محاضـــــــرات

فى

مبادئ الإقتصاد القياسى

إعداد

دكتـــــور
موافى رمضان موافى

دكتـــــورة
عبير منصور عبد الحميد

قسم الاقتصاد – كلية التجارة
جامعة جنوب الوادى

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

الْحَمْدُ لِلَّهِ الَّذِي أَنْزَلَ عَلَىٰ عَبْدِهِ الْكِتَابَ وَلَمْ يَجْعَلْ لَهُ عِوَجًا ۜ

صدق الله العظيم

(سورة الكهف : الآية ١)

بسم الله الرحمن الرحيم

مقدمة :

إن الهدف من مقرر الاقتصاد القياسى هو توضيح كيفية استخدام أدوات التحليل الاحصائى في مجال الاقتصاد. اي استخدام الوسائل والأدوات التي يستخدمها رجال الإحصاء من موضوعات الارتباط والانحدار وغيرها للمساعدة في حل المشاكل الاقتصادية التى تواجه متخذى القرار.

ولذا، فسوف نحاول في هذه المادة التعرف على ماهية الاقتصاد القياسى والتوصل إلى أن دراسة هذه المادة هى محطة هامة , ونقطة تحول لفهم أعمق وأوسع لما درسه الطالب من مقاييس لآداء الاقتصاد.

وبعد , فإننا نرجوا من الله أن نكون قد وفقنا فى إعداد هذه المذكرة آملين أن ينتفع بها أبناءنا الطلاب , وأن بستزيدوا بها علماً ومعرفةً .

مع أطيب التمنيات

فهرست الموضوعات

الفصل الأول

طبيعة علم الاقتصاد القياسي

الهدف من دراسة الفصل

بنهاية الدراسة لهذا الفصل يكون الدارس قادر على الالمام بالمفاهيم التالية:

١- تعريف الاقتصاد القياسي واهدافه

٢- تعريف النموذج الاقتصادي

٣- مكونات النموذج ومراحل اعداده

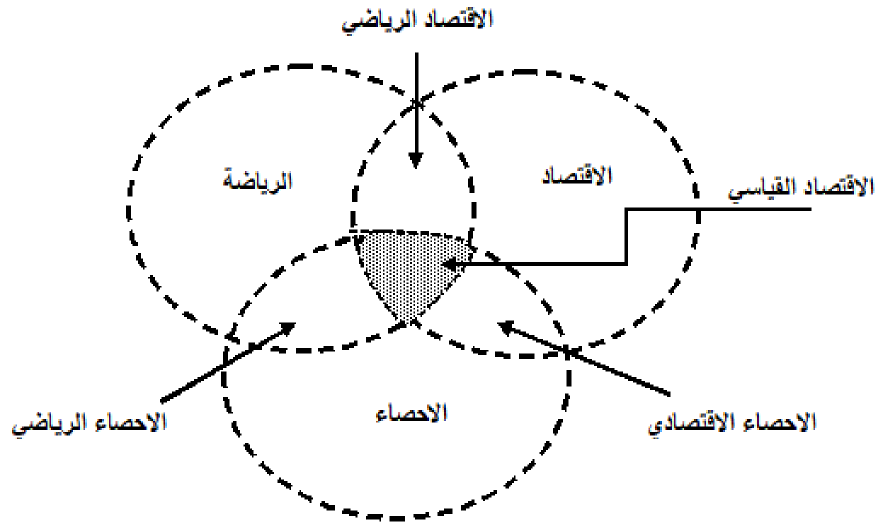
الفصل الأول

طبيعة علم الاقتصاد القياسى

- **تعريف الاقتصاد القياسي: Definition of Econometric**

كلمة إقتصاد قياسي بالإنجليزية (Econometrics) : مكونة من مقطعين : ECONOمشتقة من إقتصاد و METRICS مشتقة من كلمة قياس.

والاقتصاد القياسي Econometrics فرع من فروع علم الاقتصاد الذي يختص بالقياس (التقدير) الكمي للعلاقة بين المتغيرات مستخدما النظرية الاقتصادية والرياضيات والأساليب الاحصائية ، بهدف إختبار النظريات الاقتصادية المختلفة من ناحية ومساعدة رجال الأعمال والحكومات في إتخاذ القرارات ووضع السياسات من ناحية أخري .



أي أن الاقتصاد القياسي يهتم بتحليل الظواهر الاقتصادية الواقعية تحليلاً كمياً ، وذلك باستخدام أساليب الاستقراء الإحصائي المناسبة. أي إنه علم استعمال طرائق الاستقراء والاستدلال الإحصائي لكشف القوانين الاقتصادية الموضوعية وتحديد فعلها تحديداً كمياً.

فالتحليل الكمي للظواهر الاقتصادية هو محاولة للتحقق من العلاقات الاقتصادية والتأكد من منطقيتها في تمثيل الواقع المعقد الذي تعبر عنه النظرية الاقتصادية في صيغة فروض . ويعتمد الاقتصاد القياسي في قياس العلاقات الاقتصادية وتحليلها على دمج النظرية الاقتصادية والرياضيات والأساليب الإحصائية في نموذج متكامل ، وذلك بهدف تقويم معالم ذلك النموذج ثم إختبار الفروض حول ظاهرة إقتصادية معينة ، وأخيراً التنبؤ بقيم تلك الظاهرة.

- **علاقة الاقتصاد القياسي بالعلوم الاخرى:**

من الواضح أن علم الاقتصاد القياسي يعتمد على ثلاثة علوم هي:

١. علم الاقتصاد: وهذا أمر طبيعي ، إذ إن الاقتصاد القياسي هو أحد فروع هذا العلم. فالنظرية الاقتصادية تشير عموماً إلى وجود علاقات معينة بين متغيرات إقتصادية كالعلاقة بين الكمية المطلوبة من سلعة معينة وسعرها وأسعار السلع البديلة مثلاً ، وتحتاج عملية قياس تلك العلاقات إلى إختيار نماذج قياسية لتمثيلها.

٢. الرياضيات بما توفره من نماذج رياضية يختار الاقتصاد القياسي ما يناسب منها وفق أسس معينة للوصول إلى نموذج لتمثيل العلاقات بين المتغيرات الاقتصادية المدروسة . ومن الطبيعي أن يكون بعض تلك النماذج أقل جودة في التعبير عن الواقع المعقد من بعضها الآخر.

٣. الإحصاء بما يوفره من أدوات أساسية في القياس كالتي تتعلق بطرائق الاستدلال الإحصائي مثلاً.

إن علم الاقتصاد القياسي وفقاً لتعريف عدد من الأعلام الرواد في هذا المجال كلورنس كلاين L.Klein وإدموند مالينفو E.Malinvaud ، هو علم إستعمال طرائق الاستقراء والاستدلال الإحصائيين ، ولاسيما نظريات الاحتمال والتنبؤ والتقدير.

- **تاريخ الاقتصاد القياسي**

يعدّ علم الاقتصاد القياسي علماً حديثاً نسبياً إذا ما قورن بالعلوم الاقتصادية الأخرى ، فعلى الرغم من المحاولات التي ظهرت في القرن التاسع عشر والتي كانت ذات طابع إقتصادي قياسي ، كعمل الإحصائي الألماني أرنست إنغل (١٨٢١-١٨٩٦) Ernest Engel الذي وضع قوانينه الخاصة بالدخل والاستهلاك في ضوء بيانات ميزانية الأسرة ، وإستعمل مصطلح الاقتصاد القياسي أول مرة عام ١٩٢٦ من قبل الاقتصادي النروجي فريش Frisch.

في عام ١٩١٩ نشر الاقتصادي الأمريكي بيرسون W.M.Pearson طريقته الخاصة بتحليل الدورات الاقتصادية التي طبقت في تحليل هذه الدورات في عدد من البلدان الرأسمالية ، كما طبقت في الاتحاد السوفييتي سابقاً أيضاً في إنجاز عدد من الأبحاث التي وضعت في خدمة سياسة الدولة السوفييتية في مرحلة الانتقال من الرأسمالية إلى الاشتراكية. وتعد محاولات تقدير دوال منحنيات العرض والطلب للمنتجات الزراعية في الولايات المتحدة الأمريكية في مطلع الثلاثينات من القرن العشرين محاولات أولى أيضاً في مجال تطبيق مبادئ الاقتصاد القياسي.

أسس بعض واضعي الفكر الاقتصادي الأوائل من أمثال مور H.More ، وشولتز H.Schultz ، وفريش وستون R.Stone الجمعية الدولية للاقتصاد القياسي International Econometrics Association في عام ١٩٣٠. ثم توسع تطبيق مبادئ الاقتصاد القياسي بعد الحرب العالمية الثانية ، وأخذت أنشطة هذا العلم تشمل تقديرات لمعالم أو لثوابت نماذج اقتصادية مؤلفة من عدة معادلات. ومنذ ذلك التاريخ والاقتصاد القياسي يستخدم أداة فعالة في حل المعضلات الاقتصادية وفي عمليات التخطيط الاقتصادي . وبدأ تطبيق مبادئ هذا العلم بالانتشار حديثاً في بلدان العالم الثالث. وساعد على إنتشار طرائق الاقتصاد القياسي عاملان إثنان هما:

١. توافر الإحصاءات الاقتصادية بكميات أكبر وبدقة أفضل. وهي تؤلف المادة الأولية للبحث العلمي في الاقتصاد القياسي.

٢. التطور الكبير والسريع في مجال الحاسبات الإلكترونية الذي مكن من التوسع في النماذج الاقتصادية لتشمل عدداً كبيراً من المتغيرات بعد أن كان ذلك مقتصراً على التحليل النظري. فقد أصبح بالإمكان اليوم تقدير ثوابت نموذج مؤلف من عدة مئات من المعادلات وإختبار صلاحية النماذج الاقتصادية النظرية ومعرفة مدى ملاءمتها للواقع المعقد.

- **أهداف الاقتصاد القياسي: The Goals of Econometrics**

يهدف الاقتصاد القياسي إلي تحقيق ثلاثة أهداف رئيسية هي علي النحو التالي:

١. إختبار النظريات الاقتصادية المختلفة .

٢. مساعدة رجال الاعمال والحكومات في إتخاذ القرارات.

٣. مساعدة رجال الأعمال والحكومات في وضع السياسات.

- **مهام الاقتصاد القياسي**

تتمثل مهام الاقتصاد القياسي عامة بتحقيق ما يلي:

١. تحديد النموذج الرياضي المناسب لتمثيل العلاقة أو العلاقات القائمة بين المتغيرات الاقتصادية المدروسة ، إذ يجب على الباحث في هذه المرحلة وضع فروض النظرية الاقتصادية في نموذج رياضي عشوائي.

٢. تقدير معاملات أو ثوابت النموذج الرياضي المطبق. تبدأ هذه المهمة بجمع الإحصاءات الاقتصادية المناسبة بالدقة المطلوبة حول ظاهرة أو ظواهر يراد دراستها وتنتهي باستخدام الأساليب الإحصائية المناسبة لتقدير معالم النموذج الذي إختاره الباحث لتمثيل العلاقات بين المتغيرات.

٣. إختبار النموذج الرياضي العشوائي المطبق لمعرفة ما إذا كان يمثل فعلاً حقيقة الواقع المدروس أم أنه يجب على الباحث اختيار نموذج آخر أكثر واقعية. ويعتمد الباحث في اختيار النموذج المناسب على معايير اقتصادية، إذ من المفترض أن تنسجم قيم المعاملات المقررة في النموذج في طبيعتها وقيمها النسبية مع ما هو متوقع في إطار النظرية والفروض الاقتصادية التي تحكم الظواهر المدروسة. وكذلك من اختبارات فروض النموذج نفسها، ولاسيما تلك المتصلة بالحد العشوائي لمعرفة مدى انسجامها مع الواقع المدروس.

- **تطبيقات الاقتصاد القياسي:**

يعتبر مجال تطبيق الاقتصاد القياسي واسعًا جدًا حيث يشمل كافة الظواهر الاقتصادية:

- على مستوى الاقتصاد الجزئي: حيث يمكن إستخدام تطبيقاته لتحديد دوال الانتاج والتكاليف على مستوى المنشاة وكافة إشتقاقاتها مثل دوال الناتج المتوسط والناتج الحدي والتكلفة المتوسطة والحدية. وكذلك يقيس تاثير العوامل المؤثرة على الانتاج كميًا، ويحدد الحدود المثلى من كل عامل التي يجب إدخالها في العملية الانتاجية ، ويحدد التوليفة المثلى من العوامل مجتمعة التي تحقق أفضل عائدية.
- على مستوى الاقتصاد الكلي: يمكن باستخدام النماذج القياسية تقدير دوال الاستهلاك والطلب للسلع المختلفة على المستوى الكلي. وكذلك دوال الانتاج بصيغها غير الخطية المختلفة . كما يمكن بناء نماذج قياسية (متعددة المعادلات) توصف الاقتصاد ككل وتتضمن دوال الدخل القومي والاستثمار والاستخدام والاستهلاك والتجارة الخارجية (الصادرات والواردات).
- ويمكن إستخدام تطبيقات الاقتصاد القياسي في بعض الدراسات الاجتماعية.

- **إستخدام الاقتصاد القياسي:**

تطور إستعمال الاقتصاد القياسي مع تطور العلم نفسه ومع تغير المشكلات الاقتصادية. وبوجه عام فإن مجالات تطبيق طرق الاقتصاد القياسي هي:

١. تحليل الدورات الاقتصادية التي تعرضت لها البلدان الرأسمالية ، وخاصة الولايات المتحدة في مطلع القرن العشرين، بهدف التنبؤ بمواعيدها والتصدي للأزمات الاقتصادية ومعالجتها أو التخفيف من حدتها قبل حدوثها وتقليص الخسائر الناجمة عنها. وكانت جامعة هارفرد المركز الأول لهذا النوع من الأبحاث التي قلت أهميتها إثر عجزها عن التنبؤ بحدوث الأزمة الاقتصادية الكبرى عام ١٩٢٩.

٢. أبحاث السوق وتحديد مرونة الطلب والعرض ، إذ من الثابت عموماً أنّ هناك علاقة عكسية بين سعر المنتج والكمية المطلوبة منه. ومن المهم عند المنتجين معرفة مدى أثر تغيير محدد في سعر السلعة في الكمية المطلوبة منها. وعلى صعيد أجهزة الدولة المسؤولة عن تخطيط عملية التنمية فإن هذا النوع من الأبحاث ذو أهمية خاصة، إذ إن السياسات السعرية تؤلف أدوات لتوجيه أنماط الإنتاج والاستهلاك باتجاهات مرغوب فيها، مما يحتم ضرورة تعرُّف فعالية هذه الأدوات قبل استعمالها. ففي المجتمعات الاشتراكية مثلاً، يتطلب التخطيط الفعال للاستهلاك الفردي تعرُّف مرونات الطلب بالنسبة إلى الدخل والأسعار، لكي يستطيع المخطط تعرُّف الطلب المستقبلي في ضوء التطور المرسوم للدخول والأسعار المتوقعة للسلع وبدائلها.

٣. دراسة مستويات الإنتاج وعلاقتها بالتكلفة ، وهي دراسات ذات أهمية في مسائل تخطيط الإنتاج على صعيد الوحدات والقطاعات الإنتاجية. إذ تبين هذه الدراسات الأهمية النسبية لكل عامل من عوامل الإنتاج في العملية الإنتاجية على صعيد المؤسسة وأهميته في النمو الاقتصادي على مستوى القطاع والمجتمع. أي تحديد مصادر النمو الاقتصادي في المجتمع ودور التطور التقني في ذلك.

٤. نظرية البرمجة التي تطبق تطبيقاً واسعاً على صعيد الوحدات الإنتاجية في البلدان الرأسمالية والاشتراكية وفي تخطيط الاقتصاد الاشتراكي الشامل. وفي إطار هذه النظرية يتم تحليل النشاطات الاقتصادية المتداخلة بهدف ضمان التوازن بين جميع الوحدات المستقلة المساهمة في العمليات الإنتاجية المترابطة.

- **الاقتصاد القياسي والنماذج الرياضية:**

النموذج الاقتصادي هو تبسيط رياضي لحالة واقعية معقدة في المجتمع يفترض أن يعكس حقيقة العلاقات القائمة بين المتغيرات الاقتصادية الداخلة فيه. ويتوقف عدد هذه العلاقات على الأهداف المتوخاة من النموذج وعلى درجة التفصيل المرغوب في الحصول عليها. وتشترك النماذج الاقتصادية عامة بخصائص معينة منها:

أ. الافتراض أن سلوك المتغيرات الاقتصادية يتحدد بوساطة مجموعة معادلات تعرف بالمعادلات المتزامنة simultaneous equations.

ب. الافتراض أن النموذج المقترح تطبيقه يؤلف أكثر من مجرد تبسيط رياضي لحالة معقدة في الواقع.

ج. إفتراض أن يساعد فهم النموذج المطبق على فهم سلوك متغيرات النموذج في المستقبل. بمعنى أنه يساعد على إجراء التنبؤات المستقبلية حول مستويات تلك المتغيرات.

وتقسم النماذج الاقتصادية الرياضية إلى:

١. النموذج الخطي البسيط:

يعدّ النموذج الخطي البسيط أبسط أشكال النماذج الرياضية، فهو يتضمن متغيرين فقط أحدهما متغير تفسيري ويرمز له عادة بالرمز X، والثاني متغير تابع ويرمز له بالرمز Y. كما في النموذج ذي الرقم (١):

$$Y_i=A+BX_i+U_i \quad (1)$$

إذ إن (i) وهو الجنيب، يعبر عن رقم المشاهدة في المجتمع ($i=1,2,3,...N$) أو في العينة ($i=1,2,3,...n$)، وإن N و n تمثلان عدد وحدات المجتمع أو العينة على التوالي في الظاهرة المدروسة.

في هذا النموذج الخطي البسيط يمكن الافتراض، مثلاً، أن X_i تمثل الدخل التصرفي للأسرة (i) في حين تمثل Y_i الإنفاق الاستهلاكي الإجمالي لهذه الأسرة. أما A و B فهما معلمان أو ثابتان يمثل الأول متوسط مستوى الإنفاق الاستهلاكي عندما يكون الدخل التصرفي صفراً، ويمثل الثاني متوسط مقدار التأثير في Y عندما تتغير X بمقدار وحدة واحدة.

وأخيراً يعرف U_i بحد الخطأ أو المتغير العشوائي الذي يأخذ قيمة موجبة لدى أسرة تنفق أكثر من متوسط إنفاق الأسر المماثلة لها في الدخل وقيمة سالبة عند إنفاقها أقل من ذلك المتوسط وقيمة الصفر إذا ساوى إنفاقها متوسط إنفاق الأسر المماثلة لها في مستوى الدخل. وتبقى القيمة المتوقعة لهذا المتغير العشوائي ويرمز لها بالرمز $E(U_i)$ مساوية الصفر دائماً.

إن إدخال المتغير العشوائي U_i في النموذج الاقتصادي له عدة مسوغات أهمها:

أ. هناك الكثير من المتغيرات التي تؤثر في إنفاق الأسرة الاستهلاكي إلى جانب الدخل التصرفي في مثالنا هذا. وقد يتعذر قياس هذه المتغيرات أو ربما يحتاج ذلك إلى الكثير من الجهد والوقت والمال. فعلى سبيل المثال، إن حجم الأسرة ومكان إقامتها (مدينة أو قرية) وتركيبها النوعي وحساب أعمار أفرادها ومستواهم الثقافي، وغير ذلك كلها عوامل تؤثر في مستوى إنفاقها الاستهلاكي إلى جانب الدخل التصرفي. وقد يكون تأثير هذه المتغيرات المحذوفة في المتغير التابع موجباً أو سالباً إلا أنها في المحصلة تأثيرات يفترض أنها ثانوية يعكسها حد الخطأ.

ب. من المتعذر التنبؤ بدقة باستجابة الأفراد للتغيرات التي تطرأ على دخولهم. فإذا تضاعف دخل الأسرة مثلاً فإن التنبؤ بتغير مستوى إنفاقها الاستهلاكي وتركيبه بدقة أمر في غاية الصعوبة. ثم إن حد الخطأ يفترض فيه أن يعكس أخطاء التنبؤ هذه.

ج. أخطاء قياس متغيرات العلاقة الحقيقية في المجتمع. إذ لابد من ارتكاب أخطاء معينة في قياس قيم المتغيرات الاقتصادية في المسوح الإحصائية الميدانية. وتظهر تأثيرات أخطاء القياس هذه في المتغير العشوائي أيضاً.

ومع ذلك فإن إدخال المتغير العشوائي U_i في النموذج الاقتصادي يقتضي وضع بعض الافتراضات التي تتعلق بوسطه الحسابي (أو قيمته المتوقعة) وتباينه وتغاير قيمه المختلفة فيما بينها وتغاير قيمه المختلفة مع قيم المتغير (أو المتغيرات) التفسيري في النموذج.

٢. النموذج الخطي المتعدد المتغيرات التفسيرية:

إن الحالة التي هي أكثر شيوعاً في الاقتصاد أن يكون المتغير التابع Y تابعاً لعدد من المتغيرات التفسيرية لا لمتغير واحد. وهذه هي حال العلاقة ذات الرقم (٢) التي يطلق عليها علاقة الانحدار الخطي المتعدد.

$$Y_i=A+BX_i+CZ_i+U_i \quad (2)$$

إن Y_i في هذه العلاقة التي يفترض أنها تمثل كما في السابق الإنفاق الاستهلاكي الإجمالي للأسرة il ، تابعة ليس فقط للدخل التصرفي X_i لهذه الأسرة وإنما لمتغير آخر Z_i وهو عدد أفراد هذه الأسرة مثلاً. وقد يزيد عدد المتغيرات التفسيرية في النموذج على إثنين بحسب الظاهرة المدروسة وعلاقتها بالظواهر الأخرى.

إن الميزة الأساسية لعلاقات الانحدار الخطي المتعددة هي أنها تسمح بأن يعزل على حدة تأثير كل متغير تفسيري في النموذج. فعلى سبيل الاستئناس، تمثل B في النموذج ذي الرقم (٢) متوسط مقدار التأثير في Y عندما تتغير X بمقدار وحدة واحدة مع بقاء المتغير Z على مستواه. كذلك تمثل C متوسط مقدار التأثير فيY عندما تتغير Z بمقدار وحدة واحدة مع بقاء X على مستواه . وقد يكون التأثير موجباً أو سالباً بحسب طبيعة العلاقة بين المتغير التابع وكل من المتغيرات التفسيرية . وتدل الإشارة الموجبة (+) على العلاقة الطردية بين المتغير التابع والمتغير التفسيري (المستقل) في حين تدل الإشارة السالبة (−) على العلاقة العكسية بينهما، أي إن الإشارة تبين إتجاه التأثير.

٣. النماذج الرياضية غير الخطية:

تتعدد الصيغ غير الخطية في الاقتصاد القياسي، ويمكن دوماً ابتداع صيغ جديدة. وفيما يلي أمثلة قليلة على بعض الصيغ غير الخطية.

$$Y_i = A + BX_i^2 + U_i \qquad (3)$$

$$Y_i^2 = C + D\left(\frac{1}{X}\right) U_i \qquad (4)$$

$$Y_i = FX_i^M U_i \qquad (5)$$

إذ إن:A، B، C، D و F هي ثوابت تقدر قيمتها في النموذج المعني. وتشير هذه الصيغ إلى وجود علاقة غير خطية بين Y والمتغير التفسيري X في الصيغ الثلاث. ومع ذلك يلاحظ أن إعادة تعريف المتغير X^2 في النموذج ذي الرقم (٣)، كأن نضع X^2=W، يحول العلاقة الأصلية غير الخطية إلى علاقة خطية:

$$Y=A+BW_i+U_i$$

وإن إستعمال التحويلة الرياضية اللوغاريتمية يحول العلاقة ذات الرقم (٥) إلى علاقة خطية أيضاً:

$$\text{Log } Y_i=\text{Log } F+m \text{ Log } X+ \text{ Log } U_i$$

أما الأسس التي يتم فيها إختيار صيغة غير خطية من دون أخرى فأهمها:

(أ) انسجام الصيغة الرياضية مع النظرية الاقتصادية المتعلقة بالظاهرة المدروسة. وغالباً ما تساعد هذه النظرية في اختيار المتغيرات التي تدخل في العلاقة، كما تساعد في تحديد تأثير كل متغير تفسيري في التابع على حدة.

(ب) مراعاة العلاقة التي تعكسها المشاهدات الإحصائية حول الظاهرة أو الظواهر المدروسة، إذ قد ترجح هذه العلاقة صيغة من دون غيرها بين الصيغ المقبولة نظرياً.

(ج) البساطة التي تتجلى في اختيار أبسط الصيغ الرياضية بين الصيغ المقبولة. لعدم اختيار معادلة من الدرجة الثانية إذا كانت معادلة من الدرجة الأولى تفي بالغرض، وعدم اختيار معادلة من الدرجة الثالثة إذا كانت المعادلة من الدرجة الثانية مناسبة.

٤ . نموذج المعادلات المتزامنة simultaneous equation system

بغية توضيح مفهوم هذا النوع من النماذج الاقتصادية الرياضية نقتبس المثال التقليدي في التحليل الاقتصادي الكلي macro-analysis.

(أ) إن الكمية المعروضة من سلعة ما ويرمز لها عادة بالرمز Q_s هي الكمية التي يقبل المنتجون بإنتاجها وبيعها من أجل مستوى معين من الأسعار P مثلاً. أي إن Q_s تابع للسعر P، ويعبر عن ذلك رياضياً بالعلاقة:

$$Q_s = F(P) \qquad (6)$$

(ب) إن الكمية المطلوبة من هذه السلعة ويرمز لها بالرمز Q_d هي الكمية التي يقبل المستهلكون شراءها من أجل السعر P. أي إن Q_d تابع للسعرP، ويعبر عن ذلك رياضياً بالعلاقة:

$$Q_d = G(P) \qquad (7)$$

ويلاحظ في العلاقتين (٦) و(٧) أن المتغير التفسيري هو P ، ومع ذلك فإن كون كل من المقدارين Q_s وQ_d تابعاً لـ P لايعني بالضرورة تشابه شكل علاقة التبعية رياضياً.

(ج) يستدعي استقرار السعر في السوق تساوي الكميتين المعروضة والمطلوبة من هذه السلعة ، أي يجب تحقق العلاقة:

$$Q_d = Qs = Q_0 \qquad (8)$$

إذ تمثل Q_0 مستوى التوازن بين العرض والطلب من أجل السعر P. تؤلف هذه المعادلات الثلاث (٦) و(٧) و(٨) ما يسمى بنموذج المعادلات المتزامنة، فهو نموذج أكثر واقعية في التعبير عن العلاقات الاقتصادية القائمة في المجتمع قياساً بنماذج المعادلة الواحدة. ومع ذلك فإن صعوبة تقدير ثوابت هذا النوع من النماذج الرياضية يجعلها أقل جاذبية واستعمالاً من غيرها.

الفصل الثانى

أقسام التحليل الاقتصادى

الهدف من دراسة الفصل

بنهاية الدراسة لهذا الفصل يكون الدارس قادر على الالمام بالمفاهيم التالية:

١ـ أقسام التحليل الإقتصادي

٢ـ تعريف النموذج الاقتصادي

٣ـ مكونات النموذج ومراحل اعداده

الفصل الثاني

أقسام التحليل الإقتصادي

ينقسم التحليل الإقتصادي حسب درجة شموله للمتغيرات التي تؤثر في الظاهرة موضع الدراسة إلى مايلي:

تحليل جزئي Partial Analysis :

تحليل عام General Analysis :

فإذا كان المتغيرموضع الإهتمام هو الكمية المطلوبة من سلعة ما فان هذا المتغير يتأثر بعدد كبير من المتغيرات المستقلة مثل سعر السلعة نفسها، أسعار السلع الأخرى، أذواق المستهلكين، مستوى الدخل وهكذا.

$$y = P + Ps + T + I$$

حيث:

P= سعر السلعة نفسها

Ps= أسعار السلع الأخرى

T= أذواق المستهلكين

I= مستوى الدخل

فإذا ماتم دراسة العلاقة بين الكمية المطلوبة y من السلعة و سعرها P مع ثبات المتغيرات الأخرى فإن التحليل في هذه الحالة يطلق عليه تحليل جزئي، في حين أن دراسة العلاقة بين الكمية المطلوبة من السلعة y وكافة المتغيرات التي سبق ذكرها وذلك في وقت واحد (آنياً) يطلق عليه التحليل العام.

و ينسب التحليل العام إلى مؤسس مدرسة لوزان الإقتصادية العالم الإقتصادي ليون فالراس ***Leon Walras*** و هو فرنسي.

أدوات التحليل الإقتصادي

يمكن تقسيم التحليل الإقتصادي حسب الأدوات المستخدمة فيه أو نوع الصياغة المستخدمة إلى مايلي:

(أ) التحليل الوصفي Descriptive Analysis:

حيث يتم تحليل الظواهر الإقتصادية بطريقة وصفية كلامية دون القياس الكمي للعلاقات.ويناسب هذا النوع من التحليل الحالات التي تكون فيها العلاقات الإقتصادية بسيطة غير معقدة، كما يفيد هذا الأسلوب التحليلي في تحليل العلاقات التي يصعب صياغتها في صورة كمية.

(ب) الصيغة الرياضية Mathematical:

حيث تستخدم الأدوات الرياضية في صياغة و عرض النظريات الإقتصادية. و مع تزايد إستخدام الرياضة في الدراسات الإقتصادية ظهر فرع جديد هو "الإقتصاد الرياضي" *Mathematical Economics* وإذا كان المنهج الرياضي له فوائده و مميزاته التي يتفوق بها على الصياغة اللفظية، فان هناك العديد من المحاذير حول هذا الأسلوب مثل عدم صحة المعطيات الإحصائية والإحصاءات المستخدمة، وعدم التطابق بين المعطيات الإحصائية ووقائع موضوع الدراسة (الإلمام بكافة المتغيرات).

(ج) الصياغة القياسية Econometric:

حيث تستخدم الأدوات الرياضية والإحصائية معاً في صياغة النظريات الإقتصادية، ومع التطبيق المتزايد لهذا الأسلوب خاصة مع إستخدام الحاسبات الآلية ظهر فرع جديد في الدراسات الإقتصادية هو الإقتصاد القياسي *Econometrics*.

التحليل الإقتصادي ومتغير الزمن

يمكن تقسيم التحليل الإقتصادي وفقا لإدخال متغير عنصر الزمن في الإعتبار من عدمه إلى مايلي:

(أ) تحليل ستاتيكي (ساكن):

وهذا النوع من التحليل يتجاهل عنصر الزمن كليةً. فعند دراسة التوازن بين عرض السلعة والطلب عليها دون إشارة إلى عنصر الزمن (سعر أي فترة هو الذي يؤثر) أو الفترة يقال أن التحليل إستاتيكي "يبحث في نقطة توازن واحدة".

ولإدخال بعض الواقعية على هذا النوع من التحليل قام الإقتصاديون بدراسة الأثر النهائي لتغير أحد العوامل المستقلة على وضع التوازن الأصلي وذلك في صورة مقارنة بين وضع التوازن الجديد ووضع التوازن الأصلي دون أخذ الزمن في الإعتبار، ويعرف هذا النوع بإسم التحليل الساكن المقارن وإرتبط ظهور هذا التحليل بالنظرية العامة لكينز.

(ب) التحليل الديناميكي:

هذا النوع من التحليل يأخذ عنصر الزمن صراحة في الإعتبار ويستخدم لتوضيح مسار التغير والانتقال من وضع توازني إلى وضع توازني آخر أو إلى أوضاع غير توازنية.

ويمكن إدخال الزمن في التحليل الإقتصادي بطرق مختلفة منها:

١- *تحليل الفترات* :

حيث يكون متغير الزمن "متغيراً أو ثابتاً" ويتم تحديد العلاقات بين المتغيرات في هذا التحليل عن طريق حل مجموعة من معادلات الفروق.

٢- *تحليل العمليات:*

حيث يتم إدخال الزمن في صورة مستمرة ويتم تحديد العلاقة بين المتغيرات المختلفة عن طريق حل مجموعة من المعادلات التفاضلية.

النموذج الإقتصادي

يستعمل الإقتصاديون في تفسير وتحليل الظواهر الإقتصادية مايعرف بالنموذج، وهو عبارة عن تجسيد وتقريب للواقع، بمعنى أن يلجأ الباحث إلى تبسيط الظاهرة قيد البحث والدراسة في شكل يمكن دراسته و تحليل أسسه.

فالنموذج يعطي للإقتصادي طريقة لعرض النظرية بصورة سهلة الفهم والتحليل.

ويتكون النموذج من عدة عناصر وصيغ رياضية. فعناصره الأساسية تعرف بالمتغيرات Variables وهي رموز تأخذ قيماً مختلفة.

والمتغيرات نوعان:

متغيرات مستقلة Independent Variables

متعيرات تابعة Dependent Variables

فالزيادة أو النقص في المتغيرات المستقلة يؤدي إلى زيادة أو نقص في المتغيرات التابعة، والعلاقة التي تربطهما تعرف بصيغة النموذج. وهي في العادة صيغ رياضية إما أن تكون بشكل صريح *Explicit* أو بشكل ضمني. والأمثلة على ذلك كثيرة مثلك

من النماذج البسيطة دراسة مستويات الإستهلاك كدالة في مستوى الدخل، فهذا النموذج يبسط الواقع بصورة تجعل الإستهلاك وهو متغير تابع يتوقف أو يعتمد على متغير مستقل واحد هو الدخل. و الصيغة لهذا النموذج يمكن أن تكون :

الإستهلاك = دالة (الدخل)

$$C=f(I)$$

بدون تحديد الكيفية، في حين أن الصيغة الصريحة لهذا النموذج تبين الطريقة التي يعتمد الإستهلاك على الدخل عن طريق تحديد شكل العلاقة الرياضية.

ومن النماذج المعقدة والأكثر واقعية أن نجعل الإستهلاك، باعتباره متغيراً تابعاً يعتمد على العديد من المتغيرات المستقلة و التي تؤثر فعلاً في مستويات الإستهلاك مثل الدخل، عدد السكان، مستويات الأسعار و الأذواق وغيرها . وقد تكون صيغة النموذج ضمنية أو صريحة.

ويتعلق بالصيغة الرياضية كذلك كون العلاقة بين المتغير التابع والمتغير أو المتغيرات المستقلة علاقة خطية *Linear* أو غير خطية *Non-Linear* و سوف نبين المقصود منهما عند الحديث عن الأشكال و الرسوم البيانية.

الإفتراضات الأساسية

الإقتصاد يتعلق بدراسة السلوك البشري فيما يخص الإنفاق والإنتاج وإستعمالات الموارد ضمن إطار عالم متجدد متغير ومتشابك الأطراف، لذلك فإن محاولة معرفة كل جوانب السلوك البشري أو كل جوانب الظاهرة قيد الدراسة تعتبر محاولة مستحيلة، إذ من الصعب التنبؤ بالمتغيرات التي تتحكم في سلوك الأفراد، وكذلك من الصعب السيطرة على كل الظروف التي تؤثر في هذا السلوك.

وبناءاً على ذلك يلجأ الإقتصادي إلى الإعتماد على بعض الإفتراضات التي يمكن حصرها في الآتي:

١- إفتراض بقاء العوامل الأخرى على حالها **Ceteris Paribus** :

نتيجة لتعقد الواقع وصعوبة الإلمام بالجوانب المتعددة لأي ظاهرة في آن واحد، يستخدم الإقتصادي أثناء تحليله إفتراض معين يساعد على عزل الظاهرة قيد الدراسة لغرض معرفة العلاقة بين بعض المتغيرات فيها. ويحد هذا الإفتراض من إطار النظرية عن طريق تثبيت عوامل معينة تعتبر جزء من النموذج. فيستخدم الإقتصادي إفتراض ***Ceteris Paribus*** وهي كلمة لاتينية تعني بقاء العوامل الأخرى على حالها، أو بقائها ثابته ***Other Things (or Factors)*** **Remain Constant** و كمثال على ذلك، لو أردنا دراسة أثر تغير سعر سلعة ما على الكمية المطلوبة من تلك السلعة فإننا نلجاء إلى عزل آثار العوامل الأخرى التي تؤثر في الطلب على تلك السلعة مثل الدخل، أسعار السلع الأخرى، الأذواق وعدد السكان ... ألخ. بإفتراض أنها ثابتة. وبهذه الطريقة نستطيع معرفة تأثير تغير سعر السلعة فقط على الكميات المطلوبة بمعزل عن تأثير العوامل الأخرى.

٢- إفتراض العقلانية **Rationality Assumption**:

يعني أن الشخص في تصرفاته وقراراته منسجم مع تحقيق هدف معين. وهنا لا نبحث في طبيعة الهدف في حد ذاته من حيث كونه هدفاً سامياً أو أخلاقياً أو غير ذلك، ولكننا نقول أن الشخص يتصرف بعقلانة أو أن سلوكه عقلاني ***Rational Behavior*** إذا هو حدد هدفه و نهج النهج السليم للوصول إليه. أما إذا كان سلوكه غير متفق مع الهدف الذي حدده فإننا نقول أن هذا السلوك غير عقلاني ***Irrational Behavior*** . ويمكن التفريق بين السلوك غير العقلاني والسلوك العشوائي ***Random Behavior*** فالأخير يعني أن الشخص غير مستقر يتخبط بين الأهداف والوسائل ولايحقق هدف ولايصل لغاية.

٣- إفتراض تعظيم شئ ما **Maximization Assumption** :

يتعلق هذا الإفتراض ببيان أن هدف الشخص الذي يتصرف بعقلانية، حسب المفهوم السابق، هو تعظيم شئ ما. فقد يكون هدفه تعظيم المنفعة ***Utility*** كما في حالة المستهلك ، أو تعظيم الأرباح ***Profits*** أو المبيعات ***Sales*** كما في حالة المنتج، أو تعظيم الرفاهية ***Welfare*** بإعتباره هدف إجتماعي. ويفترض أن القرار الإقتصادي قد أتخذ لغرض تحقيق أحد هذه الأهداف.

التحليل الحدي

Marginal Analysis

تعرف الوحدات الحدية أو الهامشية بأنها الوحدات الأخيرة التي تمت إضافتها مثل الوحدة الأخيرة من السلعة المستهلكة، والوحدة الأخيرة من العنصر الإنتاجي المستخدم، والوحدة الأخيرة من السلعة المنتجة. وهكذا فكلمة حدي ***Marginal*** تعني إضافي.

وكل القرارات الإقتصادية هي قرارات حدية يستعمل فيها التحليل الحدي. فعندما يقرر المنتج إنتاج وحدة إضافية من سلعة ما فإنه ينظر إلى تكلفة الوحدة الأخيرة، أي مقدار ما تضيفة هذه الوحدة الأخيرة إلى التكلفة الكلية أو الإجمالية، وهو مايعرف بإسم التكلفة الحدية ***Marginal Cost*** . وينظر كذلك إلى فائدة الوحدة الأخيرة، أي مقدار ما تضيفة هذه الوحدة الأخيرة المنتجة إلى الإيراد الكلي أو الإجمالي، و هو ما يعرف بالإيراد الحدي ***Marginal Revenue*** . وهكذا بالنسبة لكل القرارات. فكل قراراتنا الإقتصادية نتخذها حسب التحليل الحدي حيث ننظر إلى الإضافة إلى التكلفة مقارنة بالإضافة إلى الفائدة.

فالأفراد يستخدمون التحليل الحدي دون أن تتم تسمية كذلك، فالطالب مثلاً عندما يفكر في دراسة ساعة إضافية لمادة معينة فإنه ينظر إلى المنفعة المترتبة على ذلك (والتي يمكن ان تنحصر في فهم المادة، حل الواجب، الإستعداد للإمتحان...الخ) و يقارن ذلك بالتكلفة الإضافية لتلك الساعة والتي يمكن أن تقاس بالعمل البديل الممكن القيام به خلال تلك الساعة (كزيارة صديق، الذهاب للتنزه، النوم...الخ).

الإقتصاد الوصفي، الواقعي، التقريري (الحقيقى) والإقتصاد المعيارى (المثالى)

يهتم علم الإقتصاد كواحد من العلوم الإجتماعية بالتنبؤ أو بتحديد أثر التغير في العوامل الإقتصادية على السلوك البشرى ويحاول الإقتصاد الوصفي أو الحقيقى تفسير الواقع بمحاولة الإجابة على الأسئلة من شاكلة " ماذا يكون " فالإقتصاد الحقيقى يفترض وجود علاقة يمكن بحثها وتحليلها فمثلاً إذا إرتفع سعر اللحم فإن الكمية التي يشتريها الناس سوف تقل ويمكننا إحصائياً فحص العلاقة بين أسعار اللحم كمية المشتروات لتحديد صحة هذه المقولة اى ان الإقتصاد الحقيقى يصدر أحكاماً تقريرية موضوعية يمكن إختبار صحتها أو عدم صحتها بالرجوع إلى الواقع . لذلك فهو يشمل المبادئ

والنظريات الإقتصادية التي تبحث في طبيعة الأشياء ومن ثم فهي تجيب على الأسئلة المتعقة "بما هو كائن

أما الاقتصاد المعياري (المثالي) فيستخدم أحكاماً تقديرية قيمية بالإضافة إلى المعلومات التي يمدنا بها الإقتصاد الوصفي لتأييد سياسة معينة من بين سياسات بديلة . أى أن هدف الإقتصاد المعيارى هو الوصول إلى معايير تتعلق بما يجب أن يكون ولايمكن حسم هذه الأحكام بالرجوع إلى الواقع حيث أنها تعتمد على ذات الشخص الذى يدلى بها أو يحكم فيها من حيث حالته النفسية وإنتمائه الإجتماعى والفكرى. ومن أمثلتها القول بأنه " يجب إستمرار الدعم الحكومى للقطاع الزراعي في المملكة حتى نوفر للمزارع دخل معقول يتساوى مع بقية أفراد المجتمع في القطاعات الإقتصادية الأخرى " مثل هذا القول يرتكز على قيمة إجتماعية وهي العدالة التي لايمكن التثبت من صحتها بالرجوع إلى الواقع لأن العدالة معيار أو قيمة تتعلق بقضية فلسفية .

إذاً الإقتصاد المعياري يتضمن مجموعة النظريات والمبادئ الإقتصادية التي تجيب على الأسئلة المتعلقة بما يجب أن يكون (مثل المفاضلة بين تحقيق معدل مرتفع من النمو للدخل القومي و القضاء على البطالة وما يجب أن يكون عليه النظام الضريبي مثلاً) هذا النوع من التساؤلات يمثلها الدراسات المتعلقة بإقتصاديات الرفاهيه , وواضح أن هذه الأسئلة تبحث في مواضيع يدخل فيها عنصر تقديري يختلف بإختلافات الإقتصاديين ومن ثم لايمكن إختبارها بإستخدام المشاهدات العملية كما لايمكن الحكم بأفضلية إحداها.

الفصل الثالث

تلخيص البيانات

الـهدف من دراسة الفصل

بنـهاية الدراسة لـهذا الفصل يكون الدارس قادر على الالمام بالمفاهيم التالية:

١- أنواع البيانات الاحصائية

٢- طريقة عمل الفئات المنتظمة للبيانات.

٣- أهم أشكال العرض البياني للبيانات.

الفصل الثالث

تلخيص البيانات

مقدمة

بعد جمع البيانات سواء من المصادر التاريخية أو من المصادر الميدانية، فإنها تكون بيانات خامًا غير منتظمة عدديا ويصعب دراستها أو استنتاج أى شيء منها. ولذلك دعت الحاجة إلى تنظيم وتلخيص هذه البيانات بصورة يسهل فهمها واستنتاج بعض النتائج الأولية منها. ولتوضيح ذلك نعتبر المثال التالي:

مثال (١): إذا كان لدينا تقديرات ٦٠ طالبا كالتالي:

D	B	E	C	D	B	D	C	E	A
B	E	C	D	B	D	D	A	E	C
C	D	A	C	E	D	C	C	D	B
D	E	D	D	A	D	D	C	D	C
D	A	B	D	B	D	C	D	C	E
D	B	C	C	E	D	C	C	D	A

والبيانات السابقة بوضعها الحالى قد تجعل من الصعب التعرف على الطلاب الحاصلين على تقدير مشترك مثل ممتاز (A) أو جيد جدا (B)... ومن هنا أصبحت الحاجة إلى وضع التقديرات وتلخيصها فى جدول يسهل دراسته يسمى بجدول التوزيع التكراري، وقد تكون البيانات رقمية مثل درجات الطلاب أو أوزان الطلاب، أو أجور العمال فى أحد المصانع. ونوضح ذلك بالمثال التالي:

مثال (٢): البيانات التالية تمثل درجات ٥٠ طالب فى إحدى المواد:

51	95	70	74	73	90	71	74	90	67
91	72	83	89	50	80	72	84	85	69
62	82	87	76	91	76	87	75	78	79
71	96	81	88	64	82	73	57	86	70
80	81	75	85	74	90	83	66	77	91

البيانات السابقة بوضعها الحالى يصعب دراستها أو استنتاج بعض المؤشرات منها. فمثلا ما هو عدد الطلاب الذين حصلوا على 70 درجة فأكثر؟ أو عدد الطلاب الذين حصلوا على درجات تتراوح ما بين 80 درجة و 90 درجة...إلخ؟ ولذلك فإن أول مرحلة للتحليل الإحصائى تتكون من تصميم جدول التوزيع التكراري، وقبل التعرض لكيفية تنظيم هذه البيانات فى جداول تكرارية يلزم أن نتعرف على أنواع البيانات الإحصائية.

أنواع البيانات

البيانات الإحصائية نوعان:

بيانات وصفية (Qualitative Data)وبيانات كمية (Quantitative Data).

أ- البيانات الوصفية (Qualitative Data): هى البيانات التى تصف الأفراد والمجتمع مثل لون الشعر أو العين أو البشرة أو تقديرات الطلاب فى إحدى المواد، كما ورد فى مثال (١) السابق.

ب- البيانات الكمية (Quantitative Data): هى البيانات التى تقاس فيها الأفراد والمجتمع بمقاييس كمية(رقمية) مثل أطوال الطلاب (بالسنتيمتر) ، أو أوزان الطلاب (بالكجم) وأعمار الطلاب (بالسنة) أو نتيجة الامتحان (بالدرجات) أو أجور العمال (بالجنيه).

العرض الجدولى

الجدول التكرارى

تنظم وتلخص البيانات الإحصائية سواء كانت وصفية أو كمية فيما يسمى بالتوزيع التكرارى (Frequency Distribution). وهو عبارة عن جدول يلخص البيانات الخام فيوزعها على فئات ويحدد عدد الأفراد الذين ينتمون إلى كل فئة ويسمى هذا العدد بتكرار الفئة ويرمز له عادة بالرمز f، ولإتمام ذلك ينبغى أن يصمم جدول آخر يسمى بجدول تفريغ البيانات الإحصائية. وهو يتكون من ثلاث خانات: الخانة الأولى أو العمود الأول فيكتب فيه الصفة للبيانات الوصفية أو الفئة للبيانات الكمية، وفى الخانة الثانية توضع العلامات وهى عبارة عن حزم مكونة من خمسة خطوط، أربعة منها رأسية والخامس مائل يحزم الأربعة خطوط الرأسية، وبذلك تصبح الحزمة على الصورة (𝍸)، وفى الخانة الثالثة والأخيرة يكتب مجموع العلامات أمام كل صفة أو فئة كل على حدة، ومجموع هذه العلامات فى كل فئة يسمى بالتكرار لهذه الصفة أو الفئة. وبذلك يكون جدول تفريغ البيانات الإحصائية الوصفية فى مثال (١) وهو تقديرات النجاح للطلاب فى إحدى المواد كالتالي:

جدول (٢–١): تفريغ وتوزيع التقديرات للطلاب فى مثال (١)

التكرار (عدد الطلاب)	العلامات	الصفات
6		A
8		B
16		C
22		D
8		E
60		المجموع

ومن هذا الجدول نكون جدولا آخر يسمى بالجدول التكرارى أو جدول التوزيع التكرارى للبيانات الوصفية الذى يتكون من خانتين، الأولى تمثل الصفة والثانية تمثل التكرار، كما هو مبين بجدول رقم (٢–٢) كما يلي:

جدول (٢–٢): التوزيع التكرارى لتقديرات الطلاب فى مثال (١)

التكرار (عدد الطلاب)	الفئات
6	A
8	B
16	C
22	D
8	E
60	المجموع

وأحيانا يكتب الجدول السابق (٢–٢) فى صورة أفقية كما يلي:

جدول (٢–٣): التوزيع التكرارى لتقديرات الطلاب فى مثال (١)

المجموع	E	D	C	B	A	الصفة
60	8	22	16	8	6	التكرار

وبعد إلقاء الضوء على كيفية عمل التكرارات أمام الصفات وتكوين الجداول التكرارية للبيانات الوصفية فى الجداول السابقة، فإنه يلزم عمل فئات أو فترات منتظمة (متساوية الطول) كما يلي:

طريقة عمل الفئات المنتظمة للبيانات الكمية:

الغرض من عمل الفئات هو تجميع القيم المتقاربة فى مجموعات، ولا توجد هناك قواعد ثابتة لتحديد طول الفئات وعددها، إلا أنه من المرغوب فية أن لا يكون عدد الفئات صغيرًا فتضيع معالم التوزيع وتفقد كثيرا من التفاصيل. كما لا يكون عدد الفئات كبير جدا فتضيع الحكمة من التجميع فى فئات. ولتحديد عدد الفئات وطول كل فئة فإنه يعتمد إلى حد كبير على الخبرة ومدى البيانات (Range) وهو الفرق بين أكبر قراءة وأصغر قراءة كحد أقصى، ولتوضيح كيفية عمل الفئات المنتظمة نعتبر مثال (٢) السابق وتكون الخطوات كالتالي:

أ- نحسب طول المدى للقراءات R أى أن: $R = 97 - 50 = 47$

ب- نختار مثلا عدد الفئات = 5 فئات.

ج- نحسب طول الفئة بأن نقسم المدى على عدد الفئات بحيث يقرب الكسر إن وجد من خارج القسمة إلى الواحد الصحيح مهما كانت قيمة الكسر، وبذلك يكون طول الفئة L عددا صحيحا أى أن:

$$L = 47 / 5 = 9.4 \sim 10$$

د- نختار أصغر قراءة فى البيانات لتكون بداية الفئة الأولى المقربة ويضاف إليها طول الفئة فنحصل بذلك على بداية الفئة الثانية، وفى المثال (٢) بداية الفئة الأولى المقربة 50 فتكون بداية الفئة الثانية هي: $50+10 = 60$.

هـ- تحدد بداية الفئة الثالثة المقربة بإضافة طول الفئة لبداية الفئة الثانية المقربة، وهكذا لباقى الفئات.

و- لإيجاد نهاية أى فئة نضيف إلى بدايتها طول الفئة مطروحا منه واحد، وفى هذا المثال تكون نهاية الفئة الأولى المقربة هى 59 ونهاية الفئة الثانية المقربة 69 وهكذا لباقى الفئات. ويكون جدول تفريغ البيانات كما هو موضح بالجدول التالي:

جدول (٢-٤): تفريغ وتوزيع الدرجات للطلاب فى مثال (٢)

التكرار (عدد الطلاب)	العلامات	الصفات
3	|||	50-59
5	𝍸 |||	60-69
18	𝍸 𝍸 𝍸 |||	70-79
16	𝍸 𝍸 𝍸 |	80-89
8	𝍸 |||	90-99
50		المجموع

ويلخص من جدول التفريغ (٢-٤) جدول التوزيع التكرارى للبيانات الإحصائية الكمية الذى يتكون من خانتين. الأولى يكتب بها حدود الفئات والثانية يكتب بها التكرار، كما هو مبين بالجدول التالى (٢-٥):

جدول (٢-٥): التوزيع التكرارى لدرجات الطلاب فى مثال (٢)

التكرار (عدد الطلاب)	حدود الفئات
3	50-59
5	60-69
18	70-79
16	80-89
8	90-99
50	المجموع

والجدول السابق (٢-٥) يمكن أن يكتب فى صورة أفقية وذلك لتوفير حيز الكتابة كالآتي:

المجموع	90-99	80-89	70-79	60-69	50-59	الفئات
50	8	16	18	5	3	التكرار

ويمكن تكوين جدولين آخرين من جدول التوزيع التكرارى (٢-٥) وهما:

- الجدول التكرارى النسبى **Relative Frequency Table**.
- الجدول التكرارى المئوى **Percentage Frequency Table**.

٢-٢-٢ الجدول التكرارى النسبي

يتكون الجدول التكرارى النسبى من خانتين مثل الجدول التكرارى العادى ولكن خانة التكرار يكتب بها التكرار النسبي، وهو عبارة عن التكرار لأى فئة مقسوما على مجموع التكرارات. ويكون مجموع التكرار النسبى لجميع الفئات مساويًا للواحد الصحيح، كما هو موضح بالجدول التالي:

جدول (٢-٦): التوزيع التكرارى النسبى لدرجات الطلاب فى مثال (٢)

التكرار النسبي	حدود الفئات
0.06	50-59
0.10	60-69

70-79	0.36
80-89	0.32
90-99	0.16
المجموع	1.00

٢-٢-٣ الجدول التكرارى المئوى

الجدول التكرارى المئوى للبيانات الإحصائية يتكون من خانتين أيضا مثل الجدول التكرارى النسبى السابق، ولكن فى خانة التكرارات النسبية تكتب التكرارات المئوية، ويمكن الحصول عليها بضرب التكرار النسبى فى ١٠٠. ويلاحظ أن مجموع التكرارات المئوية يساوى ١٠٠ وبذلك يكون الجدول التكرارى المئوى للبيانات فى المثال (٢) كالتالى جدول (٢-٧):

جدول (٢-٧): التوزيع التكرارى المئوى لدرجات الطلاب فى مثال (٢)

حدود الفئات	التكرار المئوي
50-59	6
60-69	10
70-79	36
80-89	32
90-99	16
المجموع	100

٢-٢-٤ الحدود الحقيقية (الفعلية) للفئات

البيانات الإحصائية المراد تلخيصها وتنظيمها فى جداول تكرارية عادة تكون مكتوبة مقربة مثلا لأقرب وحدة قياس أو لأقرب نصف وحدة قياس. فإذا كانت البيانات مقربة لأرقام صحيحة فإننا نطرح من الحد الأدنى المقرب للفئة ٠.٥ لنحصل على الحد الأدنى الحقيقي، ونضيف ٠.٥ إلى الحد الأعلى المقرب لنحصل على الحد الأعلى الحقيقى للفئة. وهكذا لباقى الفئات للحصول على الحدود الحقيقية لها. أما إذا كانت البيانات محسوبة لأقرب رقم عشرى فإننا نطرح ٠.٠٥ من الحد الأدنى المقرب للفئة لنحصل على الحد الأدنى الحقيقى لها، ونضيف ٠.٠٥ إلى الحد الأعلى المقرب لنحصل على الحد الأعلى الحقيقى للفئة، وهكذا لباقى الفئات. وبالمثل يمكن إيجاد أى حدود حقيقية مهما كانت أعداد الأرقام العشرية المقربة بنفس الطريقة السابقة، وبذلك يكون جدول التوزيع التكرارى (٢-٥) مستخدما الحدود الحقيقية للفئات كالآتي:

جدول (٢–٨): التوزيع التكرارى لدرجات الطلاب بالحدود الفعلية للفئات فى مثال (٢)

الحدود الفعلية للفئات	التكرار (عدد الطلاب)
49.5-59.5	3
59.5-69.5	5
69.5-79.5	18
79.5-89.5	16
89.5-99.5	8
المجموع	50

يعرف مركز الفئة Class Mark بالعلاقة الأتية:

مركز الفئة = (الحد الأعلى للفئة + الحد الأدنى للفئة) / ٢

ومركز الفئة لا يتأثر بحدود الفئة سواء كانت حدودًا مقربة أو حدودًا حقيقية، ففى مثال (٢) جدول (٢–٥) نحسب منه مركز الفئة الأولى من الحدود المقربة = (٥٠+٥٩) / ٢= ٥٤.٥ درجة ومن جدول (٢–٨) نحسب مركز الفئة من الحدود الحقيقية = (٤٩.٥ + ٥٩.٥) / ٢ = ٥٤.٥٥ وباستخدام نفس الطريقة يمكن حساب باقى مراكز الفئات الأخرى، أو بإضافة طول الفئة إلى مركز الفئة الأولى لنحصل على مركز الفئة الثانية وهكذا لباقى الفئات. ويمكن تنظيم الجداول السابقة للبيانات الكمية فى جدول واحد يشمل الحدود المقربة والحقيقية للفئات ومراكز الفئات والتكرار والتكرار النسبى والتكرار المئوي، ومن مثال (٢) يكون جدول شامل لتوزيع درجات الطلاب كما هو موجود فى جدول

جدول (٢–٩): التوزيع التكرارى لدرجات الطلاب فى مثال (٢) وتلخيص للجداول السابقة

الحدود المقربة للفئات	الحدود الحقيقية	مراكز الفئات	التكرار	التكرار النسبي	التكرار المئوي
50-59	49.5-59.5	54.5	3	0.06	6
60-69	59.5-69.5	64.5	5	0.10	10
70-79	69.5-79.5	74.5	18	0.36	36
80-89	79.5-89.5	84.5	16	0.32	32
90-99	89.5-99.5	94.5	8	0.16	16
المجموع			50	1.00	100

<u>ملاحظة هامة</u>: لسهولة بناء الجداول الإحصائية المستنتجة من الجدول التكرارى وكذلك الحسابات الإحصائية التى سوف نتعرض لها بالشرح فيما بعد يجب أن تكون حدود الفئات فى الجداول التكرارية حدودًا حقيقية.

٢-٢-٥ الجدول التكرارى المتجمع الصاعد "Cumulative Frequency "Less Than

فى كثير من الأحيان يكون اهتمامنا منصبا على عدد القراءات التى تكون أصغر من أو تساوى مقدارا معينا، ففى مثال (٢) يمكن أن يطلب ما هو عدد الطلاب الحاصلين على ٧٩ درجة فأقل؟ فتكون الإجابة: عدد الطلاب الحاصلين على ٧٩ درجة فأقل هو: ٣+٥+١٨=٢٦ وهذا هو التكرار المتجمع الصاعد للفئة الثالثة. وكذلك يمكن استخدام الجدول فى إيجاد عدد الطلاب الذين تنحصر درجاتهم بين حدين معلومين. ويمكن كتابة الجدول التكرارى المتجمع الصاعد المكون من خانتين، الأولى يكتب فى السطر الأول منها أقل من الحد الأدنى الحقيقى للفئة الأولى (بدلا من حدود الفئة الأولى وكذلك لباقى الفئات حتى نصل إلى الفئة الأخيرة فيكتب لها سطرين الأول منهما: أقل من الحد الأدنى الحقيقى للفئة الأخيرة والثانى منها أقل من الحد الأعلى الحقيقى للفئة الأخيرة، كما سيوضح فى جدول (٢-١٠) للتوزيع التكرارى المتجمع الصاعد لدرجات الطلاب فى مثال (٢):

جدول (٢-١٠): التوزيع التكرارى المتجمع الصاعد لدرجات الطلاب فى مثال (٢)

حدود الفئات	التكرار المتجمع الصاعد
< 49.5	0
<59.5	3
<69.5	8
<79.5	26
<89.5	42
<99.5	50

٢-٢-٦ الجدول التكرارى المتجمع الهابط "Cumulative Frequency "or More

قد يكون اهتمامنا أحيانًا منصبًا على عدد القيم التى تكون أكبر من أو تساوى قيمة معينة، ففى مثال (٢) قد يطلب معرفة عدد الطلاب الحاصلين على ٧٩ درجة فأكثر؟ فتكون الإجابة هي: عدد الطلاب الحاصلين على ٧٩ درجة فأكثر هو: ٨+١٦=٢٤ درجة والجدول التكرارى المتجمع الهابط مثل الجدول التكرارى المتجمع الصاعد مكون من خانتين، الأولى يكتب فى السطر الأول منها أكبرمن الحد الأدنى الحقيقى للفئة الأولى (بدلا من حدود الفئة الأولى) وكذلك لباقى الفئات حتى نصل إلى الفئة الأخيرة فيكتب لها سطرين الأول منهما: أكبرمن الحد الأدنى الحقيقى للفئة الأخيرة والثانى منها أكبرمن الحد الأعلى الحقيقى للفئة الأخيرة، وجدول (٢-١١) يوضح التوزيع التكرارى المتجمع الهابط لدرجات الطلاب فى مثال (٢):

جدول (٢-١١): التوزيع التكرارى المتجمع الهابط لدرجات الطلاب فى مثال (٢)

حدود الفئات	التكرار المتجمع الهابط
> 49.5	50
>59.5	47
>69.5	42
>79.5	24
>89.5	8
>99.5	0

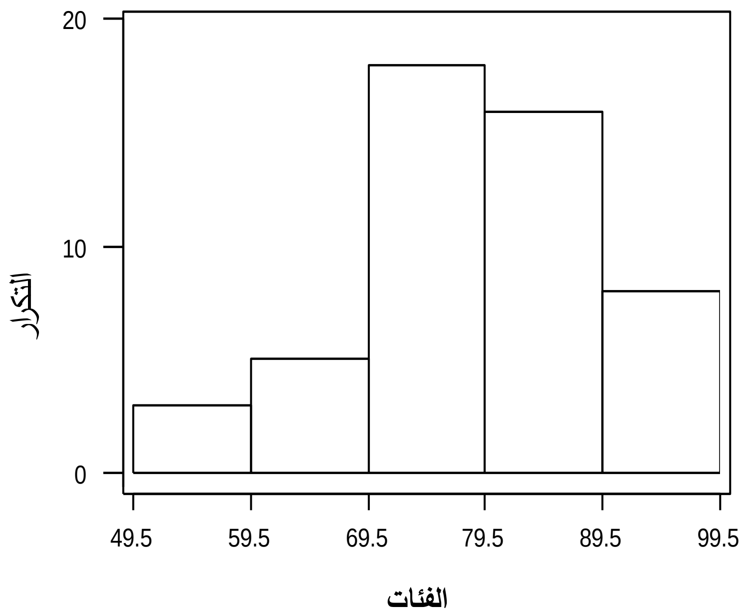
الفئات غير المنتظمة

سبق الحديث عن الفئات المنتظمة وهى غالبا ما تكون ذات أهمية كبيرة وخاصة فى العمليات الإحصائية التى سوف نتعرض لها فيما بعد، وذلك لسهولتها فى التطبيق بدون تعديل التكرار لها. ولكننا أحيانا نضطر إلى استخدام فئات غير منتظمة فى بعض الظواهر محل الدراسة لأن الفئات المنتظمة قد لا تفى بالغرض، وذلك بأن يكون تكرارها قليلا أو خاليا من التكرار مثل ظاهرة الدخول للأفراد أو الأجور أو درجات الامتحان للطلاب أو الوفيات للأطفال الرضع (أقل من سنة)، فإن عمل جدول ذى فئات غير منتظمة يكون مناسبا. ولكن عند رسم المدرج التكرارى أو غيرة من الرسوم البيانية فإنه يتطلب تعديل التكرار للفئات غير المنتظمة حتى يصبح الرسم ممثلا لهذه البيانات. وسوف نتناول طريقة تعديل التكرارات عند رسم المدرج التكرارى فيما بعد.

العرض البيانى

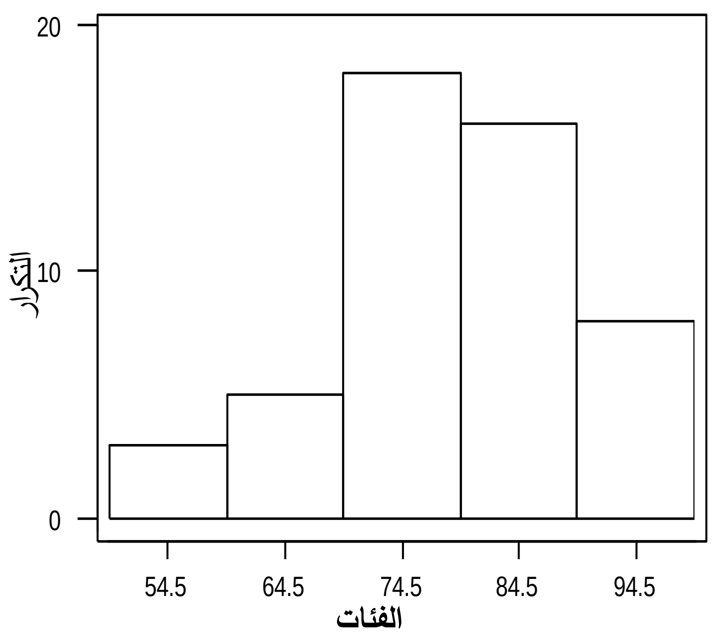
المدرج التكرارى

نرسم المدرج التكرارى على محورين متعامدين أحدهما أفقى يمثل الفئات والثانى رأسى يمثل التكرار. ونرسم مستطيلات متلاصقة على الفئات قاعدتها طول الفئة محسوبا من الحدود الحقيقية، وارتفاعاتها عبارة عن تكرار هذه الفئات. فمثلا بالنسبة إلى الفئة الأولى يكون المستطيل قاعدته بادئة من الحد الأدنى للفئة الأولى، ومنتهية بالحد الأعلى للفئة الأولى. وارتفاع المستطيل هو تكرار الفئة الأولى. وهكذا لباقى المستطيلات التى تمثل باقى التكرارات والمدرج التكرارى للبيانات الموجودة فى الجدول (٢-٥) موضح بالشكل التالى (٢-١):



شكل (٢–١): المدرج التكراري

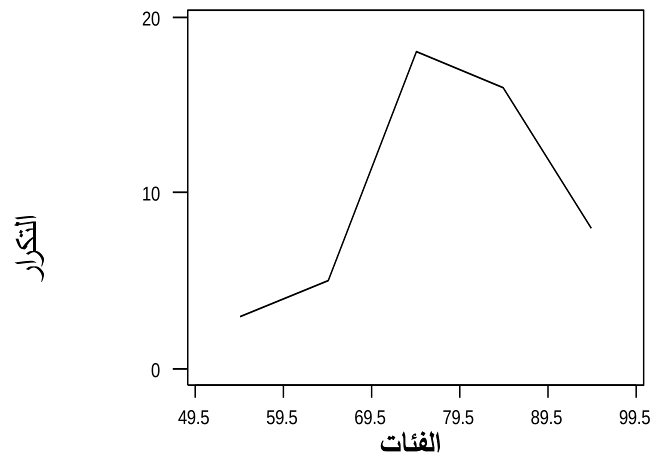
ويمكن رسم المدرج التكرارى بطريقة أخرى، وهى أن تحدد مراكز الفئات على المحور الأفقى ومنها يرسم ارتفاع المستطيل الممثل للتكرار فى منتصف القاعدة للمستطيل على أن يكون البعد من أحد جوانب مركز الفئة مساويا لبعد الجانب الآخر، ويستكمل رسم المستطبل للفئة الأولى وتتبع نفس الطريقة لباقى الفئات. ويوضح رسم المدرج التكرارى لجدول (٢–٥) بهذه الطريقة كما هو مبين بشكل (٢–٢):



شكل (٢–٢): رسم المدرج التكرارى بطريقة أخرى

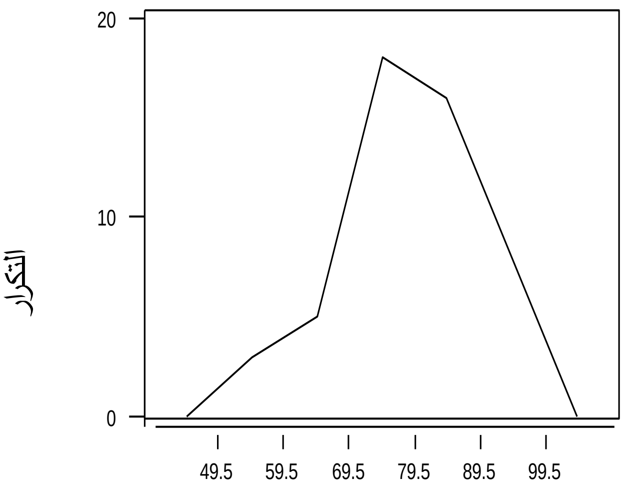
المضلع التكراري

يرسم المضلـع التكرارى علـى محورين، الأفقى يمثل الفئـات والرأسى يمثل التكرار، مثل ما ورد شرحه فى طريقة رسم المدرج التكراري، وبدلا من رسم مستطيل ارتفاعة يمثل التكرار نضع نقطة واحدة فقط على ارتفاع يمثل التكرار لهذه الفئـة وذلك عند منتصف الفئة. ويكرر رسم النقاط لباقى التكرار بحيث تكون ارتفاعاتها ممثلة لتكرار تلك الفئات وذلك من منتصفاتها، لأننا نفترض انتظام توزيع التكرارات داخل كل فئة. وبعد ذلك نصل بخط مستقيم كل نقطتين متجاورتين فنحصل على المضلع التكرارى شكل (٢-٣).



شكل (٢-٣): المضلع التكراري

ولغلق المضلع فى شكل (٢-٣) مع محور الفئات نضع نقطة على محور الفئات يسار الحد الأدنى الحقيقى للفئة الأولى على بعد يساوى نصف طول الفئة، ثم نصل بخط مستقيم هذه النقطة بالنقطة التى سبق وضعها فى مركز الفئة الأولى. ثم نضع نقطة على محور الفئات يمين الحد الأعلى الحقيقى للفئة الأخيرة تبعد مسافة قدرها نصف طول الفئة عن يمين الحد الأعلى الحقيقى للفئة الأخيرة وتبعد عنها مسافة قدرها نصف طول الفئة ثم نوصلها بخط مستقيم بالنقطة التى سبق وضعها فى منتصف الفئة الأخيرة، ولكى يكون المضلع صحيحا يجب أن يكون مغلقا. ويبين المضلع التكرارى المغلق لجدول (٢-٥) بالشكل التالى (٢-٤).



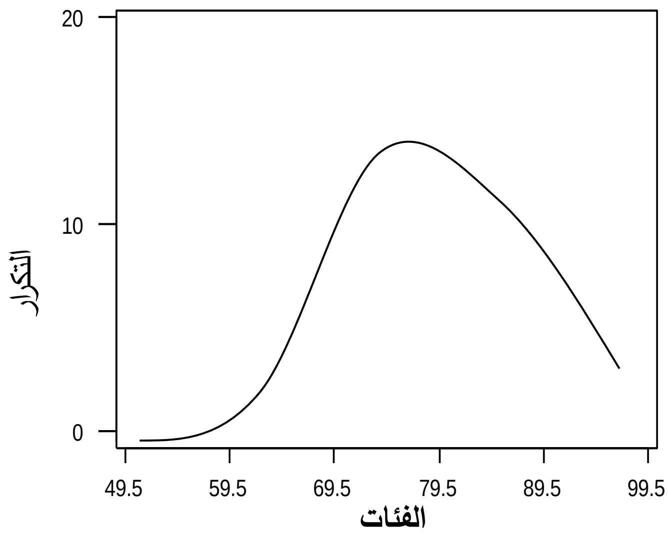
شكل (٢-٤): المضلع التكرارى المغلق

٢-٣-٣ المنحنى التكرارى الممهد

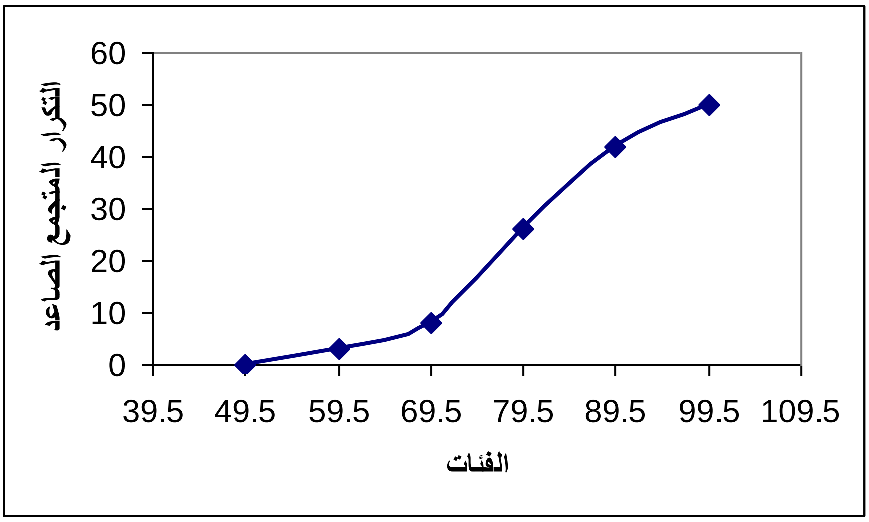
يرسم المنحنى التكرارى الممهد على محورين متعامدين الأفقى يمثل الفئات والرأسى يمثل التكرار. ويتم رسم النقاط مثل ما اتبع فى المضلع التكراري، ويمهد المنحنى التكرارى باليد كى يأخذ شكلاً انسيابيًا حتى لو لزم الأمر عدم المرور ببعض النقاط. المنحنى التكرارى الممهد للبيانات فى جدول (٢-٥) يرسم كما هو بالشكل (٢-٥).

٢-٣-٤ المنحنى التكرارى المتجمع الصاعد

يرسم المنحنى التكرارى المتجمع الصاعد على محورين متعامدين الأفقى يمثل الفئات والرأسى يمثل التكرارات المتجمعة الصاعدة وتوضع النقاط فى الرسم أعلى الحدود الدنيا الحقيقية للفئات بحيث يكون الارتفاع ممثلا للتكرار المتجمع الصاعد. والمنحنى المتجمع الصاعد الممثل فى الجدول (٢-١٠) يوضح بالرسم شكل (٢-٦).



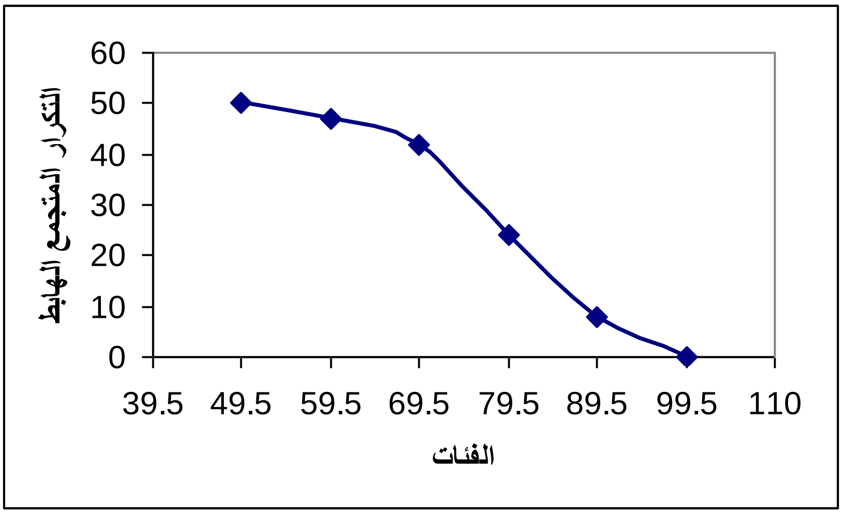
شكل (٢-٥): المنحنى التكرارى الممهد



شكل (٢-٦): المنحنى التكرارى المتجمع الصاعد

٢-٣-٥ المنحنى التكرارى المتجمع الهابط

يمثل المنحنى التكرارى المتجمع الهابط على محورين متعامدين مثل ما تم بالنسبة للمنحنى المتجمع الصاعد، بحيث يمثل المحور الأفقى الحدود الدنيا الحقيقية للفئات والرأسى يمثل التكرارات المتجمعة الهابطة، ويمهد المنحنى باليد لنحصل على المنحنى التكرارى المتجمع الهابط كما هو موضح شكل (٢-٧) للبيانات فى مثال (٢) وفى جدول (٢-١١):



شكل (٢-٧): المنحنى التكرارى المتجمع الهابط

٢-٣-٦ المدرج التكرارى فى حالة الفئات الغير منتظمة

فى حالة رسم المدرج التكرارى من فئات منتظمة كانت مساحة كل مستطيل تعبر عن التكرار الواقع فى كل فئة. وحيث ُإن الفئات متساوية فى أطوالها فإن المدرج التكرارى عبارة عن مستطيلات متلاصقة ومتساوية القاعدة وارتفاعاتها تتناسب مع التكرار. أما إذا كانت الفئات غير متساوية الطول تكون مساحة هذه المستطيلات المتلاصقة غير متناسبة مع التكرار وكذلك ارتفاعاتها، لذلك يجب تعديل التكرار قبل رسم المدرج التكرارى للفئات غير المتساوية حتى يصبح التكرار المعدل متناسبًا مع ارتفاع المستطيل الخاص بالفئة غير منتظمة الطول. وهناك طريقتان لتعديل الجدول التكراري، الطريقة الأولى: هى أن نقسم التكرار الأصلى لكل فئة على طولها فنحصل على تكرار معدل لجميع الفئات، أما الطريقة الثانية: فهى أن نعدل تكرار الفئات غير المنتظمة فقط، ويترك التكرار للفئات المنتظمة الباقية كما هو ويعدل التكرار للفئة غير المنتظمة بالعلاقة التالية:

التكرار المعدل = (التكرار الفعلى للفئة غير المنتظمة × طول الفئة المنتظمة) / طول الفئة غير المنتظمة ونوضح ذلك بالمثال التالي:

مثال (٥): الجدول التالى (٢-١٢) يبين التوزيع التكرارى لفئات غير منتظمة لدرجات الطلاب فى مثال (٢) السابق. ارسم المدرج التكرارى وذلك بعد تعديل التكرارات.

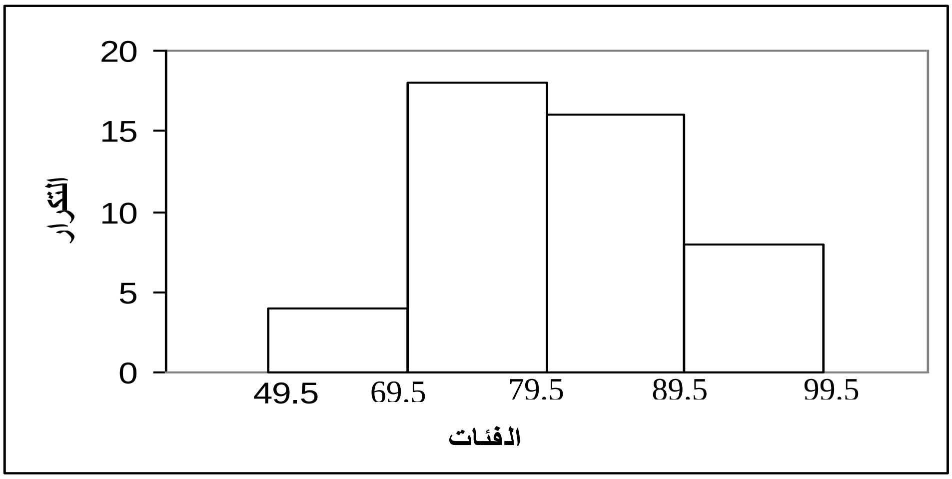
جدول (٢-١٢): التوزيع التكرارى لفئات غير منتظمة لدرجات الطلاب

المجموع	90-99	80-89	70-79	50-69	الفئات
50	8	16	18	8	التكرار

الحل: الجدول التالى (٢-١٣) يوضح كل من التكرار الأصلى والتكرار المعدل لجدول (٢-١٢) كالآتي:

89.5-99.5	79.5-89.5	69.5-79.5	49.5-69.5	حدود الفئات
8	16	18	8	التكرار
8	16	18	4	التكرار المعدل

ويفسر نقص مجموع التكرار المعدل لكبر طول الفئة غير المنتظمة عن طول الفئة المنتظمة، ثم نرسم من جدول (٢-١٣) المدرج التكرارى كما يلي:



ملاحظة: بما أن الإحداثى الصادى (المحور الرأسي) يتناسب مع مساحة المستطيل أو التكرار (فى حالة المنحنى المتجمع الصاعد) فإن استعمال التكرار الفعلى أو النسبى أو المئوى شيء واحد.

تمارين

١-فيما يلى أوزان ٨٠ فأرا من فئران التجارب بالجرام وذلك عند دراسة نقص الفيتامين.

132	125	117	124	103	117	110	127	96	129
130	122	118	114	103	119	106	125	114	100
125	128	106	111	115	123	119	114	117	143
136	92	115	118	121	137	139	120	104	125
119	115	101	129	87	108	110	133	135	126
127	103	110	126	118	82	104	137	120	95
146	126	119	105	132	126	118	100	113	119
106	125	117	102	146	129	124	113	95	148

أ- كون جدول التوزيع التكرارى مستخدما أطوال الفئات الأتية:

80-89, 90-99, 100-109, …, 140-149

ب- ارسم المدرج التكرارى والمضلع التكراري.

ج- ارسم المدرج التكرارى النسبى والمضلع التكرارى النسبي.

د- ارسم المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط لهذه البيانات.

هـ- اوجد عدد الفئران الذين تقل أوزانهم عن ١٢٥ جراما.

٢-فيما يلى بيانات ٤٠ من الطلبة:

44	98	40	60	66	71	82	64	72	68
55	69	77	78	88	60	65	68	79	69
62	64	71	66	61	75	83	70	55	62
57	72	61	62	74	62	67	66	60	50

أ- كون جدول التوزيع التكرارى مستخدما أطوال الفئات الآتية"

40-49, 50-59,…, 90-99

ب- ارسم المدرج التكرارى والمنحنى التكراري.

ج- ارسم المنحنى المتجمع الصاعد النسبى والمنحنى المتجمع الهابط النسبى لهذه البيانات.

د- إذا علم أن:

الدرجات	التقدير
0-59	E
60-69	D
70-79	C
80-89	B
90-99	A

اوجد جدول توزيع التقديرات لدرجات الطلاب

٣- فيما يلى أجور ٧٠ عاملا فى إحدى الشركات فى اليوم الواحد.

فئات الأجور	50-59	60-69	70-79	80-89	90-99	100-119	120-129
عدد العمال	8	10	16	15	10	8	3

أ- ارسم المضلع التكرارى لهذه البيانات.

ب- ارسم المنحنى التكرارى والمنحنى المتجمع الصاعد لهذه البيانات.

ج- ارسم المنحنى المتجمع الهابط لهذه البيانات.

٤- إذا أعطيت الجدول التكرارى التالى لأطوال ألف طالب مقاسة بالسنتيمترات، فارسم مدرجها التكرارى مبينا حدود الفئات.

طول الطالب	155-	158-	161-	164-	167-	170-	173-	176-
عدد الطلاب	4	10	77	235	368	220	80	6

الفصل الرابع

بعض مقاييس النزعة المركزية

الهدف من دراسة الفصل

بنهاية الدراسة لهذا الفصل يكون الدارس قادر على الالمام بالمفاهيم التالية:

١- الوسط , الوسيط , المنوال .,,,, الخ

٢- طريقة حساب الوسط , الوسيط , المنوال .,,,, الخ للعينة .

٣- طريقة حساب الوسط , الوسيط , المنوال .,,,, الخ للمجتمع .

الفصل الرابع
مقاييس النزعة المركزيـة

مقدمة

رأينا فى الفصل السابق كيفية عرض البيانات الإحصائية وتلخيصها فى جداول تكرارية أو رسوم بيانية، بهدف الحصول على بعض الخصائص للمجتمع الإحصائى محل الدراسـة. ومن المعروف عادة أن الرسوم البيانية تكون غير دقيقة، لذلك يجب أن يكون لدينا مقاييس عددية تصف لنا هذه البيانات. وسوف نتعرض فى هذا الفصل إلى نوع مهم من المقاييس الإحصائية وهو ما يسمى بمقاييس النزعة المركزية أو مقاييس الموضع أو المتوسطات. وهى مقاييس عددية تعين موقع التوزيع، وهى مهمة فى حالة المقارنة بين التوزيعات المختلفة بوجه عام. وتكون فائدتها أكثر فى حالـة التوزيعات المتشابهة فى طبيعتها وأشكالها ولكنها مختلفة فى مواقعها. فمثلا: عند دراسة عينة من البيانات الإحصائية التى تخص بعض الأسر من الريف حسب فئات الأنفاق الأستهلاكى السنوي، وعينة من البيانات الإحصائية التى تخص بعض الأسر من الحضر حسب فئات الانفاق الاستهلاكى أيضًا، فإن حساب المتوسط السنوى للإنفاق لكل من الريف والحضر يمكننا من المقارنة بينهما.

ويمكن تعريـف المتوسطات بأنهـا القيمـة النموذجيـة الممثلـة لمجموعـة مـن البيانـات. وحيث إن القيمة النموذجية تميل إلى الوقوع فى المركز، لذلك فإنه يمكن أن تسمى المتوسطات بمقاييس النزعة المركزيـة. وسـوف نتنـاول فـى هذا البـاب دراسـة المتوسطات فـى صـور مختلفـة، كل منهـا لـه مميزاتـه وعيوبه، وهذا يعتمد على طبيعة البيانـات والهدف من استخدامها. وبعض المتوسطات الأكثر شيوعا والتـى سـوف نتناولهـا بالشـرح والتفصـيل والأمثلـة وكيفيـة حسـابها، وهـى كـل مـن الوسـط الحسـابى (المتوسـط) والوسط المـرجح والوسـيط والمنـوال والوسـط الهندسـى والوسط التـوافقي، وذلـك فـى حالـة البيانات الخام والمبوبه.

تعريف رمز التجميع $\sum$

إذا كان لدينا مجموعة من المشـاهدات $(x_1, x_2, \ldots x_n)$ فإن حاصل جمـع هذه المشـاهدات يمكن التعبير عنه كما يلي:

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$$

وفى بعض الأحيان يكتب:

$$\sum x$$

وهـذا معنـاه حاصـل جمـع قـيم x. وإذا كـان لـدينا متغيـر آخـر y لمجموعـة مـن المشـاهدات $(y_1, y_2, .. y_n)$ وكذلك مقدار ثابت c فإنه يمكن إثبات بعض العلاقات الآتية:

(i) $\sum_{i=1}^{n}(x_i \pm y_i) = \sum_{i=1}^{n} x_i \pm \sum_{i=1}^{n} y_i$

(ii) $\sum_{i=1}^{n} c = nc$

(iii) $\sum_{i=1}^{n} c\,x_i = c \sum_{i=1}^{n} x_i$

وهذه العلاقات قد تكون مفيدة فى إثبات بعض الخصائص لبعض المقاييس المختلفة.

٣-١ الوسط الحسابى (المتوسط)

المتوسط أو الوسط الحسابى يعتبر من أهم مقاييس النزعة المركزية والأكثر استخداما فى الإحصاء والحياة العملية، إذ يستخدم عادة فى الكثير من المقارنات بين الظواهر المختلفة. ولو أسندت قيمة المتوسط لكل مشاهدة فإن مجموع هذه القيم الجديدة يكون مساويا لمجموع المشاهدات الأصلية (أنظر مثال (١)) ويعرف كالتالي:

إذا كان لدينا مجموعة من المشاهدات للمتغير x وهى $(x_1, x_2, .. x_n)$ فإن الوسط الحسابى يساوى حاصل جمع المشاهدات أو البيانات مقسوما على عددها ويرمز له بالرمز $\bar{x}$ وعليه فإن:

$$\bar{x} = \frac{x_1 + x_2 + .. + x_n}{n}$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i$$

مثال (١): إذا كانت درجات ٥ طلاب فى إحدى المواد هي:

60, 72, 40, 80, 63

احسب الوسط الحسابى لدرجات الطلاب.

الحل:

$$\sum_{i=1}^{5} x_i = 60 + 72 + 40 + 80 + 63$$

$$= 315$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{5}(315) = 63$$

الأن لو عوضنا بدل القراءة الأولى ٦٠ بالمتوسط ٦٣ وبالقراءة الثانية ٧٢ بالمتوسط ٦٣وبالقراءة الثالثة ٤٠ بالمتوسط ٦٣ إلخ نجد أن:

$$= 63 + 63 + 63 + 63 + 63 = 315 \sum_{i=1}^{5} x_i$$

وذلك كما ذكرنا فى الملاحظة السابقة فى تعريف المتوسط.

الوسط الحسابى للبيانات المبوبة

إذا كان لدينا عدد k من الفئات ذات المراكز $(x_1, x_2, .. x_k)$ ولها تكرارات $f_1, f_2, .. f_k$ على الترتيب فإن الوسط الحسابى يعطى بالعلاقة الآتية:

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + .. + f_k x_k}{f_1 + f_2 + .. + f_k}$$

$$= \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$$

$$= \frac{1}{n} \sum_{i=1}^{k} f_i x_i$$

حيث $n = \sum_{i=1}^{k} f_i$ ونوضح ذلك بالمثال التالي:

مثال (٢): احسب متوسط أعمار الطلاب $\bar{x}$ للبيانات التالية:

فئات العمر	5-6	7-8	9-10	11-12	13-14
عدد الطلاب	2	5	8	4	1

الحل: لسهولة الحل نضع الجدول التالي:

الفئات	مراكز الفئات (x)	التكرار (f)	$x\,f$
5-6	5.5	2	11
7-8	7.5	5	37.5
9-10	9.5	8	76
11-12	11.5	4	46
13-14	13.5	1	13.5
المجموع		20	184

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{k} f_i x_i = \frac{1}{20}(184) = 9.2 \text{ سنة}$$

بعض خصائص الوسط الحسابي

سوف نعرض بعض خصائص الوسط الحسابى وذلك فى حالة البيانات المباشرة، حيث يكون من السهل على الطالب استخدام نفس الطريقة لإثباتها فى حالة البيانات المبوبة.

الخاصية الأولى: المجموع الجبرى لانحرافات القيم عن وسطها الحسابى يساوى صفر، أى أنه إذا كانت مجموعة المشاهدات هى $(x_1, x_2, \dots x_n)$ وانحرافاتها عن وسطها الحسابى هى $(d_1, d_2, \dots d_n)$ حيث:

$$d_i = x_i - \bar{x}, \qquad i = 1, 2, \dots n$$

فإن

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} d_i = \mathbf{0}$$

الإثبات:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \Rightarrow n\,\bar{x} = \sum_{i=1}^{n} x_i \qquad \text{.....(1)}$$

$$\because \sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\bar{x} = \sum_{i=1}^{n} x_i - n\,\bar{x} = \mathbf{0} \qquad \text{.... (2)}$$

الخاصية الثانية: إذا كان للمتغير x المشاهدات $(x_1, x_2, \dots x_n)$ فإذا أضفنا أو طرحنا من القيم الأصلية للمتغير x مقدار ثابت b فإن الانحرافات $(d_1, d_2, \dots d_n)$ حيث:

$$d_i = x_i \pm b, \qquad i = 1, 2, \dots n$$

تعطى المتوسط كالتالي:

$$\bar{x} = \bar{d} + b$$

الاثبات:

$$\because d_i = x_i \pm b, \qquad i = 1, 2, \dots n$$

$$\Rightarrow \sum_{i=1}^{n} d_i = \sum_{i=1}^{n}(x_i \pm b) = \sum_{i=1}^{n} x_i \pm n\,b$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{1}{n}\sum_{i=1}^{n} x_i \pm b$$

$$\Rightarrow \bar{d} = \bar{x} \pm b$$

$$\Rightarrow \bar{x} = \bar{d} + b$$

مثال (٣): من مثال (١) إذا أخذنا درجة $b = 5($ فإن مجموع الانحرافات $\sum_{i=1}^{5} d_i$ يحسب كالتالي:

$$\sum_{i=1}^{5} d_i = \mathbf{(60 - 50) + (72 - 50\ ) + (40 - 50\ ) + (80 - 50\ ) + (63 - 50\ )}$$

$$= \mathbf{10 + 22 - 10 + 30 + 13 = 65}$$

$$\bar{d} = \frac{1}{5}\ \mathbf{(65) = 13}$$

$$\bar{x} = \bar{d} + b = \mathbf{13 + 50 = 63}$$

وهى نفس النتيجة التى حصلنا عليها فى مثال (١).

الخاصية الثالثة: إذا كان للمتغير x المشاهدات $(x_1, x_2, \dots x_n)$ وضربنا هذه المشاهدات فى مقدار ثابت حقيقى a فإن متوسط القيم الجديدة يساوى المتوسط x مضروبا فى a أى أن:

$$\overline{(ax)} = \bar{x}\ a$$

ومن الخاصيتين الثانية والثالثة ينتج أن:

$$\overline{(ax \pm b)} = a\,\overline{x} \pm b$$

<u>الخاصية الرابعة:</u> إذا كان للمتغير x المشاهدات $(x_1, x_2, \ldots x_n)$ فإن مجموع مربعات انحرافات المشاهدات عن أى قيمة حقيقية c يكون أكبر أو يساوى مجموع مربعات انحرافات المشاهدات عن وسطها الحسابي، أى أن:

$$\sum_{i=1}^{n}(x_i - c)^2 \geq \sum_{i=1}^{n}(x_i - \overline{x})^2 \qquad , \overline{x} \neq c$$

بعض مميزات الوسط الحسابي:

١- مقياس سهل حسابه ويخضع للعمليات الجبرية بسهولة.

٢- يأخذ فى الاعتبار جميع القيم محل الدراسة.

٣- أكثر المقاييس استخدامًا فى الإحصاء.

بعض عيوب الوسط الحسابي:

١- يتأثر بالقيم الشاذة (المتطرفة) وهى القيم الكبيرة جدا أو الصغيرة جدا مقارنة بباقى القيم.

٢- يصعب حسابه فى حالة الجداول التكرارية المفتوحة، حيث يتطلب ذلك معرفة مركز كل فئة.

٣- لا يمكن حسابه فى حالة البيانات الوصفية.

<u>الوسيط</u>

عند ترتيب البيانات (المشاهدات) ترتيبا تصاعديا أو تنازليا فإن الوسيط يكون البيان (المشاهدة) التى يقع ٥٠٪ من البيانات قبلها فى الترتيب و ٥٠٪ من البيانات بعدها فى الترتيب. فإذا كان عدد البيانات فرديا فإن الوسيط يكون المشاهدة التى تقع فى المنتصف، وإذا كان عدد البيانات زوجيا فإن الوسيط هو متوسط المشاهدتين اللتين تقعان فى المنتصف.

مثال(٤): اوجد الوسيط لدرجات الطلاب فى مثال (١) حيث كانت درجاتهم كالتالي:

60, 72, 40, 80, 63

الحل: يتم ترتيب البيانات تصاعديا كالتالي:

40, 60, 63, 72, 80

بما أن عدد المشاهدات فردى فيكون الوسيط هو المشاهدة التى تقع فى منتصف هذه البيانات. وعليه فإن الوسيط هو المشاهدة رقم (٣) وقيمتها 63 (نلاحظ أن هناك بيانين يقعان فى الترتيب قبل الوسيط وبيانين يقعان بعده).

مثال (٥): اوجد الوسيط لدرجات الطلاب الآتية:

72, 60, 72, 40, 80, 63

الحل: يتم ترتيب البيانات تصاعديا كالتالي:

40, 60, 63, 72, 72, 80

بما أن عدد البيانات زوجى فتكون قيمة الوسيط هى متوسط قيمتى المشاهدتين اللتين يقعان فى المنتصف.

Med = (63+72) / 2 = 67.5

الوسيط فى حالة البيانات المبوبة (الجداول التكرارية)

نتبع الخطوات التالية لحساب الوسيط حسابيا وبيانيا.

الوسيط حسابيا:

١- نكون الجدول المتجمع الصاعد (باستخدام الحدود الحقيقية).

٢- نوجد رتبة الوسيط $(\frac{n}{2})$ سواء كانت n فردية أو زوجية.

٣- نحدد مكان الوسيط بحيث يكون التكرار السابق له f_1 والتكرار اللاحق له f_2 أكبر من $\frac{n}{2}$ ونأخذ الحد الحقيقى للتكرار السابق على أنه البداية الحقيقية للفئة الوسيطية ونرمز له بالرمز A، ونعين طول الفئة الوسيطية ويساوى الحد الأدنى للفئة التالية مطروحا منه الحد الأدنى للفئة الوسيطية ونرمز له بالرمز L ويعطى الوسيط بالعلاقة:

$$\text{Med} = A + \frac{(\frac{n}{2} - f_1)}{f_2 - f_1} L$$

حيث f_1 يمثل التكرار المتجمع الصاعد السابق للتكرار المتجمع الوسيطي، f_2 يمثل التكرار المتجمع الصاعد اللاحق للتكرار المتجمع الوسيطي.

مثال (٦): احسب الوسيط لأعمار الطلاب فى المثال (٢) السابق:

الحل: نكون جدول التكرار المتحمع الصاعد كالآتي:

التكرار المتجمع الصاعد	فئات التكرار المتجمع الصاعد
0	< 4.5
2	< 6.5
7	< 8.5
15	< 10.5
19	< 12.5
20	< 14.5

نحسب $\frac{n}{2}$ وهى تساوى $(\frac{20}{2} = 10)$ ونلاحظ أن 10 تقع بين 7 ,15 فنضع خطًا أفقيًا يمثل تكرار الوسيط المتجمع 10 وعليه فيكون:

$$A = 8.5, \quad f_1 = 7, \qquad f_2 = 15, \qquad L = 10.5 - 8.5 = 2$$

وبتطبيق قانون الوسيط الحسابى نحصل على:

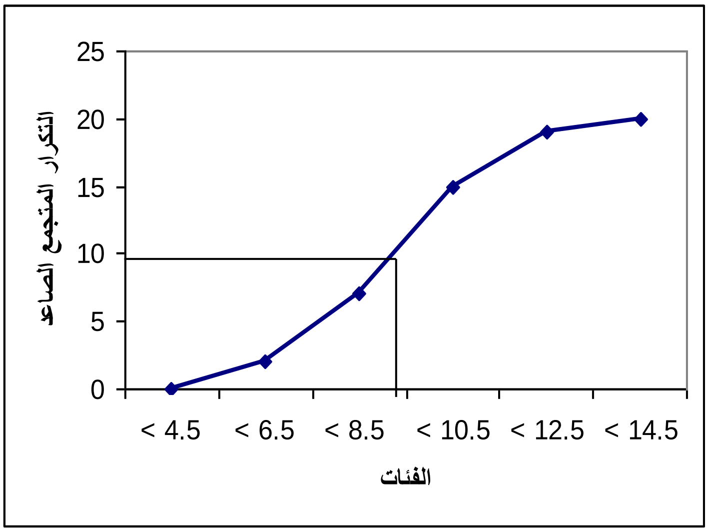
$$\text{Med} = 8.5 + \frac{10-7}{15-7} \cdot 2 = 9.25,$$ سنة

الوسيط بيانيا

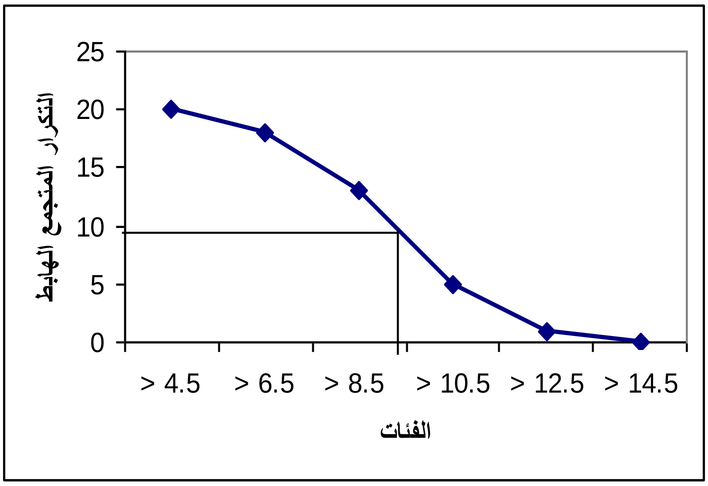
يمكن إيجاد الوسيط بيانيا من المنحنى المتجمع الصاعد أو المنحنى المتجمع الهابط كل على حده أو تقاطع المنحنيين الصاعد والهابط معا فى رسم واحد. فى حالة المنحنى المتجمع الصاعد نحدد نقطة $\frac{n}{2}$ على المحور الرأسى للتكرارات ونرسم منها خطا أفقيا موازيا لمحور الفئات إلى أن يلتقى بالمنحنى فى نقطة، نسقط من تلك النقطة عمودا رأسيا يلاقى محور الفئات فى نقطة تكون قيمتها هى قيمة الوسيط بيانيا. أما فى حالة المنحنى المتجمع الهابط نتبع نفس الخطوات السابقة للمتجمع الصاعد لتحديد قيمة الوسيط بيانيا. أما فى حالة تقاطع المنحنيين الصاعد والهابط فى نقطة فنقوم بإسقاط عمود رأسى على محور الفئات تكون نقطة تقاطعه مع محور الفئات هى قيمة الوسيط. وسوف نوضح ذلك من خلال المثال التالي:

مثال (٧): احسب الوسيط بيانيا من مثال (٢) لأعمار الطلاب.

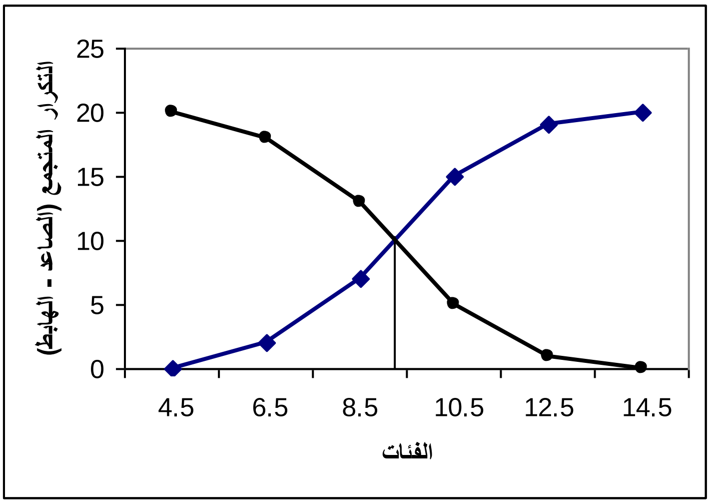
الحل: أولا: من الجدول المتجمع الصاعد فى مثال (٦) يمكن رسم المنحنى المتجمع الصاعد كالآتي:



نرسم من هذا الجدول المنحنى الهابط كالتالى ونعين منه الوسيط بيانيا:



نرسم المنحنيين الصاعد والهابط السابقين كالتالي:



مميزات الوسيط:

١- لايتأثر بالقيم المتطرفة.

٢- يمكن حساب الوسيط فى حالة الجداول التكرارية المفتوحة للبيانات الكمية.

٣- يمكن إيجاده فى حالة البيانات الوصفية التى يمكن ترتيبها.

٤- مجموع الانحرافات المطلقة عن الوسيط أقل ما يمكن مقارنة بأى قيمة حقيقية a، أى أن:

$$\sum_{i=1}^{n}|x_i - Me| \le \sum_{i=1}^{n}|x_i - a|$$

حيث: $a \neq Me$

<u>عيوب الوسيط:</u>

١- لا يأخذ جميع القيم فى الاعتبار عند حسابه.

٢- لا يسهل التعامل معه فى التحاليل الإحصائية والرياضية.

المنوال

يعرف المنوال على أنه القيمه الأكثر شيوعا (تكرارا) فى مجموعة البيانات. وقد يكون لمجموعة البيانات منوال واحد ولذلك يطلق عليها وحيدة المنوال، أو يكون لها أكثر من منوال وتسمى متعددة المنوال. وقد لا يكون لمجموعة البيانات أى منوال وبذلك تسمى عديمة المنوال.

مثال (٨): احسب المنوال من البيانات التالية:

2, 6, 9, 4, 6, 10, 6

الحل: يوجد لهذة البيانات منوال واحد وهو القيمة 6 لأنها تكررت ثلاث مرات أكثر من غيرها.

مثال (٩): احسب المنوال من البيانات التالية:

4, 3, 7, 9, 4, 4, 7, 4

الحل: يوجد لهذة البيانات منوال واحد وهو القيمة 4 لأنها تكررت أربع مرات أكثر من أى قيمة أخرى.

مثال (١٠): احسب المنوال من البيانات التالية:

5, 7, 5, 7, 8, 9, 7, 5, 10

الحل: نجد من خلال هذه البيانات أن القيمة 5 تكررت ثلاث مرات والقيمة 7 تكررت ثلاث مرات وعليه فإن هذه البيانات يوجد لها منوالان هما 5,7.

مثال (١١): احسب المنوال من البيانات التالية:

4, 9, 8, 12, 11, 7, 15

الحل: لا يوجد فى هذه البيانات أى قيمة تكررت أكثر من مرة وعليه فإنه لا يوجد منوال لهذه البيانات.

المنوال فى حالة البيانات المبوبة (الجداول التكرارية)

فى حالة البيانات المبوبة أو الجداول التكرارية لا يمكن القول بأن قيمة معينة يكون لها أكبر تكرار لأن القيم تذوب داخل الفئات المختلفة، ولذلك يمكن القول بأن هناك فئات منوالية وهى الفئات التى يقابلها أعلى تكرار، وفى حالة تساوى تكرار الفئة السابقة للفئة المنوالية مع تكرار الفئة اللاحقة للفئة المنوالية فإنه يمكن حساب قيمة المنوال بمركز الفئة المنوالية أى فى منتصفها، وفى حالة عدم تساويهما فى التكرار فإنه يمكن حساب المنوال كالآتي:

١- نوجد أكبر تكرار f وعليه يمكن إيجاد التكرار السابق له وهو f_1 والتكرار اللاحق له وهو f_2.

٢- نأخذ بداية الفئة المنوالية ويرمز لها بالرمز A وهى الفئة التى يقابلها أعلى تكرار f.

٣- نحدد طول الفئة المنوالية L وهو يساوى الفرق بين بداية الفئة المنوالية وبداية الفئة التالية لها ويتم تطبيق القانون التالي:

$$Mod = A + \frac{f - f_1}{2f - f_1 - f_2} \cdot L$$

مثال (١٢): اوجد المنوال لأعمار الطلاب فى مثال (٢).

الحل: الجدول التالى يوضح أعمار الطلاب كما يلي:

فئات الأعمار	5 – 6	7 – 8	9 – 10	11 – 12	13 – 14
التكرار	2	5	8	4	1

من الجدول نجد أن:

$f = 8, \quad f_1 = 5, \quad f_2 = 4, \quad A = 8.5$

$L = 10.5 - 8.5 = 2$

وبالتعويض فى القانون ينتج أن:

$$Mod = 8.5 + \frac{8-5}{16-5-4} \cdot 2$$

$$Mod = 9.36$$ **سنة**

<u>مميزات المنوال:</u>

١- مقياس سهل حسابه ولا يتأثر بالقيم الشاذة.

٢- يمكن إيجاده للقيم الوصفية والتوزيعات التكرارية المفتوحة.

<u>عيوب المنوال:</u>

١- عند حساب المنوال لا تؤخذ جميع قيم البيانات فى الاعتبار.

٢- قد يكون لبعض البيانات أكثر من منوال وبذلك لا يمكن تحديد قيمة وحيدة للمنوال.

<u>العلاقة بين الوسط الحسابى والوسيط والمنوال</u>

توجد علاقة تجريبية بين المقاييس الثلاثة (الوسط الحسابى – الوسيط – المنوال) وذلك فى حالة التوزيعات التكرارية أحادية المنوال وغير المتماثلة والمتماثلة وذات الالتواء البسيط. وتعطى هذه العلاقة من خلال المعادلة التالية:

الوسط الحسابى – المنوال = ٣ (الوسط الحسابى – الوسيط) (٩)

وقد وجد أن الوسيط تقع قيمته بين قيمتى الوسط الحسابى والمنوال.

وفى حالة التوزيعات التكرارية المتماثلة الوحيدة المنوال فإن قيمة الوسط الحسابى تكون مساوية لقيمة الوسيط تكون مساوية لقيمة المنوال أى أن:

الوسط الحسابى = الوسيط = المنوال

أما فى حالة التوزيعات التكرارية الملتوية (إلتواء موجب – التواء سالب) فإن العلاقة (٩) تكون غير صحيحة.

الوسط الهندسى

الوسط الهندسى .G.M لمجموعة من القيم $x_1, x_2, ..., x_n$ هو الجذر النونى لحاصل ضرب هذه القيم: $G.M.= \sqrt[n]{x_1 x_2 ... x_n}$، يمتاز الوسط الهندسى عن الوسط الحسابى بأنه أقل تأثرا بالقيم الشاذة فى البيانات لأنه معلوم رياضيا بأن الوسط الهندسى لمجموعة من القيم أقل من وسطها الحسابي، وعادة يحسب الوسط الهندسى باستخدام اللوغاريتمات كالتالي:

$$\text{Log G.M.} = \frac{1}{n}\left(\sum_{i=1}^{n} Log(x_i)\right)$$

مثال (١٣): احسب الوسط الحسابى والوسط الهندسى للبيانات

3, 5, 6, 6, 7, 10, 12

الحل:

$$G.M. = \sqrt[7]{3.5.6.6.7.10.12}$$

وباستخدام اللوغاريتمات يكون الوسط الهندسي:

$$\text{Log G.M.} = \frac{1}{7}(\text{Log } 3 + \text{Log } 5 + \text{Log } 6 + \text{Log } 6 + \text{Log } 7 + \text{Log } 10 + \text{Log } 12)$$

$$= \frac{1}{7}(0.4771 + 0.699 + 0.7782 + 0.7782 + 0.8451 + 1 + 1.0729)$$

$$= 0.8081$$

$$G.M. = 6.43$$

$$\bar{x} = \frac{1}{7}(3 + 5 + 6 + 6 + 7 + 10 + 12) = 7$$

ونلاحظ أن قيمة الوسط الحسابي $\bar{x}$ من المثال أكبر من الوسط الهندسى .G.M وهذا يوضح الحقيقة بأن الوسط الهندسى لمجموعة أرقام موجبة غير متساوية أقل من وسطها الحسابي.

الوسط الهندسى فى حالة البيانات المبوبة (الجداول التكرارية)

فى هذه الحالة يحسب الوسط الهندسى للفئات التى عددها k ومراكزها هى $x_1, x_2, ..., x_k$ والتى يقابلها بالترتيب تكرارات $f_1, f_2, ..., f_k$ فإن الوسط الهندسى يعطى بالعلاقة التالية:

$$G.M. = \sqrt[n]{x_1^{f_1} x_2^{f_2} ... x_n^{f_n}}$$

حيث: $n = \sum_{i=1}^{k} f_i$

الربيعات والعشيرات والمئينات

إذا رتبت عينة من البيانات حسب قيمتها تصاعديا أو تنازليا فإن القراءة التى تكون فى المنتصف والتى تقسم العينة إلى مجموعتين متساويتين فى العدد هى الوسيط كما سبق تعريفة. وبتعميم الفكرة وتقسيم البيانات بعد ترتيبها إلى أربعة أجزاء متساوية ويرمز لها بـ Q_1، Q_2، Q_3 فإن Q_1 يسمى الربيع الأول و Q_2 يسمى الربيع الثانى (الوسيط) و Q_3 يسمى الربيع الثالث. وكذلك يمكن

إيجاد القيم التى تقسم البيانات السابقة بعد ترتيبها إلى عشرة أقسام ونرمز لها بالرمز بـ D_1، D_2، ....، D_9 حيث D_1 يسمى العشير الأول وهو يمثل القيمة التى يسبقها عشر القراءات، D_2 يسمى العشير الثانى وهو يمثل القيمة التى يسبقها ٠.٢ من القراءات وهكذا. كما يمكن إيجاد القيم التى تقسم البيانات السابقة بعد ترتيبها إلى مائة قسم ونرمز لها بالرمز P_1، P_2، ....، P_{99} حيث P_1 يسمى المئين الأول وهو يمثل القيمة التى يسبقها ٠.٠١ من القراءات، P_2 يسمى المئين الثانى وهو يمثل القيمة التى يسبقها ٠.٠٢ من القراءات، وهكذا لباقى المئينات. ويعطى قانون حساب الربيعات والعشيرات والمئينات فى حالة البيانات المبوبة مثل قانون الوسيط السابق (٨) مع استبدال $\frac{n}{2}$ بـ $\frac{n}{4}$ للربيع الأول، $\frac{2n}{4}$ للربيع الثانى وهكذا. كذلك استبدال $\frac{n}{2}$ بـ $\frac{n}{100}$ للمئين الأول، $\frac{2n}{100}$ للمئين الثانى وهكذا. وسوف نوضح ذلك من خلال المثال التالي:

مثال (١٥): احسب لأعمار الطلاب فى مثال (٢) السابق كل من العشير الثانى والمئين التسعين.

الحل: يمكن تكوين الجدول المتجمع الصاعد لأعمار الطلاب كالتالي:

حدود الفئات	التكرار المتجمع الصاعد
< 4.5	0
< 6.5	2
D_2	
< 8.5	7
< 10.5	15
P_{90}	
< 12.5	19
< 14.5	20

لإيجاد العشير الثانى D_2 نستخدم القانون:

$$D_2 = A + \frac{(\frac{2n}{10} - f_1)}{f_2 - f_1} L$$

حيث A بداية الفئة للعشير الثاني، n مجموع التكرارات، f_1 التكرار السابق، f_2 التكرار اللاحق، L طول فئة العشير الثاني:

$$D_2 = 6.5 + \frac{4-2}{7-2}.2 = 7.3$$

لإيجاد المئين التسعين P_{90} نستخدم القانون:

$$P_{90} = A' + \frac{(\frac{90n}{100} - f'_1)}{f'_2 - f'_1} L$$

$$P_{90} = 10.5 + \frac{18-15}{19-15}2 = 12$$

تمارين

١- إذا كان لدينا المتغيران x، y لهما القيم التالية:

x	1	2	4	-2	5
y	-1	1	5	-7	7

احسب قيمة كل مما يأتي:

i) $\sum x,\ \sum y,\ \sum xy,\ \sum x^2,\ \sum y^2,\ \sum x^3,\ \sum y^3,\ \sum xy^2$

ii) $(\sum x)^2,\ ((\sum x)+2)^2,\ \sum (x-3y)^2,\ \sum x(x+4)$

٢- أى من العلاقات الآتية صحيحة:

(i) $\sum \left(\frac{y}{x}\right) = \frac{\sum y}{\sum x}$, ii) $\sum xy = \sum x \sum y$

(iii) $\sum (x-c)(x+c) = \sum x^2 - nc^2$, (iv) $(\sum x)^2 = \sum x^2$

٣- فيما يلى أعمار مجموعة من الطلاب بإحدى المدارس الابتدائية:

6,6,9,8,6,10,9,9,8,7,8,6,7,8,8,11,10,11,8,8

١- احسب الوسط الحسابى لأعمار هؤلاء الطلاب.

٢- اوجد الوسيط لأعمار هؤلاء الطلاب.

٣- اوجد المنوال لأعمار هؤلاء الطلاب.

٤- ما قيمة المقاييس الثلاثة بعد ثلاث سنوات بفرض بقائهم جميعا على قيد الحياة؟

٤- عند فحص مجموعة من الأرقام يتكون كل منها من رقم واحد، كانت البيانات على النحو التالي:

التكرار	الرقم
8	2
10	3
20	5
20	7
6	8
6	9

احسب الوسط الحسابى والوسيط والمنوال لهذه البيانات.

٥- فيما يلى توزيع درجات ٦٠ طالب فى أحد الامتحانات:

90-99	80-89	70-79	60-69	50-59	40-49	الفئات
4	6	17	17	10	6	عدد الطلاب

احسب الوسط الحسابى والوسيط والمنوال لدرجات الطلاب.

٦- فيما يلى الأجر اليومى لعدد من العمال بالجنيه فى أحد المصانع:

فئات الأجر	20-29	30-39	40-49	50-59	60-69	70-79
عدد العمال	9	12	15	8	4	2

١ – احسب الوسط الحسابى لأجور العمال.

٢ – اوجد الوسيط والمنوال حسابيا وبيانيا.

٣– إذا كان الأجر اليومى لكل عامل يزيد بمقدار خمسة جنيهات كل ستة أشهر، فما قيمة المقاييس السابقة بعد سنة؟

٤ – اوجد الوسط الهندسى والوسط التوافقى للأجور.

٧– اوجد الوسط الحسابى و الوسط الهندسى والوسط التوافقى لمجموعة الأرقام:

0, 2, 4, 6

٨– احسب للتمرين رقم (٦) كل من:

١ – الربيع الأول و الربيع الثالث.

٢ – العشير الرابع والعشير السابع.

٣– المئين ٣٣ و المئين ٧٥.

الفصل الخامس

بعض مقاييس التشتت

الهدف من دراسة الفصل

بنهاية الدراسة لهذا الفصل يكون الدارس قادر على الالمام بالمفاهيم التالية:

١- المدى , مميزاته و عيوبه .

٢- نصف المدى الربيعى , مميزاته و عيوبه .

٣- طريقة حساب الوسط , الوسيط , المنوال ,,,,, الخ للمجتمع .

٤- نصف المدى الربيعى للبيانات المباشرة

٥- نصف المدى الربيعى للبيانات المبوبة

الفصل الخامس

بعض مقاييس التشتت

مقدمة

لقد سبق لنا دراسة طرق عرض البيانات جدوليا وبيانيا والتعرف على أشكالها وتوزيعاتها المختلفة وكذلك دراسة مقاييس النزعة المركزية (المتوسطات) وذلك لوصف البيانات عدديا لهذه التوزيعات المختلفة. ولكن طرق عرض البيانات وحساب المتوسطات للمجموعات المختلفة من البيانات غير كافٍ للمقارنة بين هذه المجموعات. ولتوضيح ذلك نأتى بمثال لدراسة ثلاث مجموعات مختلفة من الطلاب X، Y، Z وكانت الدرجات كالآتي:

X	**59,61,62,58,60**
Y	**50,60,66,54,70**
Z	**39,65,46,78,72**

وبحساب الوسط الحسابى للثلاث مجموعات نجدة يساوى ٦٠ درجة لكل منها. ولكن عند النظر لدرجات المجموعة الأولى نجدها متقاربة، ودرجات المجموعة الثانية أقل تقاربا من المجموعة الأولى، ودرجات المجموعة الثالثة أقل تقاربا من درجات المجموعة الثانية. أى أن الثلاث مجموعات مختلفة التجانس رغم أن الوسط الحسابى لهم متساو. وبذلك تكون مقاييس النزعة المركزية غير كافية للمقارنة بين طبيعة البيانات الإحصائية. لذلك نشأت الحاجة إلى إيجاد مقاييس تقيس درجة تجانس (تقارب) أو تشتت (تباعد) مفردات البيانات عن بعضها البعض، وتعرف هذه المقاييس بمقاييس التشتت **Measures of Dispersion**. وسوف نستعرض منها كل من المدى، نصف المدى الربيعي، الانحراف المتوسط، التباين والانحراف المعياري، معامل الاختلاف ومقاييس الالتواء والتفرطح. وسوف نتناول كل منها بالتفصيل والأمثلة كل على حدة.

المدى

يعرف المدى بأنه الفرق بين أكبر قراءة وأصغر قراءة فى مجموعة القراءات أى أن:

المدى **R** = أكبر قيمة – أصغر قيمة

وذلك فى حالة البيانات المباشرة. أما فى حالة البيانات المبوبة فإن المدى يعرف بأكثر من طريقة، نذكر منها الطريقتين الآتيتين:

١- المدى = الفرق بين مركزى الفئة العليا والفئة الدنيا.

٢- المدى = الحد الأعلى للفئة العليا مطروحا منه الحد الأدنى للفئة الدنيا.

وسوف نوضح ذلك من خلال الأمثلة التالية:

مثال (١): احسب المدى R لدرجات الطلاب الآتية:

82, 40, 62, 70, 30, 80

الحل: أكبر قيمة = ٨٢ درجة، أصغر قيمة = ٣٠ درجة

$R = 82 - 30 = 52$ درجة

مثال (٢): اوجد المدى R لدرجات مجموعة من الطلاب معطاة بالجدول التالي:

90-99	80-89	70-79	60-69	50-59	40-49	الفئات
1	2	11	15	9	2	عدد الطلاب

الحل: يمكن إيجاد المدى بطريقتين:

الطريقة الأولى: نقوم بحساب مركز الفئة العليا = ٩٤.٥ ومركز الفئة الدنيا = ٤٤.٥

$R = 94.5 - 44.5 = 50$ درجة

الطريقة الثانية: الحد الأعلى للفئة العليا الحقيقى = ٩٩.٥

الحد الأدنى للفئة الدنيا الحقيقى = ٣٩.٥

$R = 99.5 - 39.5 = 60$ درجة

ونلاحظ اختلاف كل من الطريقتين فى حساب قيمة المدى. وغالبا ما تستخدم الطريقة الأولى فى إيجاد المدى.

مميزات المدى:

١- سهل الحساب.
٢- يعطى فكرة سريعة عن طبيعة البيانات ويستخدم كثيرا فى مراقبة جودة الإنتاج وكذلك فى وصف طبيعة الأحوال الجوية.

عيوب المدى:

١- يعتمد فى حسابه على قيمتين فقط من البيانات مع إهمال باقى القيم.
٢- يتأثر كثيرا بالقيم الشاذة (المتطرفة) لذلك فهو مقياس تقريبى لا يعتمد عليه.

نصف المدى الربيعي

من أهم عيوب المدى أنه يتأثر بالقيم الشاذة وبالتالى فهو لا يعطى صورة صادقة عن طبيعة البيانات. لذلك دعت الحاجة إلى إيجاد مقياس آخر يتم من خلاله التخلص من تأثير القيم الشاذة وهو ما يسمى بنصف المدى الربيعي. ويعرف كما يلي:

إذا كانت لدينا مجموعة من البيانات عددها n قيمة فيتم ترتيب القيم تصاعديا وتقسم إلى أربعة أقسام متساوية كما هو موضح على الخط الأفقى التالي:

$$\_\_\_\_\_\_\_\_\_\_ Q_3 \_\_\_\_\_\_\_\_\_\_ Q_2 \; Q_1$$

تسمى القيمة التى يسبقها ربع البيانات بالربيع الأول ويرمز له بالرمز Q_1 ورتبته $\frac{n}{4}$، وتسمى القيمة التى يسبقها ثلاثة أرباع البيانات بالربيع الثالث ويرمز له بالرمز Q_3 ورتبته $\frac{3n}{4}$ كما يسمى المقدار الناتج من الفرق بين Q_1، Q_3 بالمدى الربيعى وهو يمثل النصف الأوسط للقيم، ويؤخذ نصف هذا المدى مقياسا للتشتت ويسمى بنصف المدى الربيعى ويرمز له بالرمز Q ويعطى من خلال العلاقة:

$$Q = \frac{Q_3 - Q_1}{2}, \qquad (1)$$

ويلاحظ أن Q_2 هو الربيع الثانى وهو القيمة التى يسبقها نصف البيانات ورتبته $\frac{n}{2}$ أى أن Q_2 هو الوسيط الذى سبق شرحه فى مقاييس النزعة المركزية. وسوف نتناول بالشرح نصف المدى الربيعى Q فى كل من البيانات المباشرة والبيانات المبوبة كالتالي.

<u>نصف المدى الربيعى للبيانات المباشرة</u>

يحسب من خلال الخطوات التالية:

١- ترتب البيانات تصاعديا.

٢- نوجد قيمة Q_1 وهى القيمة التى يسبقها ربع البيانات.

٣- نوجد قيمة Q_3 وهى القيمة التى يسبقها ثلاثة أرباع البيانات.

٤- يتم تطبيق القانون (١) لحساب نصف المدى الربيعي.

مثال (٣): اوجد نصف المدى الربيعى لأوزان مجموعة الطلاب التالية:

67,65,69,58,55,71,72,70

الحل: ترتب البيانات تصاعديا لنحصل على:

55,58,65,67,69,70,71,72

$Q_1 = \frac{58+65}{2} = 61.5,\qquad Q_3 = \frac{70+71}{2} = 70.5$

$Q = \frac{Q_3 - Q_1}{2} = 4.5$

مثال (٤): اوجد نصف المدى الربيعى لأوزان مجموعة الطلاب التالية:

59,67,65,69,58,55,70,72,74

الحل: ترتب البيانات تصاعديا لنحصل على:

55,58,59,65,67,69,70,72,74

$Q_1 = 59,$

$Q_3 = 70$

$$Q = \frac{Q_3 - Q_1}{2} = 5.5$$

نصف المدى الربيعى للبيانات المبوبة

يحسب نصف المدى الربيعى لهذه البيانات بنفس الطريقة التى سبق شرحها لحساب الوسيط وهو طريقة الفروق ويحسب الربيع الأول Q_1 بوضع $\frac{n}{4}$ بدلا من $\frac{n}{2}$ فى قانون الوسيط ويحسب الربيع الثالث Q_3 بوضع $\frac{3n}{4}$ بدلا من $\frac{n}{2}$ فى قانون الوسيط وبعد ذلك نحسب نصف المدى الربيعى من القانون (١) ويوضح طريقة حساب Q_1 و Q_3 بالعلاقتين الآتيتين، أى أن:

$$Q_1 = A_1 + \frac{(\frac{n}{4} - f_1)}{f_2 - f_1} L, \qquad (2)$$

$$Q_3 = A_2 + \frac{(\frac{3n}{4} - f_1')}{f_2' - f_1'} L, \qquad (3)$$

مثال (٥): اوجد نصف المدى الربيعى حسابيا لدرجات الطلاب فى مثال (٢).

الحل: نكون جدول التكرار المتجمع الصاعد كما يلي:

حدود الفئات	التكرار المتجمع الصاعد
< 39.5	0
< 49.5	2
< 59.5	11
< 69.5	26
< 79.5	37
< 89.5	39
< 99.5	40

$n = 40,\quad \frac{n}{4} = 10,\quad \frac{3n}{4} = 30,\quad L = 10$

$$Q_1 = 49.5 + \left(\frac{10-2}{11-2}\right) 10$$

$$Q_1 = 49.5 + 8.89 = 58.39$$

$$Q_3 = 69.5 + \left(\frac{30-26}{37-26}\right) 10$$

$$Q_3 = 69.5 + 3.64 = 73.14$$

$$Q = \frac{Q_3 - Q_1}{2} = 7.38$$

مميزات نصف المدى الربيعي:

١- يتخلص من القيم الشاذة نحو الكبر أو الصغر.

٢- يمكن حسابه من التوزيعات التكرارية المفتوحة من الطرفين.

عيوب نصف المدى الربيعي:

١- لا يأخذ جميع القيم فى الاعتبار.

٢- لا يسهل التعامل معه فى التحليل الإحصائي.

٤-٣ الانحراف المتوسط

يعرف الانحراف المتوسط بأنه متوسط الانحرافات المطلقة للبيانات عن وسطها الحسابى x ويرمز له بالرمز .M.D ويعرف رياضيا كالتالي:

$$\text{M.D.}= \frac{1}{n}\sum_{i=1}^{n}|x_i - x| , \qquad (4)$$

والسبب فى أخذ القيم المطلقة للانحرافات هو أن مجموع انحرافات القيم عن وسطها الحسابى يساوى صفر (برهن ذلك). والتعريف السابق (٤) يكون بالنسبة للبيانات المباشرة أما فى حالة البيانات المبوبة يعطى الانحراف المتوسط من خلال العلاقة الآتية:

$$\text{M.D.}= \frac{\sum_{i=1}^{n} f_i |x_i - x|}{n} , \qquad (5)$$

حيث n هى مجموع التكرارات

ونوضح ذلك بالأمثلة التالية.

مثال (٦): اوجد الانحراف المتوسط لأعمار مجموعة الطلاب

6,5,7,7,8,9,9,5

الحل: نكون جدول الحل كالتالي:

x	$x-x$	$\lvert x-x \rvert$
6	-1	1
5	-2	2
7	0	0
7	0	0
8	1	1
9	2	2
9	2	2
5	-2	2
56	0	10

$$x = \frac{\sum x}{n}$$

$\bar{x} = \frac{56}{8} = 7$

$M.D. = \frac{10}{8} = 1.25,$ سنة

مثال (٧): اوجد الانحراف المتوسط لدرجات الطلاب فى مثال (٢).

الحل: نكون جدول الحل كالتالي:

Classes	x	f	fx	$x-\bar{x}$	$\lvert x-\bar{x}\rvert$	$\lvert x-\bar{x}\rvert f$
40-49	44.5	2	89	-21.25	21.25	42.5
50-59	54.5	9	490.5	-11.25	11.25	101.25
60-69	64.5	15	967.5	-1.25	1.25	18.75
70-79	74.5	11	819.5	8.75	8.75	96.75
80-89	84.5	2	169	18.75	18.75	37.5
90-99	94.5	1	94.5	28.75	28.75	28.75
		40	2630			325

$\bar{x} = \frac{\sum fx}{n}$

$\bar{x} = \frac{2630}{40} = 65.75$

$M.D. = \frac{325}{40} = 8.125,$ درجة

ملاحظة: أحيانا يعرف الانحراف المتوسط باستخدام الوسيط بدلا من الوسط الحسابى أو أى متوسطات أخرى غير الوسط الحسابي.

٤-٤ الانحراف المعيارى

من الصعب التعامل رياضيا (تحليليا) مع الانحراف المتوسط، ولذلك دعت الحاجة إلى استخدام مقياس للتشتت بنفس قوة الانحراف المتوسط، ولكى يكون من السهل التعامل معه تحليليًا، وبما أن الفكرة هى التخلص من الإشارات للانحرافات فإن تربيع الانحرافات يخلصنا من الإشارة. ولهذا فإن الانحراف المعيارى يعرف عن طريق التباين والذى يعرف على أنه متوسط مربع انحرافات القيم عن وسطها الحسابى ويرمز له بالرمز σ^2، والجذر التربيعى للتباين ينتج عنه مقياس من أهم وأدق مقاييس التشتت وهو ما يسمى بالانحراف المعيارى ويرمز له بالرمز σ، وسوف نتناول طريقة حسابه فى حالة البيانات المباشرة والبيانات المبوبة كما يلي:

الانحراف المعيارى فى حالة البيانات المباشرة

إذا كان لدينا مجموعة من البيانات من مجتمع إحصائى عدد مفرداته N على الصورة $X_1, X_2, \ldots X_N$ ومتوسط هذه البيانات $\bar{X}$ فإن مربعات انحرافات هذه القيم عن وسطها الحسابى يكون على الصورة:

$$(X_1-\bar{X})^2, (X_2-\bar{X})^2, \ldots, (X_N-\bar{X})^2$$

ويعرف التباين σ^2 كالتالى:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^2 , \qquad (6)$$

والانحراف المعيارى هو:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i-\bar{X})^2} , \qquad (7)$$

ويفضل عند حساب الانحراف المعيارى أن يحسب التباين من المعادلة (٦) ويأخذ الجذر التربيعى للنتيجة النهائية لنحصل على المعادلة (٧). أما فى حالة العينة التى حجمها n المأخوذة من المجتمع فإن الانحراف المعيارى فى هذه الحالة يرمز له بالرمز S والتباين S^2 ويعرف بقسمة مجموع مربعات الانحرافات على $(n-1)$ ويكتب كما يلي:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{N}(x_i-\bar{x})^2 , \qquad (8)$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{N}(x_i-\bar{x})^2} , \qquad (9)$$

وذلك فى حالة البيانات المباشرة، و S يعطى تقديرا أفضل للانحراف المعيارى للمجتمع الذى أخذت منه العينة. وإذا كان عدد المفردات كبير (أكبر من ٣٠) فإن قيمة S^2 و σ^2 تكون متساوية تقريبا وذلك من الناحية العملية.

مثال (٨): احسب الانحراف المعيارى لأعمار مجموعة من الطلاب فى المرحلة الابتدائية:

8,9,7,6,5

الحل:

x	$x-\bar{x}$	$(x-\bar{x})^2$
8	1	1
9	2	4
7	0	0
6	-1	1
5	-2	4
35	0	10

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{35}{5} = 7$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

$$S^2 = \frac{1}{5-1}(10) = 2.5$$

$$S = \sqrt{2.5} = 1.581,$$ سنة

ويمكن تبسيط العلاقة (٨) كالتالي:

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} \right), \qquad (10)$$

الإثبات:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2)$$

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2\right)$$

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)$$

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} \right)$$

ويلاحظ أن الصورة (١٠) تحتاج فى الحساب إلى $\sum x$ ، $\sum x^2$ فقط.

مثال (٩): احسب الانحراف المعيارى لأعمار الطلاب فى مثال (٨) باستخدام العلاقة (١٠).

الحل: نكون جدول الحل كما يلي:

x	x^2
8	**64**
9	**81**
7	**49**
6	**36**
5	**25**
35	**255**

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} \right)$$

$$S^2 = \frac{1}{4}\left(255 - \frac{(35)^2}{5}\right) = 2.5$$

$$S = \sqrt{2.5} = 1.581,$$ سنة

وهى نفس النتيجة لمثال (٨) السابق

بعض خصائص الانحراف المعياري

الخاصية الأولى: إذا أضفنا أو طرحنا مقدارًا ثابتًا c من جميع القراءات لمجموعة البيانات فإن الانحراف المعيارى للقيم الجديدة هو نفسه الانحراف المعيارى للقيم الأصلية ويمكن إثبات ذلك بالآتي:

نفرض أن القيم الأصلية هى $x_1, x_2, ..x_n$ ونفرض أن القيم الجديدة هى $d_1, d_2, ..d_n$ حيث:

$d_1 = x_1 + c$, $d_2 = x_2 + c, \ldots, d_n = x_n + c$

$$\because S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} [(d_i \pm c) - (\bar{d} \pm c)]^2 \right) = \frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2, \qquad (11)$$

ويمكن أن تستخدم هذه الخاصية فى تبسيط البيانات، وخاصة عندما تكون قيمتها كبيرة كما يوضح ذلك المثال التالي:

مثال (١٠): استخدم الخاصية الأولى فى حل المثال (٨) باختيار الثابت c يساوى ٥.

الحل: نطرح المقدار الثابت من كل القراءات كما هو موضح بالجدول الآتي:

x	$d = x - 5$	d^2
8	3	9
9	4	16
7	2	4
6	1	1
5	0	0
	10	30

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} d_i^2 - \frac{(\sum_{i=1}^{n} d_i)^2}{n} \right)$$

$$S^2 = \frac{1}{4}\left(30 - \frac{(10)^2}{5}\right) = 2.5$$

$$S = \sqrt{2.5} = 1.581,$$ سنة

وهى نفس النتيجة لمثال (٨) السابق.

الخاصية الثانية: إذا ضربنا جميع القيم فى مقدار ثابت أو قسمناها على مقدار ثابت فإن الانحراف المعيارى يتأثر بذلك وسوف نثبت ذلك، فى حالة الضرب هذا، ويمكن اتباع نفس الخطوات فى حالة القسمة كما يلي:

نفرض أن القراءات الأصلية هى $x_1, x_2, ..x_n$ فإذا ضربنا هذه القيم فى مقدار ثابت c فتكون القيم الجديدة على الصورة $d_1, d_2, ..d_n$ حيث:

$$d_1 = cx_1, d_2 = cx_2, ..., d_n = cx_n$$

وعليه فإن:

$$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{d_i}{c} - \frac{\bar{d}}{c}\right)^2 = \frac{1}{n-1}\frac{1}{c^2}\sum_{i=1}^{n}(d_i - \bar{d})^2$$

$$S_x^2 = \frac{1}{c^2}S_d^2, \Rightarrow S_x = \frac{1}{c}S_d, \qquad \textbf{(12)}$$

أى أن الانحراف المعيارى للقيم الأصلية فى حالة الضرب يساوى الانحراف المعيارى للقيم الجديدة مقسوما على المقدار الثابت، كما هو موضح بالعلاقة (١٢) أما فى حالة القسمة فإنه يمكن إثبات أن:

$$S_x = c\,S_d, \qquad \textbf{(13)}$$

أى أن الانحراف المعيارى للقيم الأصلية يساوى الانحراف المعيارى للقيم الجديدة مضروبا فى المقدار الثابت كما هو موضح بالعلاقة (١٣).

الخاصية الثالثة: مجموع مربعات انحرافات القيم عن وسطها الحسابى $\bar{x}$ تكون أصغر من مجموع مربعات انحرافات القيم عن أى وسط فرضى أخر a حيث $\bar{x} \neq a$.

الإثبات:

$$\sum(x-a)^2 = \sum(x + \bar{x} - \bar{x} - a)^2$$
$$= \sum[(x-\bar{x}) + (\bar{x}-a)]^2$$
$$= \sum(x-\bar{x})^2 + n(\bar{x}-a)^2 + 2(\bar{x}-a)\sum(x-\bar{x})$$
$$= \sum(x-\bar{x})^2 + n(\bar{x}-a)^2$$

ونلاحظ أن المقدار $n(\bar{x}-a)^2$ مقدار موجب دائما، ونستنتج من ذلك أن:

$\sum(x-\bar{x})^2 < \sum(x-a)^2$ وهو المطلوب إثباته

الخاصية الرابعة: إذا كانت هناك عينتان مجموع تكرارهما هو n_1 و n_2 وتباينهما S_1^2 و S_2^2 على الترتيب ولهما نفس المتوسط $\bar{x}$ فإن التباين المشترك هو:

$$S^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-1}$$

الإثبات:

نفرض أن المجموعتين هما:

$$x_1, x_2, \ldots x_{n_1} \ , \ y_1, y_2, \ldots y_{n_2}$$

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2$$

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{x})^2$$

$$(n_1 - 1)\, S_1^2 = \sum_{i=1}^{n_1} (x_i - \bar{x})^2$$

$$(n_2 - 1)\, S_2^2 = \sum_{i=1}^{n_2} (y_i - \bar{y})^2$$

$$(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2 = \sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2 = \sum_{i=1}^{n_1+n_2} (z_i - \bar{x})^2$$

$$\therefore S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 1}$$

الخاصية الخامسة: الانحراف المعيارى لمجموعة من البيانات أكبر من الانحراف المتوسط لها (على الطالب التأكد من ذلك).

الانحراف المعيارى للبيانات المبوبة

إذا كان لدينا عدد k من الفئات ذات المراكز $x_1, x_2, \ldots x_k$ ولها التكرارات $f_1, f_2, \ldots f_k$ على الترتيب فإن المعادلات السابقة (٨)، (٩)، (١٠) تصبح كالتالي:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{k} f_i (x_i - \bar{x})^2 \ , \ S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{k} f_i (x_i - \bar{x})^2} \ , \qquad \textbf{(14)}$$

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^{k} f_i x_i^2 - \frac{(\sum_{i=1}^{k} f_i x_i)^2}{n} \right), \qquad \textbf{(15)}$$

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^{k} f_i d_i^2 - \frac{(\sum_{i=1}^{k} f_i d_i)^2}{n} \right), \qquad \textbf{(16)}$$

وسوف نبين طريقة حساب الانحراف المعيارى باستخدام العلاقات السابقة بالأمثلة التالية:

مثال (١١): اوجد الانحراف المعيارى لدرجات الطلاب فى مثال (٢) وذلك باستخدام العلاقات (١٤)، (١٥)، (١٦)

الحل: ١- باستخدام العلاقة (١٤) نكون جدول الحل كالتالي:

Classes	x	f	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
40-49	**44.5**	**2**	**89**	**-21.25**	**451.56**	**903.13**
50-59	**55.5**	**9**	**490.5**	**-11.25**	**126.56**	**1139.06**
60-69	**65.5**	**15**	**967.5**	**-1.25**	**1.56**	**23.44**
70-79	**75.5**	**11**	**819.5**	**8.75**	**76.56**	**842.19**
80-89	**85.5**	**2**	**169**	**18.75**	**351.56**	**703.13**
90-99	**95.5**	**1**	**94.5**	**28.75**	**826.56**	**826.56**
Total		**40**	**2630**			**4437.5**

$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i x_i = \frac{1}{40}$**(2630) = 65.75**

$S^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i (x_i - \bar{x})^2 = \frac{1}{40-1}$**(4437.5) = 113.78**

S**= 10.67,** درجة

٢- باستخدام العلاقة (١٥) نكون جدول الحل كالتالي:

Classes	x	f	xf	$x^2 f$
40-49	**44.5**	**2**	**89**	**3960.5**
50-59	**55.5**	**9**	**490.5**	**26732.25**
60-69	**65.5**	**15**	**967.5**	**62403.75**
70-79	**75.5**	**11**	**819.5**	**61052.75**
80-89	**85.5**	**2**	**169**	**12280.5**
90-99	**95.5**	**1**	**94.5**	**8930.25**
Total		**40**	**2630**	**177360**

$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} f_i x_i^2 - \frac{(\sum_{i=1}^{k} f_i x_i)^2}{n} \right) = \frac{1}{40-1}$**(177360 – 172922.5) = 113.78**

S**= 10.67,** درجة (وهى نفس النتيجة السابقة)

٣- باستخدام العلاقة (١٦) وبأخذ المقدار الثابت (الوسط الفرضي) $c = 64.5$ وهو مركز الفئة التى يقابلها أكبر تكرار (وذلك لتبسيط الحسابات) كما هو موضح بجدول الحل التالي:

Classes	x	f	$d = x - 64.5$	df	$d^2 f$
40-49	**44.5**	**2**	**-20**	**-40**	**800**
50-59	**55.5**	**9**	**-10**	**-90**	**900**
60-69	**65.5**	**15**	**0**	**0**	**0**
70-79	**75.5**	**11**	**10**	**110**	**1100**
80-89	**85.5**	**2**	**20**	**40**	**800**
90-99	**95.5**	**1**	**30**	**30**	**900**
Total		**40**		**50**	**4500**

$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{k} f_i d_i^2 - \frac{(\sum_{i=1}^{k} f_i d_i)^2}{n} \right) = \frac{1}{40-1}$**(4500 – 62.5) = 113.78**

S**= 10.67,** درجة

ملاحظة: باستخدام الخاصية الثانية نقوم بحل المثال السابق، ولقد لاحظنا أنه باستخدام الوسط الفرضى وهو مركز الفئة التى يقابلها أعلى تكرار قد بسطت الحسابات كثيرا، هذا ويمكن تبسيط الحسابات أكثر وذلك بقسمة انحرافات القيم عن الوسط الفرضى على طول الفئة (تستخدم هذه الطريقة فى حالة الفئات المنتظمة) وبذلك يكون الحل على النحو التالي:

Classes	x	f	$d=x-65.5$	$d'=d/10$	$d'f$	$d'^2 f$
40-49	44.5	2	-20	-2	-4	8
50-59	55.5	9	-10	-1	-9	9
60-69	65.5	15	0	0	0	0
70-79	75.5	11	10	1	11	11
80-89	85.5	2	20	2	4	8
90-99	95.5	1	30	3	3	9
Total		40			5	45

$$S_{d'}^2=\frac{1}{n-1}\left(\sum_{i=1}^{n} f_i d_i'^2 - \frac{\left(\sum_{i=1}^{n} f_i d_i'\right)^2}{n}\right) = \frac{1}{40-1}(45 - 0.625) = 1.1378$$

$S_{d'}= 1.067$, $S_x=10\, S_{d'}= 10.67$ درجة (وهى نفس النتيجة السابقة)

نلاحظ أن مميزات وعيوب الانحراف المعيارى هى نفس مميزات وعيوب الوسط الحسابى البسيط والتى سبق ذكرها فى الباب الثالث.

المتغير المعيارى والقيم المعيارية

إذا كان لدينا المتغير X والذى له القيم $x_1, x_2, ..., x_n$ والتى لها المتوسط $\bar{x}$ والانحراف المعيارى s فإن المتغير Z الذى يعطى بالعلاقة:

$$Z_i= \frac{x_i-\bar{x}}{s}, \qquad i=1,2,...,n$$

حيث Z_i تقيس الانحرافات عن الوسط الحسابى بوحدات من الانحراف المعيارى يسمى " المتغير المعياري" (القيمة المعيارية).

مثال (١٢): حصل طالب على ٨٢ درجة فى مادة الإحصاء حيث كان متوسط الدرجات هو ٧٥ درجة وذلك بانحراف معيارى ١٠ درجات ثم حصل على ٨٩ درجة فى مادة الرياضيات بمتوسط درجات ٨١ درجة وانحراف معيارى ١٦ درجة. فى من المقررين كانت درجة استيعاب هذا الطالب أعلى؟

الحل: إذا كانت Z_1 تمثل الدرجة المعيارية للإحصاء فإن:

$$Z_1 = \frac{82-75}{10} = 0.7$$

إذا كانت Z_2 تمثل الدرجة المعيارية للإحصاء فإن:

$$Z_2 = \frac{89-81}{16} = 0.5$$

وهذا يدل على أن درجة استيعاب الطالب لمادة الإحصاء أفضل منها لمادة الرياضيات.

تمارين

١- احسب المدى ونصف المدى الربيعى والانحراف المتوسط والانحراف المعيارى من البيانات التالية:

6,3,5,5,9,4,6,7,1,2,4,8

٢- فيما يلى توزيع أوزان ٥٠ طالب من طلاب جامعة القاهرة:

فئات الوزن	58-60	61-63	64-66	67-69	70-72	73-75
عدد الطلاب	2	7	14	15	8	4

اوجد: أ- مدى أوزان الطلاب.

ب-نصف المدى الربيعى للأوزان.

ج-الانحراف المتوسط والانحراف المعياري.

٣- احسب مقاييس الالتواء والتفرطح من البيانات فى المسألة (٢).

٤- الجدول التالى يمثل دخل مجموعة من الأسر بمئات الجنيهات:

فئات الدخل	< 10	10-14	15-19	20-24	25-29	$30 \leq$
عدد الأسر	5	20	35	19	13	8

أى من المقاييس التالية يمكن إيجادها وأى منها لا يمكن إيجادها مع ذكر السبب؟

المدى – نصف المدى الربيعى – الانحراف المعياري.

٥- اوجد العزم الأول والثانى والثالث والرابع لمجموعة البيانات:

2, 5, 9, 4, 3, 6

٦- عند دراسة أطوال مجموعة من الأطفال حديثى الولادة كانت أطوالهم كالتالي:

70, 70, 70, 70, 70, 70, 70

احسب مقاييس التشتت لهذه الأطوال.

الفصل السادس

العلاقات الاحصائية

<u>الهدف من دراسة الفصل</u>

<u>بنهاية الدراسة لهذا الفصل يكون الدارس قادر على الالمام بالمفاهيم التالية:</u>

١ـ معامل الارتباط الخطى لبيرسون وكيفية حسابه.

٢ـ معامل ارتباط الرتب (سبيرمان) وكيفية حسابه.

٣ـ معامل الاقتران ومعامل التوافق وكيفية حسابه.

الفصل السادس

العلاقات الإحصائية

(الارتباط والانحدار)

مقدمة

تناولنا فى الفصول السابقة طرق دراسة متغير واحد لأى ظاهرة محل الدراسة، مثل أوزان مجموعة من الطلاب أو أجور مجموعة من العمال... إلخ. وعرضنا كيف يمكن تلخيص البيانات فى جداول توزيعات تكرارية وكيفية عرضها بيانيا. كذلك دراسة بعض المقاييس العددية التى تساعد على معرفة بعض خصائص التوزيعات التكرارية، ومنها مقاييس النزعة المركزية ومقاييس التشتت والالتواء والتفرطح.

سوف نتناول الآن دراسة البيانات التى يكون لأفرادها متغيران يتغيران معا فى وقت واحد، وذلك لمعرفة نوع العلاقة التى تربط بينهما، مثل دراسة العلاقة بين أوزان وأطوال مجموعة من الطلاب أو أعمار ودرجات مجموعة من الطلاب، وهكذا.... ثم إيجاد مقاييس تقيس درجة هذه العلاقة.

كذلك سوف نقوم بدراسة العلاقة بين المتغيرين (x,y)، فإذا كانت هناك علاقة بين المتغير x والمتغير y فكيف يمكن التعبير عنها بمعادلة رياضية ومنها يمكن التنبؤ بقيمة أحد المتغيرين إذا علمت قيمة المتغير الآخر. وسوف نتناول فى هذا الباب إيجاد مقاييس لقياس قوة الارتباط بين المتغيرين(x,y) فى الحالة الخطية فقط. وسندرس منها معامل الارتباط الخطى لبيرسون **(Pearson)**، ومعامل ارتباط الرتب لسبيرمان (**Spearman**)، كما سوف ندرس معامل الاقتران ومعامل التوافق، وكذلك دراسة معادلة الانحدار الخطى البسيط.

معامل الارتباط الخطى لبيرسون

يستخدم معامل الارتباط الخطى لبيرسون لقياس التغير الذى يطرأ على المتغير y عندما تتغير قيم x أو العكس. ويستخدم عادة فى حالة البيانات الكمية.

إذا كان لدينا أزواج المشاهدات التالية:

$$(x_n, y_n), \ldots, (x_2, y_2), (x_1, y_1)$$

فإن معامل الارتباط r لبيرسون يعطى من خلال العلاقة:

$$r = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \quad , \qquad (1)$$

ويكون لمعامل الارتباط (r) الخصائص التالية:

١- قيمته تساوى صفرا عندما تكون الظاهرتان مستقلتان تماما.

٢- قيمته مقدار موجب عندما يكون الارتباط بين المتغيرين طرديًا. ويكون قويا عندما يكون المقدار الموجب قريبا من الواحد الصحيح، وضعيفا عندما يكون المقدار الموجب قريبا من الصفر.

٣- قيمته مقدار سالب عندما يكون الارتباط بين المتغيرين عكسيا. ويكون قويا عندما يكون المقدار السالب قريبًا من (-١)، وضعيفًا عندما يكون المقدار السالب قريبًا من الصفر.

مثال (١): الجدول التالى يوضح درجات مجموعة مكونة من ثمانية طلاب فى كل من مادتى الإحصاء والرياضيات فى إحدى الامتحانات. هل هناك علاقة بين تحصيل الطالب فى المادتين؟

الإحصاء x	13	9	19	15	11	8	16	11
الرياضيات y	15	7	17	15	10	9	14	10

الحل: لتبسيط البيانات بالجدول نطرح مقدار ثابت = ١٠ من كل من قيم x وقيم y ونكون الجدول التالي:

x	y	$x = x$ 10 -	$y = y$ 10-	$y\,x$	x^2	y^2
13	15	3	5	15	9	25
9	7	-1	-3	3	1	9
19	17	9	7	63	81	49
15	15	5	5	25	25	25
11	10	1	0	0	1	0
8	9	-2	-1	2	4	1
16	14	6	4	24	36	16
11	10	1	0	0	1	0
		22	17	132	158	125

$$r=\frac{\sum xy-\frac{(\sum x)(\sum y)}{n}}{\sqrt{(\sum x^2-\frac{(\sum x)^2}{n})(\sum y^2-\frac{(\sum y)^2}{n})}}$$

$$r=\frac{132-\frac{(22\times 17)}{8}}{\sqrt{(158-\frac{(22)^2}{8})(125-\frac{(17)^2}{8})}}$$

$r=$**0.93**

أى يوجد ارتباط طردى (قوى جدا) بين درجات تحصيل الطالب فى المادتين.

مثال (٢): البيانات التالية توضح العلاقة بين قيمة الاستهلاك y والدخل x

الاستهلاك (y)	٢	٣	٤	٥	٦	١٠
الدخل (x)	٣	٥	٦	٨	٩	١١

والمطلوب: حساب معامل الارتباط بين الاستهلاك والدخل، وما هو مدلوله ؟

حساب المجاميع:

x	y	x^2	y^2	xy
3	2	9	4	6
5	3	25	9	15
6	4	36	16	24
8	5	64	25	40
9	6	81	36	54
11	10	121	100	110
4	30	336	190	249

$\sum x=42\ ,\ \sum y=30$

$\sum x^2=336$

$\sum y^2=190$

$\sum xy=249$

يوجد ارتباط طردي قوي بين الاستهلاك والدخل.

- **حساب معامل الارتباط:**

باستخدام المجاميع السابقة، نجد أن معامل الارتباط قيمته هي

$$r=\frac{\sum xy-\frac{\sum x\sum y}{n}}{\sqrt{\left(\sum x^2-\frac{(\sum x)^2}{n}\right)\left(\sum y^2-\frac{(\sum y)^2}{n}\right)}}$$

$$=\frac{249-\frac{(42)(30)}{6}}{\sqrt{\left(336-\frac{(42)^2}{6}\right)\left(190-\frac{(30)^2}{6}\right)}}$$

$$=\frac{39}{\sqrt{(42)(40)}}=\frac{39}{40.9878}=-0.9515$$

•يوجد ارتباط طردي قوي بين الاستهلاك والدخل.

معامل ارتباط الرتب (سبيرمان)

معامل الارتباط الخطى لبيرسون الذى سبق الحديث عنه يقيس مقدار قوة الارتباط بين متغيرين وذلك فى حالة البيانات الكمية. لكن فى بعض الأحيان يكون مطلوب إيجاد قوة الارتباط بين متغيرين على صورة بيانات وصفية يمكن وضعها فى صورة ترتيبية، مثال على هذا تقديرات الطلاب فى مادتين مختلفتين، فيكون من الصعب حساب معامل ارتباط بيرسون. لذلك نشأت الحاجة إلى إيجاد مقياس يعطى قوة الارتباط للبيانات الوصفية. وهذا المقياس هو ما يسمى بمعامل ارتباط الرتب لسبيرمان، وهو يعطى مقياسا للارتباط فى كل من البيانات الكمية والوصفية التى لها صفة الترتيب مثل تقديرات الطلاب، فإنه يمكن إعطاء رتب لها من حيث كبر التقدير وصغره وكذلك البيانات الكمية. نلاحظ أن رتب المتغيرين (x,y) تزيد وتنقص حسب زيادة ونقص كل من قيم المتغيرين (x,y). لذلك فإن حساب معامل الارتباط للرتب يقترب كثيرا من معامل ارتباط بيرسون، ولكن يمتاز عنه فى السهوله والدقة خاصة عندما تكون أزواج القيم أقل من ١٥.

ويعطى معامل ارتباط الرتب بالعلاقة التالية:

$$r_s = 1- \frac{6\sum d^2}{n(n^2-1)}, \qquad (2)$$

حيث r_s معامل ارتباط الرتب لسبيرمان، n تمثل عدد أزواج القيم (x,y)، d هى الفرق بين رتب أزواج القيم (x,y) ويمكن توضيح ذلك من خلال المثال التالي:

مثال (٢): اوجد معامل ارتباط الرتب لتقديرات الطلاب فى كل من مـادتى الإحصـاء والرياضيات كمـا هو موضح بالجدول التالي:

الرياضيات x	A	C	C	C	B	D
الإحصاء y	B	B	D	C	A	E

الحل: نقوم بتلخيص الحل فى الجدول التالي:

الرياضيات x	الإحصاء y	رتبة $x = a$	رتبة $y = b$	$d = a - b$	d^2
A	B	6	4.5	1.5	2.25
C	B	3	4.5	-1.5	2.25
C	D	3	2	1	1
C	C	3	3	0	0
B	A	5	6	-1	1
D	E	1	1	0	0
					6.5

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$$r_s = 1 - \frac{6 \times 6.5}{6(36-1)}$$

$$r_s = 1 - 0.186$$

$$r_s = 0.814$$

أى يوجد ارتباط طردى قوى بين تقديرات مادتى الرياضيات والإحصاء.

معامل الاقتران ومعامل التوافق

لقد سبق أن وضحنا بأن معامل ارتباط بيرسون يعطى قوة الارتباط فى حالة البيانات الكمية، وكذلك معامل ارتباط الرتب لسبيرمان، وهو يستخدم لإيجاد قوة الارتباط للرتب فى حالة البيانات الكمية والوصفية التى لها صفة الترتيب، ولكن قد تكون هناك بيانات وصفية لها صفات مميزة ولكن لا يمكن ترتيبها مثل الحالة الاجتماعية، وكذلك لون البشرة....إلخ. ولقياس قوة الارتباط لهذه البيانات نشأت الحاجة إلى إيجاد مقياس مناسب يقيس الارتباط بين هذه الصفات، ومنها معامل الاقتران الذى يستخدم عندما يكون لكل من الظاهرتين صفتين فقط. كذلك دراسة معامل التوافق فى حالة تكون كل من الظاهرتين أو إحداهما على الأقل من أكثر من صفتين كما سنوضح ذلك بالتفصيل كما يلي:

١ـ معامل الإقتران

يستخدم معامل الاقتران لقياس قوة الارتباط بين ظاهرتين كل منهما ذات صفتين فقط، وسوف يرمز له بالرمز C.C مثل دراسة قوة الارتباط بين التدخين والتعليم، حيث يوضح الجدول التالى التكرار للصفات:

التدخين / التعليم	يدخن	لا يدخن
متعلم	A	B
غير متعلم	C	D

فيكون معامل الإقتران C.C كالتالي:

$$C.C = \frac{AD - BC}{AD + BC} \qquad (3)$$

ونوضح ذلك بالمثال الآتي:

مثال (٣): عند دراسة العلاقة بين التعليم والتدخين فى إحدى الشركات أخذت عينة مكونة من ١٧ شخصا وكانت النتائج على النحو التالي:

التدخين / التعليم	يدخن	لا يدخن
متعلم	5	5
غير متعلم	3	4

احسب معامل الاقتران بين التدخين والتعليم.

الحل:

$$C.C = \frac{AD - BC}{AD + BC}$$

$$C.C = \frac{(5\times4)-(3\times5)}{(5\times4)+(3\times5)}$$

C.C = 0.14

وهو ارتباط ضعيف

٢ـ معامل التوافق

إذا كانت بيانات الظاهرتين التى لدينا عبارة عن بيانات وصفية لكل منهما أو إحداهما على الأقل وكانت مقسمة لأكثر من صفتين، فإن معامل الاقتران السابق لا يصلح فى هذه الحالة ويتم استخدام مقياس آخر هو معامل التوافق C. لحساب معامل التوافق نفرض أن لدينا الظاهرة X والتى لها r من الصفات والظاهرة الثانية Y والتى لها s من الصفات، ويوضح جدول الاقتران بين الظاهرتين كما يلي:

الصفة Y / الصفة X	Y_1	Y_2	$\cdots$	Y_s	المجموع
X_1	f_{11}	f_{12}	$\cdots$	f_{1s}	$f_{1.}$
X_2	f_{21}	f_{22}	$\cdots$	f_{2s}	$f_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
X_r	f_{r1}	f_{r2}	$\cdots$	f_{rs}	$f_{r.}$
المجموع	$f_{.1}$	$f_{.2}$	$\cdots$	$f_{.s}$	$f_{..}$

نحسب المقدار B ومنه نحسب معامل التوافق C بالعلاقة التالية:

$$C = \sqrt{\frac{B-1}{B}}, \qquad (4)$$

حيث:

$$B = \frac{(f_{11})^2}{f_{.1} f_{1.}} + \frac{(f_{12})^2}{f_{.2} f_{1.}} + \ldots + \frac{(f_{rs})^2}{f_{.s} f_{r.}}$$

ونوضح ذلك بالمثال التالي:

مثال (٤): عند دراسة العلاقة بين الرائحة ولون الزهرة لعينة مكونة من ٣٠ زهرة كانت لدينا النتائج التالية:

الرائحة Y / اللون X	بدون رائحة	له رائحة	المجموع
أصفر	6	4	10
أبيض	7	2	9
أحمر	6	5	11
المجموع	19	11	30

احسب معامل التوافق C بين اللون ورائحة الزهور.

الحل: نحسب قيمة B كالتالي:

$$B = \frac{6^2}{19\times 10}+\frac{7^2}{19\times 9}+\frac{6^2}{19\times 11}+\frac{4^2}{11\times 10}+\frac{2^2}{11\times 9}+\frac{5^2}{11\times 11}$$

B = 1.05

ويكون معامل التوافق C كالتالي:

$$C = \sqrt{\frac{B-1}{B}}$$

$$C = \sqrt{\frac{1.05-1}{1.05}}$$

C = 0.22

ويلاحظ أن قيمة معامل التوافق تبين مقدار قوة الارتباط وهى ضعيفة فى هذا المثال.

خط الانحدار البسيط

لقد سبق لنا دراسة العلاقة بين متغيرين (x, y) وإيجاد معامل الارتباط بينهما بعدة طرق وذلك لقياس قوة الارتباط وإتجاه العلاقة بينهما (طردية – عكسية) كما فى معاملى ارتباط بيرسون وسبيرمان ومدى قوة العلاقة كما فى حالة معاملى الاقتران والتوافق. وفيما يلى نبحث عن إيجاد معادلة رياضية تمثل أفضل توفيق لخط مستقيم يعبر عن البيانات فى شكلها الخطي. والغرض من إيجاد معادلة خط الانحدار هو التنبؤ بقيمة المتغير التابع لقيمة محددة من قيم المتغير المستقل. وتسمى العلاقة بين المتغير المستقل x والمتغير التابع y بمعادلة خط الانحدار البسيط. وعليه فإذا كان x متغيرًا مستقلاً، y متغيرًا تابعًا فإن المعادلة التى نحصل عليها تسمى بمعادلة خط انحدار y على x وهى على الصورة التالية:

$$y = a + bx, \qquad \textbf{(5)}$$

حيث يعرف a على أنه ثابت الانحدار (الجزء المقطوع من محور y) ويحسب من خلال العلاقة:

$$a = y - bx, \qquad \textbf{(6)}$$

كذلك يعرف b بمعامل انحدار y على x ويحسب من خلال العلاقة التالية:

$$b = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}, \qquad \textbf{(7)}$$

مثال (٥): اوجد معادلة خط انحدار درجات الإحصاء y على درجات الرياضيات x فى مثال (١).

الحل: نقوم بتلخيص الحسابات من خلال الجدول التالي:

x	y	xy	x^2	y^2
15	**13**	**195**	**225**	**169**
7	**9**	**63**	**49**	**81**
17	**19**	**323**	**289**	**361**
15	**15**	**225**	**225**	**225**
10	**11**	**110**	**100**	**121**
9	**8**	**72**	**81**	**64**
14	**16**	**224**	**196**	**256**
10	**11**	**110**	**100**	**121**
97	**102**	**1322**	**1265**	**1398**

$$b = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$b = \frac{1322 - \frac{97 \times 102}{8}}{1265 - \frac{(97)^2}{8}} = 0.96$$

$$a = \bar{y} - b\bar{x}$$

$$a = \frac{102}{8} - 0.96\left(\frac{97}{8}\right)$$

$$a = 1.11$$

أى أن معادلة خط انحدار y على x هي:

$$y = 1.11 + 0.96x$$

تمارين

١ - فى أحد أماكن بيع السيارات كانت المبيعات كالتالي:

4	1	6	5	1	1	2	3	عمر السيارة x
29	71	17	18	70	60	44	31	ثمن البيع y

- اوجد معامل الارتباط بين عمر السيارة (بالسنوات) وثمن البيع (بآلاف الجنيهات) بطريقة بيرسون.
- اوجد خط انحدار y على x.

٢ - الجدول التالى يمثل الدخل x والإنفاق y لمجموعة من الأسر بمئات الجنيهات.

40	39	27	38	44	42	66	56	x
28	20	25	19	22	27	38	31	y

- اوجد معامل ارتباط بيرسون وسبيرمان للدخل والإنفاق.
- اوجد خط انحدار y على x.
- اوجد قيمة الإنفاق عندما يصبح الدخل ٦٠٠٠ جنيه.

٣ - البيانات التالية تمثل تقديرات ثمانية من الطلاب فى مادتى الكيمياء والطبيعة.

B	E	D	C	E	D	B	A	الكيمياء
B	E	D	C	D	E	C	A	الطبيعة

اوجد معامل الارتباط المناسب لتقديرات الكيمياء والطبيعة.

الفصل السابع

بعض التوزيعات الاحتمالية وتطبيقاتها

الهدف من دراسة الفصل

بنهاية الدراسة لهذا الفصل يكون الدارس قادر على الالمام بالمفاهيم التالية:

١- توزيع ذى الحدين

٢- توزيع بواسون

٣- التوزيع فوق الهندسي

٤- التوزيع الطبيعى القياسي

٥- توزيع ت

٦- توزيع ف

الفصل السابع
بعض التوزيعات الاحتمالية وتطبيقاتها

مقدمة

إذا كان مدى المتغير العشوائى X مجموعة محدودة من القيم فى هذه الحالة يمكن القول بأن X متغير عشوائى منفصل، مثال على ذلك عدد الوحدات المنتجة لإحدى الآلات، عدد أطفال الأسرة، عدد مرات ظهور الكتابة عند إلقاء قطعة نقود،....وهكذا.

أما إذا كان عدد القيم التى يأخذها المتغير العشوائى X غير محدود (لا يمكن عده) فى هذه الحالة يقال إن المتغير العشوائى X متغير عشوائى متصل. مثال على ذلك أوزان أو أطوال مجموعة من الطلاب أو الدخل السنوى لمجموعة من الأسر، أو درجات الحرارة فى فترة ما،... إلخ.

إن دراسة المتغيرات العشوائية المتصلة والمنفصلة والتوزيعات الاحتمالية المصاحبة لكل نوع من أنواع المتغيرات العشوائية لتساعدنا فى الحصول على نتائج يمكن استخدامها فى تقديرات معالم المجتمع كذلك اختبارات الفروض المتعلقة باتخاذ القرارات المهمة حيث يتم اتخاذ مثل هذه القرارات على أساس علمى صحيح.

وفيما يلى بعض التوزيعات الاحتمالية المهمة التى لها العديد من التطبيقات المهمة فى الحياة العملية مثل توزيع ذى الحدين وتوزيع بواسون والتوزيع فوق الهندسي، وذلك للمتغير العشوائى المنفصل (المتقطع) والتوزيع الطبيعى وتوزيع ت وتوزيع ف للمتغير العشوائى المتصل.

توزيع ذى الحدين

توجد العديد من ظواهر الحياة تكون النتيجة الممكنة لها إما نجاحًا وإما فشلاً، واحتمال أن تكون نتيجة التجربة نجاحًا هو p بينما إحتمال نتيجة الفشل هو q بحيث $(p+q=1)$. فإذا تم تكرار مثل هذه التجربة n من المرات فإننا نحصل كل مرة على نجاح أو فشل (نتيجة كل محاولة مستقلة عن نتيجة المحاولة الأخرى) بمعنى أنه يتم إجراء التجربة n من المحاولات المستقلة. المتغير العشوائى x الذى يمثل عدد مرات الحصول على نجاح خلال n من المحاولات يقال أنه يتبع توزيع ذى الحدين الذى له داله التوزيع الاحتمالى التالية:

$$P(X=x) = \binom{n}{x} p^x q^{n-x}, \qquad x=0,1,2,\ldots n$$

ويستخدم هذا التوزيع فى الحياة العملية لكثير من الظواهر، مثل النجاح والرسوب فى الامتحان لمجموعة من الطلبة، إصابة هدف معين من عدمه أو التدخين وعدم التدخين لمجموعة من الأشخاص فى مدينة ما...إلخ.

مثال (١): إذا كانت نسبة النجاح فى أحد المقررات هى ٠.٨ فإذا تقدم لهذا الامتحان ١٥ طالب ما هو احتمال أن ينجح:

١- جميع الطلاب
٢- ٨ طلاب
٣- ٦ طلاب
٤- ولاطالب

الحل:

$n = 15, \qquad p = 0.8, \qquad q = 0.2$

$$P(X=x) = \binom{15}{x} (0.8)^x (0.2)^{15-x}, \qquad x = 0,1,2,\ldots,15$$

١- نجاح جميع الطلاب:

$$P(X=15) = \binom{15}{15} (0.8)^{15} (0.2)^{15-15}$$

$$P(X=15) = 1 \times 0.035 \times 1$$

$$P(X=15) = 0.035$$

٢- نجاح ٨ طلاب:

$$P(X=8) = \binom{15}{8} (0.8)^8 (0.2)^{15-8}$$

$$P(X=8) = 6435 \times 0.1677722 \times 0.0000128$$

$$P(X=8) = 0.013819$$

٣- نجاح ٦ طلاب:

$$P(X=6) = \binom{15}{6} (0.8)^6 (0.2)^{15-6}$$

$$P(X=6) = 5005 \times 0.262144 \times 0.000000512$$

$$P(X=6) = 0.000672$$

٤- عدم نجاح أى طالب:

$$P(X=0) = \binom{15}{0} (0.8)^0 \; (0.2)^{15-0}$$

$P(X=0)$ = 1 × 1 × 0

$P(X=0)$= 0

مثال (٢): إذا كانت نسبة الإصابة بمرض الإنفلوانزا فى إحدى المدن فى فصل الشتاء هى ٠.٠٦. تم إختيار ٢٠ شخص من هذه المدينة، ما احتمال أن يكون:

١- ٧ أشخاص مصابون بالإنفلونزا.

٢- جميعهم أصحاء.

٣- جميعهم مرضى.

٤- ما إحتمال أن يكون نصفهم مرضى.

الحل:

n = **20,** p = **0.6,** q = **0.4**

$$P(X=x) = \binom{20}{x} (0.6)^x \; (0.4)^{20-x}, \qquad x = 0, 1, 2, \ldots, 20$$

١- ٧ أشخاص مصابون بالإنفلونزا:

$$P(X=7) = \binom{20}{7} (0.6)^7 \; (0.4)^{20-7}$$

$P(X=7)$= 77520 × 0.0279936 × **0.00000671**

$P(X=7)$= 0.014563

٢- جميعهم أصحاء:

$$P(X=0) = \binom{20}{0} (0.6)^0 \; (0.4)^{20-0}$$

$P(X=0)$ = 1 × 1 × 0.000

$P(X=0)$= 0

٣- جميعهم مرضى:

$$P(X=20) = \binom{20}{20} (0.6)^{20} \; (0.4)^{20-20}$$

$P(X=20)$ = 1 × 0.00004 × 1

$P(X=20)$= 0.00004

٤- إحتمال أن يكون نصفهم مرضى:

$$P(X=10) = \binom{20}{10} (0.6)^{10} \; (0.4)^{20-10}$$

$P(X=10)$ = 184756 × 0.006046618 × 0.0001

$P(X=10)$= 0.111715

مثال (٣): صندوق يحتوى على ٧ مصابيح فإذا كان احتمال أن يكون المصباح جيدًا هو ٠.٨، تم اختيار ٣ مصابيح عشوائيا، ما احتمال:

١- أن تكون جميع المصابيح جيدة.
٢- أن يكون هناك مصباح تالف.
٣- أن تكون جميع المصابيح تالفة.
٤- أن يكون هناك مصباح جيد على الأقل.
٥- أن يكون هناك مصباح جيد على الأكثر.

الحل:

$n = 3, \qquad p = 0.8, \qquad q = 0.2$

$$P(X=x) = \binom{3}{x} (0.8)^x \, (0.2)^{3-x}, \qquad x = 0,1,2,3$$

١- أن تكون جميع المصابيح جيدة:

$$P(X=3) = \binom{3}{3} (0.8)^3 \, (0.2)^{3-3}$$
$$P(X=3) = 1 \times 0.512 \times 1$$
$$P(X=3) = 0.512$$

٢- أن يكون هناك مصباح تالف:

$$1- P(X=2) = 1- \binom{3}{2} (0.8)^2 \, (0.2)^{3-2}$$
$$= 1- (3 \times 0.64 \times 0.2)$$
$$= 1- 0.384$$
$$1- P(X=2) = 0.616$$

٣- أن تكون جميع المصابيح تالفة:

$$P(X=0) = \binom{3}{0} (0.8)^0 \, (0.2)^{3-0}$$
$$P(X=0) = 1 \times 1 \times 0.008$$
$$P(X=0) = 0.008$$

٤- أن يكون هناك مصباح جيد على الأقل:

$$P(X\geq 1) = P(X=1) + P(X=2) + P(X=3)$$
$$P(X\geq 1) = \binom{3}{1} (0.8)^1 \, (0.2)^{3-1} + \binom{3}{2} (0.8)^2 \, (0.2)^{3-2} + \binom{3}{3} (0.8)^3 \, (0.2)^{3-3}$$
$$P(X\geq 1) = 0.096 + 0.384 + 0.512$$
$$P(X\geq 1) = 0.992$$

٥- أن يكون هناك مصباح جيد على الأكثر

$P(X \leq 1) = P(X=1) + P(X=0)$
$P(X \leq 1) = 0.096 + 0.008$
$P(X \leq 1) = 0.104$

توزيع بواسون

يستخدم هذا التوزيع لحساب احتمال وصول عدد معين إلى مركز الخدمة، مثال على ذلك:

- ماكينة السحب الآلي.
- شباك البنك.
- طلمبة بنزين فى محطة الوقود.
- سويتش التليفون.
- وصول السيارات إلى أماكن الانتظار.

فى هذا التوزيع سوف يعبر المتغير العشوائى X عن أعداد الواصلين خلال فترة زمنية معينة (دقيقة، ساعة،...) $x = 0, 1, 2, \ldots$ وسوف تمثل λ متوسط أعداد الواصلين إلى محطة الخدمة فى الفترة الزمنية المحددة.

وسوف تعبر دالة الاحتمال عن وجود عدد X فى فترة زمنية محددة فى محطة الخدمة. وتأخذ دالة الاحتمال الشكل التالي:

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x = 0, 1, 2, \ldots$$

مثال (٤): إذا كان متوسط عدد طالبى استخدام ماكينة السحب الآلى فى أحد البنوك هو ٥ أفراد كل نصف ساعة.

أ- احسب الاحتمالات التالية لإعداد الواصلين كل نصف ساعة بأن يكون:

١- ١٠ أشخاص.
٢- يقل عن ٣ أشخاص.
٣- أكثر من شخص واحد.
٤- يتراوح العدد بين ٤ و ٨ أشخاص.

ب- احسب نفس الاحتمالات السابقة إذا كان معدل الوصول كل ربع ساعة.

ج- احسب نفس الاحتمالات السابقة إذا كان معدل الوصول كل ساعة.

الحل: أ- معدل الوصول كل نصف ساعة.

$\lambda = 5$,

$$P(X=x) = \frac{e^{-5}5^{x}}{x!}, \qquad x = 0, 1, 2, \ldots$$

1- $P(X=10)= \frac{e^{-5}5^{10}}{10!}$

$P(X=10)=$ **0.018132789**

2- $P(X<3)= P(X=0) + P(X=1) + P(X=2)$

$P(X<3) = \frac{e^{-5}5^{0}}{0!} + \frac{e^{-5}5^{1}}{1!} + \frac{e^{-5}5^{2}}{2!}$

$P(X<3) =$ **0.006737947 + 0.033689735 + 0.084224337**

$P(X<3)=$ **0.124652**

3- $P(X>1) = 1- P(X\leq 1)$

$P(X>1) = 1- (P(X=0) + P(X=1))$

$P(X>1) =$ **1- (0.006737947 + 0.033689735)**

$P(X>1) =$ **1- 0.040428**

$P(X>1) =$ **0.959572318**

4- $P(4\leq X\leq 8) = P(X=4) + P(X=5)+ P(X=6) + P(X=7) + P(X=8)$

$P(4\leq X\leq 8) = \frac{e^{-5}5^{4}}{4!} + \frac{e^{-5}5^{5}}{5!} + \frac{e^{-5}5^{6}}{6!}+\frac{e^{-5}5^{7}}{7!} + \frac{e^{-5}5^{8}}{8!}$

$P(4\leq X\leq 8) =$ **0.17546737 + 0.17546737 + 0.146222808 + 0.104444863 +**

+ 0.065278039

$P(4\leq X\leq 8) =$ **0.66688045**

ب– معدل الوصول كل ربع ساعة.

إذا متوسط معدل الوصول:

$\lambda = 5 \times \frac{1}{2} = 2.5$

$P(X=x) = \frac{e^{-2.5}2.5^{x}}{x!}, \qquad x=0,1,2,\ldots$

1- $P(X=10)= \frac{e^{-2.5}2.5^{10}}{10!}$

$P(X=10)=$ **0.000215725**

2- $P(X<3)= P(X=0) + P(X=1) + P(X=2)$

$P(X<3) = \frac{e^{-2.5}2.5^{0}}{0!} + \frac{e^{-2.5}2.5^{1}}{1!} + \frac{e^{-2.5}2.5^{2}}{2!}$

$P(X<3) =$ **0.082084999 + 0.205212497 + 0.256515621**

$P(X<3)=$ **0.543813**

3- $P(X>1) = 1- P(X\leq 1)$

$P(X>1) = 1- (P(X=0) + P(X=1))$

$P(X>1) =$ **1- (0.082084999 + 0.205212497)**

$P(X>1) =$ **1- 0.287297**

$P(X>1) =$ **0.712703**

4- $P(4\leq X\leq 8) = P(X=4) + P(X=5)+ P(X=6) + P(X=7) + P(X=8)$

$P(4 \le X \le 8) = \frac{e^{-2.5}2.5^4}{4!} + \frac{e^{-2.5}2.5^5}{5!} + \frac{e^{-2.5}2.5^6}{6!} + \frac{e^{-2.5}2.5^7}{7!} + \frac{e^{-2.5}2.5^8}{8!}$

$P(4 \le X \le 8)$ = **0.133601886 + 0.066800943 + 0.027833726 + 0.009940617 +**

+ 0.003106443

$P(4 \le X \le 8)$ = **0.241284**

ج– معدل الوصول كل ساعة

إذا متوسط معدل الوصول:

$\lambda = 5 \times 2 = 10$

$P(X=x) = \frac{e^{-10}10^x}{x!}, \qquad x = 0,1,2,\ldots$

1- $P(X=10) = \frac{e^{-10}10^{10}}{10!}$

$P(X=10)$ = **0.125110036**

2- $P(X<3) = P(X=0) + P(X=1) + P(X=2)$

$P(X<3) = \frac{e^{-10}10^0}{0!} + \frac{e^{-10}10^1}{1!} + \frac{e^{-10}10^2}{2!}$

$P(X<3)$ = **0.0000454 + 0.000454 + 0.00227**

$P(X<3)$ = **0.002769**

3- $P(X>1) = 1 - P(X \le 1)$

$P(X>1) = 1 - (P(X=0) + P(X=1))$

$P(X>1)$ = **1- (0.0000454 + 0.000454)**

$P(X>1)$ = **1- 0.000499**

$P(X>1)$ = **0.999501**

4- $P(4 \le X \le 8) = P(X=4) + P(X=5) + P(X=6) + P(X=7) + P(X=8)$

$P(4 \le X \le 8) = \frac{e^{-10}10^4}{4!} + \frac{e^{-10}10^5}{5!} + \frac{e^{-10}10^6}{6!} + \frac{e^{-10}10^7}{7!} + \frac{e^{-10}10^8}{8!}$

$P(4 \le X \le 8)$ = **0.01891664 + 0.03783327 + 0.06305546 + 0.09007923 +**

+ 0.11259903

$P(4 \le X \le 8)$ = **0.322484**

مثال (٥): إذا كان متوسط وصول السفن إلى أحد الموانيء سفينتين فى اليوم. اوجد احتمال أن يصل إلى هذا الميناء فى يوم معين ثلاث سفن.

الحل:

$\lambda = 2$,

$P(X=x) = \frac{e^{-2}2^x}{x!}, \qquad x = 0,1,2,\ldots$

$P(X=3) = \frac{e^{-2}2^3}{3!}$

$P(X=3)$ = **0.18044704**

التوزيع فوق الهندسى

يفيد التوزيع فوق الهندسى فى حالة المجتمعات التى يتم تقسيمها إلى صفتين أو جزءين حيث تتميز هذه المجتمعات بكونها محدودة وصغيرة. وتتم عملية سحب العينة من المجتمع بدون إرجاع. وعليه فإن شرط استقلال المحاولات يكون غير متحقق، حيث يؤثر السحب بدون إرجاع على نسبة إحدى الصفتين وذلك لصغر حجم المجتمع. مثال على ذلك سحب عينة من الطلاب فى إحدى الشعب، سحب عينة من السيارات لدراسة عدد السيارات المعيبة...إلخ.

والمتغير العشوائى X فى هذه الحالة يمثل عدد حالات النجاح وتكون دالة التوزيع الاحتمالى له على الصورة:

$$P(X=x)=\frac{\binom{a}{x}\binom{b}{n-x}}{\binom{N}{n}}, \qquad x=0,1,2,..n$$

حيث:

- N تمثل حجم المجتمع.
- n تمثل حجم العينة المسحوبة من المجتمع.
- $a+b=N$.

مثال (٦): معرض سيارات به ٤٨ سيارة من بينها ٨ سيارات معيبة. أختيرت عينة عشوائية من ٥ سيارات اوجد:

أ- احتمال أن تكون العينة المسحوبة كلها سليمة

ب- احتمال أن تكون سيارة واحدة معيبة

ج- احتمال وجود سيارتين معيبتين على الأقل بالعينة

الحل: سوف يمثل المتغير العشوائى X عدد السيارات المعيبة فى العينة المسحوبة، a يمثل عدد السيارات المعيبة فى المجتمع وb يمثل عدد السيارات السليمة فى المجتمع. وسوف تأخذ دالة التوزيع الإحتمالى الصورة:

$$P(X=x)=\frac{\binom{8}{x}\binom{40}{5-x}}{\binom{48}{5}}, \qquad x=0,1,2,3,4,5$$

أ- احتمال أن تكون العينة المسحوبة كلها سليمة

$$P(X=0) = \frac{\binom{8}{0}\binom{40}{5-0}}{\binom{48}{5}}$$

$$P(X=0) = \frac{1\times 6580}{17123}$$

$$P(X=0) = \mathbf{0.38}$$

ب- احتمال أن تكون سيارة واحدة معيبة

$$P(X=1) = \frac{\binom{8}{1}\binom{40}{5-1}}{\binom{48}{5}}$$

$$P(X=1) = \frac{8\times 9139}{17123}$$

$$P(X=1) = \mathbf{0.43}$$

ج- احتمال وجود سيارتين معيبتين على الأقل بالعينة

$$P(X\geq 2) = 1 - (P(X=0) + P(X=1))$$

$$P(X\geq 2) = 1 - (0.38 + 0.43)$$

$$P(X\geq 2) = 0.19$$

مثال (٧): شحنة من ٨٠ جهاز كهربائى من بينها ٤ أجهزة معيبة، تم سحب عينة عشوائية من ٣ أجهزة.اوجد إحتمال أن تحتوى هذه العينة على جهاز واحد متعطل.

الحل: سوف يمثل المتغير العشوائى X عدد الأجهزة المعيبة فى العينة المسحوبة، a يمثل عدد الأجهزة المعيبة فى الشحنة وb يمثل عدد الأجهزة السليمة فى الشحنة. وسوف تأخذ دالة التوزيع الاحتمالى الصورة:

$$P(X=x) = \frac{\binom{4}{x}\binom{76}{3-x}}{\binom{80}{3}}, \qquad x=0,1,2,3$$

احتمال أن يكون هناك جهاز واحد معيب بالعينة هو:

$$P(X=1) = \frac{\binom{4}{1}\binom{76}{3-1}}{\binom{80}{3}}$$

$$P(X=1) = \frac{4\times 286}{8216}$$

$$P(X=1) = \mathbf{0.14}$$

التوزيع الطبيعى القياسى

التوزيع الطبيعى هو عبارة عن توزيع له الشكل الجرسى وهو متماثل حول المتوسط μ وتساوى المساحة الكلية تحت المنحنى الواحد الصحيح. ولهذا التوزيع العديد من التطبيقات المهمة فى الحياة العملية لمعظم الظواهر الطبيعية مثل الأطوال والأوزان ودرجات الطلاب ومقياس ضغط الدم وغيرها من الظواهر البيولوجية والعلمية. وتعطى داله الكثافة الحتمالية للمتغير العشوائى X الذى يتبع التوزيع الطبيعى ($X \approx N(\mu, \sigma^2)$) كالتالي:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

حيث:

$$-\infty < X < \infty, \qquad -\infty < \mu < \infty, \qquad \sigma > 0$$

هذا ويعتبر التوزيع الطبيعى القياسى حالة خاصة من التوزيع الطبيعى وذلك حيث يكون المتوسط له يساوى الصفر والانحراف المعيارى يساوى الواحد الصحيح. وسوف نرمز للمتغير العشوائى الذى يتبع التوزيع الطبيعى القياسى بالرمز Z أى أن $Z \approx N(0,1)$ وتعطى دالة الكثافة الاحتمالية لهذا التوزيع من خلال العلاقة:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \qquad -\infty < z < \infty$$

وكما هو معلوم هناك جداول خاصة بهذا التوزيع يمكن من خلالها حساب الاحتمالات المطلوبة.

هذا ويمكن تحويل جميع القيمة غير القياسية (قيم تتبع التوزيع الطبيعى) إلى قيم قياسية وذلك من خلال العلاقة التالية:

$$Z = \frac{x-\mu}{\sigma}$$

فمثلاً إذا كان لدينا متغير عشوائى يتبع التوزيع الطبيعى $X \approx N(16,16)$ فيمكن تحويله إلى متغير عشوائى له التوزيع الطبيعى القياسى كالتالي:

$$Z = \frac{x-16}{4}$$

المتغير العشوائى الجديد Z يتوزع توزيعا طبيعيا قياسيا له المتوسط صفر والانحراف المعيارى واحد.

مثال (٨): باستخدام الجدول الإحصائى الذى يعطى المساحة تحت المنحنى الطبيعى القياسي، اوجد الاحتمالات التالية:

1- $P(z \le 1.72)$, 2- $P(z \le 1.07)$

3- $P(z \ge 0.29)$ 4- $P(-1.91 \le z \le 0.45)$

الحل:

1- $P(z \leq 1.72)$ = 0.9573
2- $P(z \leq 1.07)$ = 0.8577
3- $P(z \geq 0.29)$ = 1- $P(z < 0.29)$
$P(z \geq 0.29)$ = 1- 0.6141
$P(z \geq 0.29)$ = 0.3859
4- $P(-1.91 \leq z \leq 0.45)$= $P(z \leq 0.45)$ - $P(z \leq -1.91)$
$P(-1.91 \leq z \leq 0.45)$ = 0.6736 – 0.0281
$P(-1.91 \leq z \leq 0.45)$ = 0.6455

مثال (٩): إذا كانت $X \approx N(16,16)$ اوجد الاحتمالات الأتية:

1- $P(X \leq 14)$
2- $P(X \geq 22)$

الحل: نلاحظ أن المتغير العشوائى X هو متغير عشوائى غير قياسى وعليه سوف نقوم بتحويله إلى المتغير العشوائى القياسى كالتالي:

1- $x = 14 \Rightarrow Z = \frac{x-\mu}{\sigma}$
$\Rightarrow\ z = \frac{14-16}{4}$ = -0.5
$\Rightarrow P(X \leq 14)$ = $P(Z \leq -0.5)$ = 0.3085
2- $x = 22 \Rightarrow Z = \frac{x-\mu}{\sigma}$
$\Rightarrow\ z = \frac{22-16}{4}$ = 1.5
$\Rightarrow P(X \geq 22)$ = $P(Z \geq 1.5)$ = 1- $P(Z < 1.5)$
$\Rightarrow P(Z \geq 1.5)$ = 1- 0.9332 = 0.0668

مثال (١٠): إذا كانت درجة ذكاء الطلاب فى الجامعة تتبع التوزيع الطبيعى بمتوسط ١٠٥ وانحراف معيارى ١٠. ما هى نسبة الطلبة الذين تقع درجة ذكائهم بين ١١٤،١٠٠.

الحل:

$X \approx N(105,100)$
$P(100 \leq X \leq 114) = P\left(\frac{100-105}{10} \leq \frac{X-\mu}{\sigma} \leq \frac{114-105}{10}\right)$
$P(100 \leq X \leq 114) = P(-0.5 \leq Z \leq 0.9)$
$P(100 \leq X \leq 114) = P(z \leq 0.9)$ - $P(z \leq -0.5)$
$P(100 \leq X \leq 114)$= 0.8159 – 0.3085
$P(100 \leq X \leq 114)$= 0.5074

توزيع ت

إذا كانت دالة الكثافة الإحتمالية للمتغير العشوائى t على الصورة:

$$f(t)= c(1+\frac{t^2}{\nu})^{-\frac{\nu+1}{2}}, \qquad -\infty<t<\infty$$

فإن هذا التوزيع يسمى توزيع t حيث ν تمثل درجات الحرية و c ثابت يعتمد على ν ليجعل المساحة تحت المنحنى تساوى الواحد الصحيح.

ويتشابه توزيع t مع التوزيع الطبيعى القياسى من حيث الشكل الجرسى إلا أنه أكثر انخفاضا منه وعندما تزداد درجات الحرية فإن توزيع t يقترب من التوزيع الطبيعى القياسي.

وهناك جداول خاصة لهذا التوزيع مثل التوزيع الطبيعى القياسى إلا أن جداول توزيع t تختلف بعض الشيئ حيث يعتمد الجدول على درجات الحرية التى تمثل العمود الرأسى والمساحات التى تمثل الخط الأفقى بينما الأعداد داخل الجدول تمثل قيم t المناظرة لدرجات الحرية والمساحة.

مثال (١١): اوجد قيمة t لكل من:

$t\,(0.975, 20)$ $\qquad$ $t\,(0.995, 12)$

$t\,(0.95, 5)$ $\qquad$ $t\,(0.90, 7)$

الحل: باستخدام جدول t نحصل على:

$t\,(0.975, 20) = 2.086$

$t\,(0.995, 12) = 3.055$

$t\,(0.95, 5) = 2.015$

$t\,(0.90, 7) = 1.415$

مثال (١٢): اوجد درجات الحرية المناظرة للقيم التالية:

$t\,(0.975, \nu) = 2.228$ $\qquad$ $t\,(0.995, \nu) = 2.921$

$t\,(0.95, \nu) = 1.721$ $\qquad$ $t\,(0.90, \nu) = 1.337$

الحل: باستخدام جدول t نحصل على:

$t\,(0.975, \nu) = 2.228 \Rightarrow \nu = 10$

$t\,(0.995, \nu) = 2.921 \Rightarrow \nu = 16$

$t\,(0.95, \nu) = 1.721 \Rightarrow \nu = 21$

$t\,(0.90, \nu) = 1.337 \Rightarrow \nu = 16$

توزيع ف

من التوزيعات المهمة التى تستخدم فى اختبارات الفروض هو توزيع **F**. هذا التوزيع دالة الكثافة الاحتمالية له تعطى من خلال العلاقة التالية:

$$f(F)= \frac{cF^{(\nu_1-2)/2}}{(\nu_2+\nu_1F)^{(\nu_1+\nu_2)/2}}, \qquad F>0$$

يسمى هذا التوزيع بتوزيع **F** وعبر عنه بالرمز $F(\nu_1,\nu_2)$ حيث ν_1 و ν_2 يمثلان درجات الحرية و c هو ثابت يعتمد على درجات الحرية حتى تصبح المساحة تحت المنحنى مساوية للواحد الصحيح.

ويلاحظ فى هذا التوزيع بأنه ملتو ناحية اليمين هذا وكلما ازدادت درجات الحرية اقترب هذا التوزيع من التوزيع الطبيعى القياسي.

هذا وسوف نستخدم الرمز $F(\alpha,\nu_1,\nu_2)$ (α ترمز إلى المساحة) لإيجاد قيمة **F** المناظرة لدرجات الحرية والمساحة المعطاة.

مثال (١٣): اوجد:

F (0.01, 11 , 15)   **F (0.05, 10 , 7)**

الحل:

F (0.01, 11 , 15) = 0.235
F (0.05, 10 , 7) = 0.318

تمارين

١- إذا كانت نسبة المعيب فى الإنتاج تمثل ١٠٪ سحبت عينة مكونة من ٥ وحدات. اوجد الاحتمالات التالية:

أ- لا يوجد فى العينة وحدة معيبة.

ب-توجد وحدة معيبة فقط.

ت-توجد وحدة معيبة على الأكثر.

ث-توجد وحدتان معيبتان على الأقل.

٢- إذا كان احتمال أن يتخرج طالب من الجامعة هو ٠.٠٦. سحبت عينة مكونة من ٤ طلاب، اوجد الاحتمالات الآتية:

- أن يتخرج جميع الطلاب فى العينة.
- أن يتخرج طالبان فقط.
- أن يتخرج طالبان على الأقل.

٣- مصنع به ١٥ عامل و ٥ مهندسين، سحبت عينة عشوائية مكونة من ٣ أفراد. اوجد الاحتمالات الآتية:

١-العينة كلها من المهندسين.

٢-العينة بها عامل واحد ومهندسان.

٣-العينة كلها من العمال.

٤- إذا كان المتغير العشوائى Z له توزيع طبيعى قياسي، فاوجد الاحتمالات التالية:

- $P(Z<1.8)$
- $P(Z>-0.5)$
- $P(-0.2<Z<0.5)$

٥- إذا كان المتغير العشوائى X الذى يتبع التوزيع الطبيعى بمتوسط يساوى ٨٠ وانحراف معيارى يساوى ٤.٨ اوجد الاحتمالات التى يأخذها المتغير العشوائى للقيم التالية:

١- أقل من ٨٧.٢.

٢- أكبر من ٧٦.٤.

٣- بين ٨١.٢، ٨٦.

٤- بين ٧١.٦، ٨٨.٤.

٦- إذا كانت درجة الحرارة خلال شهر مارس تتبع التوزيع الطبيعى بمتوسط ٢٠ درجة وانحراف معيارى ٣.٣٣ درجة.

اوجد احتمال أن تكون درجة الحرارة بين ٢١.١١، ٢٦.٦٦ فى هذا الشهر.

٧- إذا كان إحتمال أن يكسب فريق مباراة هو ٠.٧٥ فإذا لعب هذا الفريق ٤ مباريات. اوجد احتمال أن يكسب هذا الفريق:

- **مباراة واحدة على الأقل.**
- **مبارتان فقط.**
- **أكثر من نصف المباريات.**

٨- إذا كان ٢٠٪ من إنتاج ماكينة مسامير تالفًا. اوجد احتمال أن يكون من بين ٤ مسامير تم اختيارها عشوائيا:

- **مسمار واحد تالف.**
- **لا توجد مسامير تالفة.**
- **يوجد مسماران تالفان على الأكثر.**

٩- صندوق يحتوى على ٤ كرات بيضاء و ٦ كرات حمراء. تم سحب عينة عشوائية مكونة من ٣ كرات بدون إرجاع. افرض أن المتغير العشوائى X يمثل عدد الكرات البيضاء فى العينة المسحوبة اوجد التوزيع الاحتمالى للمتغير العشوائى X.

١٠- إذا كان متوسط عدد الحوادث على الطريق الصحراوى هو ٣ حوادث يوميا. ما هو احتمال وقوع أكثر من ٥ حوادث فى يوم ما.؟

١١- اوجد قيمة t لكل من:

t **(0.95, 20)**
t **(0.90, 28)**
t **(0.99,12)**
t **(0.975, 7)**
t **(0.995, 13)**

١٢- اوجد قيمة **b** فى كل مما يأتي:

(b , 5) = 2.015
t **(b , 20) = 1.325**
t **(b , 23) = 2.069**
t **(b , 12) = 2.681**
t **(b , 15) = 2.131**

١٣- اوجد:

F (0.01, 7 , 12)
F (0.05, 12 , 5)
F (0.01, 5 , 8)
F (0.05, 5 , 5)

١٤- اوجد قيمة **b** فى كل مما يأتي:

F (b, 8 , 9) = 3.23
F (b, 9 , 11) = 4.63
F (b, 3 , 24) = 3.72
F (b, 2 , 24) = 3.

ملحوظة هامة:

من باب الأمانة العلمية ومبدأ رد الفضل الى أهله فقد تم الاعتماد في اعداد هذه النسخة بشكل كبير أن لم يكن أساسى على كتاب الإحصاء الخاص ببرنامج الطرق المؤدية الى التعليم العالى بجامعة عين شمس فللسادة المؤلفون جزيل الشكر، وعلى رأسهم الأستاذ الدكتور سيد كاسب.