائن المطعم عشوائياً، احسب الاحتمالات التالية:

أ- أن يكون طالباً ويطلب وجبات ساخنة.

ب- إذا كان من تم اختياره يطلب وجبات ساخنة، ما هو احتمال أن يكون طالباً؟

ج- ليس طالباً ولا يطلب وجبات ساخنة.

ملحوظة: من أجل حل أيسر، يمكنك تكوين جدول ٢ × ٢ لعرض البيانات مُفترضاً أن إجمالي عدد الزبائن هو ١٠٠٠ زبون.

(٢) هل الحدثان **"طالب"** و **"يطلب وجبات ساخنة" مستقلان؟** برِّر إجابتك.

السؤال الثاني:

أولاً: إذا كنت بصدد القيام بالاستثمار في أحد نشاطيْن (س) أو (ص). فإذا كان التوزيع الاحتمالي للعائد من استثمار ١٠٠٠ في كل منهما كما يلي:

العائد (بالجنيه المصري)	**النشاط (س)**	٥٠	١٠٠	٢٠٠
	النشاط (ص)	٨٠	١٢٥	١٤٠
الاحتمال		٠,١	٠,٦	٠,٣

المطلوب:

(١) أوجد الاحتمالات التالية:

(أ) ع(ص = ١٠٠) **(ب)** ع(س ≤ ٥٠)

(٢) احسب تباين المتغير ف، حيث ف = ٢ص – ١٥٠.

(٣) أي النشاطيْن الأفضل لك من حيث الاستثمار فيه؟ برّر إجابتك.

ثانياً: تم رمي زهرة نرد ٣ مرات. وتعتبر الرمية حالة نجاح إذا ظهر الرقم ١ أو الرقم ٢، بينما تعتبر حالة فشل إذا ظهر أي من الأرقم ٣ ، ٤ ، ٥ أو ٦. احسب الاحتمالات التالية:

(١) الحصول على حالتيْ نجاح. **(٢)** الحصول على حالة فشل واحدة.

وضِّح العلاقة بين إجابتك في (١) وإجابتك في (٢) مُبدياً ما تراه مناسباً من تعليق.

ثالثاً: أفاد مدير مركز للحاسب الآلي بأن نظام الحاسب لديه قد تعرض الي ٣ مرات للتوقف خلال الـ ١٠٠ يوم الماضية. احسب الاحتمالات التالية:

(١) عدم توقف النظام خلال يوم معين.

(٢) توقف النظام مرتان على الأقل خلال مدة ٣ أيام.

رابعاً: الطلب المُتوقَّع (س) على أحد المُنتَجات خلال شهر قادم يمكن تمثيله بمتغير عشوائي له توزيع طبيعي متوسطه ١٢٠٠ وحدة وانحرافه المعياري ١٠٠ وحدة.

المطلوب:

(١) ما هو احتمال أن يزيد الطلب على ١٣٠٠ وحدة؟

(٢) احسب احتمال أن يتراوح الطلب ما بين ١٠٠٠ وحدة و ١٣٠٠ وحدة.

(٣) إذا كان ح(س > جـ) = ٠,٩ ، حدِّد مستوي الطلب جـ.

السؤال الثالث:

أولاً: يرغب مدير التسويق في إحدى الشركات في إجراء مقارنة بين سعر مُنتَج الشركة (ب) وسعر مُنتَج آخر (جـ) منافس له. ولتحقيق ذلك قام بأخذ عينة عشوائية من كل مُنتَج فكانت النتائج التالية بالنسبة للأسعار:

	المنتَج (ب)	**المنتَج (جـ)**
الوسط الحسابي للعينة (س)	٥٠	٤٦
الانحراف المعياري للعينة (ع)	١٢	١٠
حجم العينة (ن)	١٠٠	١٢٥

فإذا كان: م$_{ب}$ هو متوسط السعر للمُنتَج (ب) ، م$_{جـ}$ هو متوسط السعر للمُنتَج (جـ).

المطلوب:

(١) إيجاد فترة الثـقة ٩٥٪ للمتوسط م$_{جـ}$.

(٢) هل يمكن للباحث استنتاج أن متوسط المُنتَج جـ (م$_{جـ}$) يختلف عن ٤٥؟ مستخدماً في ذلك مستوي المعنوية ٠,٠٥

(٣) علِّق على نتيجتك في (٢) وذلك بالمقارنة بما حصلت عليه في (١).

(٤) بافتراض أن طول فترة الثـقة في (١) أكبر مما ينبغي، وأنه يُراد إنقاص هذا الطول بمقدار ٠,٥، ما هو حجم العينة المطلوب في هذه الحالة؟ ملحوظة: بقية المُعطيات كما هي في (١)، وأنه يتم استخدام الانحراف المعياري للعينة كتقدير للانحراف المعياري للمجتمع.

(٥) ما هو التقدير بنقطة للفرق بين المتوسطيْن (م$_{جـ}$ – م$_{ب}$) ؟

(٦) كوِّن فترة الثـقة ٩٩٪ للفرق بين متوسطي المجتمعيْن (م$_{ب}$ – م$_{جـ}$). هل من الممكن أن نكون على ثـقة ٩٩٪ بأن المتوسطيْن مختلفان؟ برِّر إجابتك.

(٧) استخدم اختبار الفرض الملائم لمعرفة مدي صحة الادعاء القائل بأن متوسط السعر للمُنتَج جـ أقل من نظيره للمُنتَج ب وذلك بمستوي معنوية ٠,٠١

ثانياً: يود باحث اجتماعي معرفة آراء الناس فيما يتعلق بظاهرة العنف المُبالغ فيه في أجهزة الإعلام المرئية هذه الأيام. وقد وجد هذا الباحث أن ٦٠٪ من أفراد عينة عشوائية حجمها

٢٠٠ رجل يروْن وجود هذه الظاهرة فعلاً، بينما بلغت تلك النسبة ٨٠٪ من أفراد عينة عشوائية حجمها **٢٥٠ امرأة**.

فإذا كانت **ل**$_١$ و **ل**$_٢$ هما **نسبة** الذين يُقِرون بوجود هذه الظاهرة وذلك **للرجال والنساء ككل** علي الترتيب.

المطلوب:

(١) تحديد فترة الثقة ٩٩٪ للفرق بين النسبتيْن (**ل**$_١$ – **ل**$_٢$).

(٢) باستخدام مستوي المعنوية ٠,٠٥ ، هل يمكنك استنتاج أن هناك فرقاً جوهرياً بين آراء الرجال ككل وآراء النساء ككل تجاه تلك القضية ؟

(٣) وضِّح ما إذا كان هناك تغيير في نتيجتك في (٢) إذا تغير مستوي المعنوية ليكون ٠,٠١.

الامتحان رقم (١٥)

المادة: إحصاء تطبيقي **الفرقة:** الثانية
الكلية: التجارة بقنا **دور:** مايو ٢٠١٩م

طلاب الدراسة باللغة العربية

اختر الإجابة الصحيحة لكل سؤال من الأسئلى التالية:

(١) تُعتبر القيمة العددية لـ "متوسط عينة"

(A) معلمة مجتمع **(B)** معلمة عينة **(C)** إحصاء عينة **(D)** لا شيء مما سبق

(٢) حيث أن حجم المجتمع يكون دائماً أكبر من حجم العينة، فإن إحصاء العينة لا يمكن أن يكون ...... معلمة المجتمع. املأ الفراغ باختيار الإجابة الملائمة مما يلي:

(A) أكبر من **(B)** مساوياً **(C)** أقل من **(D)** لا شيء مما سبق

(٣) إذا كان حدوث حدث ما يعني عدم حدوث آخر، فإن هذيْن الحدثيْن

(A) مستقلان **(B)** متنافيان (مانعان) **(C)** تجريبيان **(D)** لا شيء مما سبق

(٤) المساحة تحت أحد جانبي المنحنى الطبيعي تساوي

(A) ٠,٥ **(B)** ١ **(C)** قيمة الوسط الحسابي **(D)** ٢

(٥) قيمة الوسيط للتوزيع الطبيعي المعياري (القياسي) هي

(A) م (متوسط المجتمع) **(B)** صفر **(C)** ١ **(D)** ٠,٥

(٦) عند قياس احتمال حدوث أي حدث، فإن الصفر يمثل

(A) أحداث مستحيلة **(B)** أحداث ممكنة **(C)** أحداث مُعيَّنة **(D)** أحداث عينة

(٧) العبارة القائلة بأن "ع(ب | جـ) = ع(جـ | ب)" حيث ب و جـ حدثان مستقلان

(A) صحيحة دائماً

(B) غير صحيحة على الإطلاق

(C) المعلومات غير كافية؛ يجب معرفة ما إذا كان الحدثان متنافييْن (مانعيْن)

(D) المعلومات غير كافية؛ يجب معرفة ما إذا كان الحدثان متساوييْن من حيث احتمال الحدوث

(٨) لنفترض أنه تم رفض فرضية العدم عند مستوى المعنوية ٥٪ ، فإنه عند أي مستوى من مستويات المعنوية التالية يتم أيضا رفض فرضية العدم؟

(A) ٤٪ **(B)** ٢,٥٪ **(C)** ٦٪ **(D)** ٣٪

(٩) ما هي الصيغة غير الصحيحة مما يلي عند صياغة فرضية العدم؟

(A) ف٠: = $\hat{ل}_1 - \hat{ل}_2$ = صفر **(B)** ف٠: م = ١٠

(C) ف٠: = $م_1 - م_2$ = صفر **(D)** ف٠: ل = ٠,٥

(١٠) أفضل طريقة لتحديد ما إذا كان هناك فرق جوهري ما بين متوسطي المرتبات للرجال والنساء هي

(A) إيجاد الفرق بين متوسطي العينتيْن

(B) إجراء اختبار: "تساوي متوسطي المجتمعيْن" مقابل: "متوسط الرجال هو الأكبر"

(C) إجراء اختبار: "تساوي متوسطي المجتمعيْن" مقابل: "متوسط النساء هو الأكبر"

(D) إجراء اختبار: "تساوي متوسطي المجتمعيْن" مقابل: "المتوسطان مختلفان"

(١١) التوزيع الطبيعي المعياري (القياسي)

(A) يكون ملتوياً جهة اليسار (B) تباينه يساوي واحد

(C) انحرافه المعياري يساوي الصفر (D) وسطه الحسابي يساوي واحد

(١٢) إذا زاد حجم العينة فإن دقـة فترة الثـقة

(A) تزيد (B) تنقص

(C) أحياناً تزيد وأحياناً تنقص (D) لا يطرأ عليها تغيير

(١٣) فرضية العدم والفرضية البديلة تتعلقان بـ:

(A) إحصاءات العينة (B) معلمات العينة

(C) معلمات المجتمع

(D) أحياناً معلمات المجتمع وأحياناً أخرى إحصاءات العينة

(١٤) العبارة القائلة " إذا كان هناك دليل كاف لرفض فرضية العدم عند مستوى المعنوية ٥٪، فإنه يكون هناك دليل كاف لرفضها عند مستوى المعنوية ١٪ " هي:

(A) دائماً صحيحة

(B) غير صحيحة على الإطلاق

(C) أحياناً تكون صحيحة، يجب معرفة قيمة إحصاء الاختبار للوصول إلى نتيجة

(D) المعلومات غير كافية، حيث يعتمد ذلك على نوع الاختبار

(١٥) جميع ما يلي من شأنه أن يُـزيد طول فترة الثـقة ما عدا

(A) زيادة درجة الثـقة (B) زيادة التشتت

(C) زيادة حجم العينة (D) نقصان حجم العينة

إذا تم إلـقاء زهرتـيْ نرد (ب و جـ).

أجب عن السؤاليْن التاليين (السؤال ١٦ والسؤال ١٧):

(١٦) ما هو الاحتمال: ع(ب + جـ = ٨ | ب = رقم زوجي)

(A) $\frac{٥}{١٨}$ (B) $\frac{١}{٦}$ (C) $\frac{٥}{٦}$ (D) $\frac{٤}{٩}$

(١٧) احتمال ظهور الرقم ٤ على زهرة واحدة على الأقل أو أن مجموع رقمي الزهرتيْن يساوي ٧ هو:

(A) $\frac{١}{٩}$ **(B)** $\frac{٤}{٩}$ **(C)** $\frac{٥}{١٢}$ **(D)** $\frac{٧}{١٢}$

إذا تم اختيار ورقة واحدة من مجموعة أوراق اللعب.
أجب عن السؤاليْن التاليين (السؤال ١٨ والسؤال ١٩):

(١٨) ما هو احتمال الحصول على ١٠؟

(A) $\frac{١٠}{٥٢}$ **(B)** $\frac{٤}{٥٢}$ **(C)** $\frac{١}{٥٢}$ **(D)** $\frac{٢}{١٣}$

(١٩) إذا كانت الورقة تحمل اللون الأحمر، ما هو احتمال أن تكون من نوع القلب؟

(A) ٠,٥ **(B)** ٠,٣ **(C)** ٠,٦ **(D)** ٠,٤

(٢٠) صندوق يحتوي على كرة واحدة حمراء و ٢ كرة بيضاء و ٢ كرة سوداء. فإذا تم سحب كرتيْن (بدون إحلال)، ما هو احتمال أن يكون لونهما مختلف؟

(A) ٠,٦ **(B)** ٠,٤ **(C)** ٠,٨ **(D)** ٠,٢

(٢١) إذا كان لديك اختبار فيه ٤ اختيارات للإجابة على كل من أسئلة الاختبار، من بينها اختيار واحد صحيح. فإذا كان احتمال معرفة الإجابة الصحيحة للسؤال تساوي ٠,٦ وفي حالة عدم معرفة الإجابة يتم اختيارها عشوائيا من بين الاختيارات الأربعة. ما هو احتمال أن تكون قد عرفت الإجابة على السؤال إذا كانت إجابتك عليه صحيحة؟

(A) ٠,٨٢ **(B)** ٠,٨٤ **(C)** ٠,٨٨ **(D)** ٠,٨٦

(٢٢) يتعرض مُنتَج معين عند إنتاجه لنوعيْن من العيوب ($ب_١$ و $ب_٢$). فإذا كان:

$ب_١$ = المُنتَج به العيب الأول ، $ب_٢$ = المُنتَج به العيب الثاني،

جـ = المُنتَج خالٍ من العيوب.

فإذا علمت أن: ع($ب_١$) = ٠,٠٥ ، ع($ب_٢$) = ٠,٠٧ ، ع(**جـ**) = ٠,٨

ما هو احتمال أن يحتوي المُنتَج علي العيبيْن معا؟

(A) ٠,٠٨ **(B)** ٠,٠٤ **(C)** ٠,٠٦ **(D)** ٠,١٢

فيما يلي التوزيع الاحتمالي للمتغير س:

س	صفر	١	٢	٣
ع(س)	٠,٠٥	٠,١٥	٠,٦٠	؟

أجب عن السؤاليْن التاليين (السؤال ٢٣ والسؤال ٢٤):

(٢٣) الاحتمال: ع(٣ > س ≥ ١) هو

(A) ٠,٨ **(B)** ٠,٢٠ **(C)** ٠,٧٥ **(D)** ٠,٦٥

(٢٤) إذا كان لديك المُتغيِّر ص حيث: ص = ٤ س − ٢ ، ما هو تباين ص؟

مع **عدم التقريب في** حساباتك.

(A) ٩,٢٤ **(B)** ٨,٧٦ **(C)** ٦,٢٥ **(D)** ٧,٨٥

من خبرات سابقة، تَبيَّن أن ٧٠٪ من المُتقدمين لإحدى الوظائف يجتازون اختباراً معيناً. وأنه تقدم لتلك الوظيفة **خمسة** أشخاص.

أجب عن السؤاليْن التاليين (السؤال ٢٥ والسؤال ٢٦):

(٢٥) ما هو احتمال اجتياز جميع المُتقدمين لهذا الاختبار؟

(A) ٠,١٨٥ **(B)** ٠,١٧٥ **(C)** ٠,١٨٢ **(D)** ٠,١٦٨

(٢٦) احتمال أن واحداً على الأقل سوف يفشل في اجتياز هذا الاختبار.

(A) ٠,٨١٨ **(B)** ٠,٨٣٢ **(C)** ٠,٨١٥ **(D)** ٠,٨٢٥

إذا كان متوسط عدد العيوب في المتر الواحد من القماش الذي تم إنتاجه بواسطة إحدى الآلات هو ٠,٢

أجب عن السؤاليْن التاليين (السؤال ٢٧ والسؤال ٢٨):

(٢٧) احتمال عدم وجود أي عيوب في قطعة قماش طولها ٤ أمتار هو

(A) ٠,٤٥ **(B)** ٠,٥٢ **(C)** ٠,٤٨ **(D)** ٠,٥٦

(٢٨) ما هو الطول (بالمتر) الذي يجب أن تكون عليه قطعة قماش بحيث يكون احتمال عدم وجود أي عيوب بها يساوي ٠,٠١٨٣؟

(A) ١٨ **(B)** ١٩ **(C)** ٢٠ **(D)** ٢٢

يتبع عمر نوع معين من المصابيح الكهربائية المُنتَجة في أحد المصانع توزيعاً طبيعياً متوسطه ٥٠٠ ساعة وانحرافه المعياري ٦٠ ساعة.

أجب عن السؤاليْن التاليين (السؤال ٢٩ والسؤال ٣٠):

(٢٩) ما هو احتمال أن يعمل مصباح كهربائي ٥٦٠ ساعة على الأكثر؟

(A) ٠,٨٤١٣ **(B)** ٠,١٥٧٨ **(C)** ٠,٨٤٢٢ **(D)** ٠,١٥٨٧

(٣٠) إذا ظل مصباح يعمل إلى ما بعد ٤٤٠ ساعة، ما هو احتمال أن يزيد عمره على ٥٦٠ ساعة؟

(A) ٠,١٦٨٨ **(B)** ٠,١٨٨٦ **(C)** ٠,١٨٦٨ **(D)** ٠,١٦٨٥

(٣١) خصـــصـــت إحدى المدارس الخاصـــة اختباراً معيناً للالتحاق بها. وكانت درجات المتقدمين لها تتبع توزيعاً طبيعياً متوســـطه ٧٢ درجة وانحرافه المعياري ٥ درجات. وأنه لكي يُقبل المُتقدم في هذه المدرسة يجب أن يكون من أفضل ٢,٥٪ بين المتقدمين. ما هو الحد الأدنى للدرجة (مُقرَّباً لأقرب درجة صحيحة) الذي يجب أن يحصل عليه المُتقدم حتى يتم قبوله بتلك المدرسة؟

(A) ٨٦ **(B)** ٧٦ **(C)** ٧٨ **(D)** ٨٢

فيما يلي بيانات عن أوزان عينتيْن من الأطفال حديثي الولادة في مدينتيْن (ب و جـ):
المدينة (ب): حجم العينة = ٧٥ ، متوسط العينة = ٦,١٢ ، تباين للعينة = ٩
المدينة (جـ): حجم العينة = ١٠٠ ، متوسط العينة = ٧,٢٦ ، تباين للعينة = ١٣
فإذا كان $م_ب$ و $م_{جـ}$ هما متوسطي وزن الطفل في المدينتيْن ب و جـ على الترتيب.
أجب عن الأسئلة الثلاثة التالية (السؤال ٣٢ حتى السؤال ٣٤):

(٣٢) ما هو الحد الأعلى لفترة الثقة ٩٥٪ للفرق بين المتوسطيْن ($م_{جـ} - م_ب$)؟

(A) ٢,١٨ **(B)** ١,٩٤ **(C)** ٢,١٢ **(D)** ١,٨٨

(٣٣) يدَّعي أحد الباحثين بأن متوســـط وزن الطفل في المدينة (جـ) يفوق نظيره في المدينة (ب). وللوقوف على مدى صحة هذا الادعاء تم إجراء الاختبار الإحصائي التالي عند مستوى المعنوية ١٪:

$ف_٠$: $م_{جـ} - م_ب$ = صفر مقابل $ف_١$: $م_{جـ} - م_ب$ > صفر ،

قيمة إحصاء الاختبار في هذه الحالة هي:

(A) ١,٩٦ **(B)** ٢,٢٨ **(C)** ٢,٤٢ **(D)** ١,٨٦

(٣٤) فيما يتعلق بالسؤال (٣٣)، ما هو القرار النهائي بشأن ادعاء الباحث؟

(A) رفض فرضية العدم؛ هناك دليل كافٍ يؤيد صحة ادعاء الباحث

(B) هناك من الأدلة ما يكفي لرفض فرضـــية العدم؛ ليس هناك ما يؤيد صـــحة ادعاء الباحث

(C) عدم رفض فرضية العدم؛ هناك دليل كافٍ يؤيد صحة ادعاء الباحث

(D) ليس هناك من الأدلة ما يكفي لرفض فرضـــية العدم؛ ليس هناك ما يؤيد صـــحة ادعاء الباحث

(٣٥) ما هو حجم العينة (الحد الأدنى) اللازم إذا كنت تريد أن تكون على ثـــــقة ٩٩٪ بأن متوســـط العينة لا يختلف عن متوســـط المجتمع بأكثر من ٣ علماً بأن الانحراف المعياري للمجتمع هو σ = ٨؟

(A) ٤٨ **(B)** ٤٧ **(C)** ٤٦ **(D)** ٥٠

(٣٦) تم تقدير فترة ثـــقة للنسبة في مجتمع لتكون ٠,٤١٧٧٥ و ٠,٥٨٢٢٥ فإذا كان حجم العينة هو ١٠٠، ما هي درجة الثقة التي تم على أساسها تقدير هذه الفترة؟

<u>مع عدم إجراء أي تقريب في حساباتك.</u>

(A) ٪٩٥ **(B)** ٪٩٨ **(C)** ٪٩٠ **(D)** ٪٩٩

(٣٧) إذا كانت فترة الثـــــقة ٩٥٪ لمتوســـط درجات الطلاب هي ٢٠ و٢٥ درجة وذلك باستخدام عينة حجمها ١٤٤ طالباً. ما هي قيمة الانحراف المعياري؟

(A) ١٥,٣ **(B)** ١٥,٦ **(C)** ١٦,٥ **(D)** ١٦,٢

يرغب باحث اجتماعي في الوقوف على مدى اعتقاد الناس بوجـود مشـاهد عنف مُبالغٍ فيه في بعض الأفلام العربية. وفي عينة عشـــوائية حجمها ١٢٠ رجلاً أقَرَّ ٧٠٪ منهم بوجود هذا العنف. وفي عينة عشـــوائية أخري مُكوَّنة من ٨٠ امرأة، أفادت ٥٠٪ منهن بوجود هذا العنف. فإذا كانت $ل_١$ و $ل_٢$ هما نسبتيْ من يقرون بوجود هذا العنف في مجتمعي الرجال والنسـاء على الترتيب. **أجب عن الأسئلة <u>الثلاثة</u> التالية (السؤال ٣٨ حتى السؤال ٤٠):**

(٣٨) فترة الثقة ٩٥٪ للنسبة ($ل_٢$) هي:

(A) ٠,١ و ٠,٩ **(B)** ٠,٣ و ٠,٧ **(C)** ٠,٢ و ٠,٨ **(D)** ٠,٤ و ٠,٦

(٣٩) لمعرفة ما إذا كان هناك اختلاف بين آراء الرجال وآراء النساء فيما يتعلق بهذا الأمر، تم إجراء اختبار إحصائي، ما هي الصيغة الملائمة لكل من فرضية العدم والفرضية البديلة لإجراء مثل هذا الاختبار؟

(A) $ف_٠$: $ل_١ - ل_٢ =$ صفر مقابل $ف_١$: $ل_١ - ل_٢ >$ صفر

(B) $ف_٠$: $ل_١ - ل_٢ =$ صفر مقابل $ف_١$: $ل_١ - ل_٢ <$ صفر

(C) $ف_٠$: $ل_١ - ل_٢ =$ صفر مقابل $ف_١$: $ل_١ - ل_٢ \neq$ صفر

(D) $ف_٠$: $ل_١ - ل_٢ =$ صفر مقابل $ف_١$: $ل_١ - ل_٢ \leq$ صفر

(٤٠) فيما يتعلق بالسؤال (٣٩)، وعند مستوى معنوية ٠,٠١ ، القرار النهائي هو:

(A) قيمة إحصاء الاختبار هي ٢,١٥، وأنه ليس هناك اختلاف بين آراء الرجال والنساء

(B) قيمة إحصاء الاختبار هي ٢,٨٦، وأن هناك اختلافاً بين آراء الرجال والنساء

(C) قيمة إحصاء الاختبار هي ٢,٢٤، وأنه ليس هناك اختلاف بين آراء الرجال والنساء

(D) قيمة إحصاء الاختبار هي ٣,٢٠، وأن هناك اختلافاً بين آراء الرجال والنساء

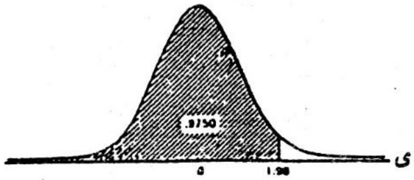
الملاحق

(الجداول الإحصائية)

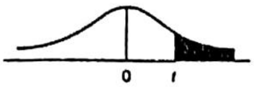
١- جدول المساحة تحت منحنى التوزيع المعتاد المعياري
٢- جدول المساحة تحت منحنى توزيع ت
٣- جدول المساحة تحت منحنى توزيع مربع كاي

الملحق (١)

المساحة تحت المنحنى المعتاد المعياري: ϕ(ي)



ى	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	ى
0.00	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359	0.00
0.10	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753	0.10
0.20	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141	0.20
0.30	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517	0.30
0.40	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879	0.40
0.50	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224	0.50
0.60	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549	0.60
0.70	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852	0.70
0.80	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133	0.80
0.90	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389	0.90
1.00	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621	1.00
1.10	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830	1.10
1.20	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015	1.20
1.30	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177	1.30
1.40	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319	1.40
1.50	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441	1.50
1.60	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545	1.60
1.70	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633	1.70
1.80	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706	1.80
1.90	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767	1.90
2.00	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817	2.00
2.10	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857	2.10
2.20	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890	2.20
2.30	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916	2.30
2.40	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936	2.40
2.50	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952	2.50
2.60	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964	2.60
2.70	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974	2.70
2.80	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981	2.80
2.90	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986	2.90
3.00	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990	3.00
3.10	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993	3.10
3.20	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995	3.20
3.30	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997	3.30
3.40	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998	3.40
3.50	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	3.50
3.60	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	3.60
3.70	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	3.70
3.80	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	3.80



الملحق (٢)

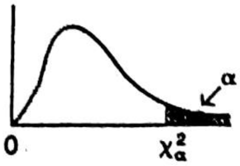
المساحة عند الذيل الأيمن تحت منحنى توزيع ت

df	Area in the Right Tail under the *t* Distribution Curve					
	.10	.05	.025	.01	.005	.001
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
31	1.309	1.696	2.040	2.453	2.744	3.375
32	1.309	1.694	2.037	2.449	2.738	3.365
33	1.308	1.692	2.035	2.445	2.733	3.356
34	1.307	1.691	2.032	2.441	2.728	3.348
35	1.306	1.690	2.030	2.438	2.724	3.340
36	1.306	1.688	2.028	2.434	2.719	3.333
37	1.305	1.687	2.026	2.431	2.715	3.326

تابع: الملحق (٢)

المساحة عند الذيل الأيمن تحت منحنى توزيع ت

	Area in the Right Tail under the *t* Distribution Curve					
df	.10	.05	.025	.01	.005	.001
38	1.304	1.686	2.024	2.429	2.712	3.319
39	1.304	1.685	2.023	2.426	2.708	3.313
40	1.303	1.684	2.021	2.423	2.704	3.307
41	1.303	1.683	2.020	2.421	2.701	3.301
42	1.302	1.682	2.018	2.418	2.698	3.296
43	1.302	1.681	2.017	2.416	2.695	3.291
44	1.301	1.680	2.015	2.414	2.692	3.286
45	1.301	1.679	2.014	2.412	2.690	3.281
46	1.300	1.679	2.013	2.410	2.687	3.277
47	1.300	1.678	2.012	2.408	2.685	3.273
48	1.299	1.677	2.011	2.407	2.682	3.269
49	1.299	1.677	2.010	2.405	2.680	3.265
50	1.299	1.676	2.009	2.403	2.678	3.261
51	1.298	1.675	2.008	2.402	2.676	3.258
52	1.298	1.675	2.007	2.400	2.674	3.255
53	1.298	1.674	2.006	2.399	2.672	3.251
54	1.297	1.674	2.005	2.397	2.670	3.248
55	1.297	1.673	2.004	2.396	2.668	3.245
56	1.297	1.673	2.003	2.395	2.667	3.242
57	1.297	1.672	2.002	2.394	2.665	3.239
58	1.296	1.672	2.002	2.392	2.663	3.237
59	1.296	1.671	2.001	2.391	2.662	3.234
60	1.296	1.671	2.000	2.390	2.660	3.232
61	1.296	1.670	2.000	2.389	2.659	3.229
62	1.295	1.670	1.999	2.388	2.657	3.227
63	1.295	1.669	1.998	2.387	2.656	3.225
64	1.295	1.669	1.998	2.386	2.655	3.223
65	1.295	1.669	1.997	2.385	2.654	3.220
66	1.295	1.668	1.997	2.384	2.652	3.218
67	1.294	1.668	1.996	2.383	2.651	3.216
68	1.294	1.668	1.995	2.382	2.650	3.214
69	1.294	1.667	1.995	2.382	2.649	3.213
70	1.294	1.667	1.994	2.381	2.648	3.211
71	1.294	1.667	1.994	2.380	2.647	3.209
72	1.293	1.666	1.993	2.379	2.646	3.207
73	1.293	1.666	1.993	2.379	2.645	3.206
74	1.293	1.666	1.993	2.378	2.644	3.204
75	1.293	1.665	1.992	2.377	2.643	3.202
∞	1.282	1.645	1.960	2.326	2.576	3.090



الملحق (٣)

المساحة عند الذيل الأيمن تحت منحنى توزيع كا٢

df	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$	df
1	.000	.000	.001	.004	3.841	5.024	6.635	7.879	1
2	.010	.020	.051	.103	5.991	7.378	9.210	10.597	2
3	.072	.115	.216	.352	7.815	9.348	11.345	12.838	3
4	.207	.297	.484	.711	9.488	11.143	13.277	14.860	4
5	.412	.554	.831	1.145	11.070	12.832	15.086	16.750	5
6	.676	.872	1.237	1.635	12.592	14.449	16.812	18.548	6
7	.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2,733	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819	13
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319	14
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582	19
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401	21
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796	22
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.558	24
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928	25
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	40.113	43.194	46.963	49.645	27
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993	28
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336	29
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672	30

المراجع

1. Anderson, R. A., Sweeney, D. J. and Williams, T. A. (2002). Statistics for Business and Economics (8^{th} edition). South - Western Thomson Learning.

2. Erichson, B. H. and Nosanchuk. T. A. (1977). Understanding Data. McGraw-Hill.

3. Lind, D. A., Marchal, W. G. and Wathen, S.A. (2008). Basic Statistics for Business and Economics (13^{th} edition). McGraw-Hill.

4. Mann, P. S. (2007). Introductory Statistics (2007). John Wiley.