حامعة جنوب الوادى

كلية التربية بالغردقة

الفرقة الثانية عام رياضيات (Math)

المادة : (Pure 6)

إستاذ المادة / د. اسماعيل جاد امين

الفصل الدراسي الأول

Ordinary Differential Equations (I)

Ordinary Differential Equations

Chapter 1
Introduction to ODEs

- Objectives of Lesson

- Recall basic definitions of ODEs:
  - Order
  - Linearity
  - Initial conditions
  - Solution
- Classify ODEs based on:
  - Order, linearity, and conditions.
- Classify the solution methods.

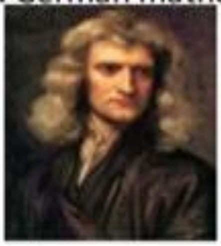
Ordinary Differential Equations (ODEs)

Chapter1

- History of differential equations:

INVENTION OF DIFFERENTIAL EQUATION:

- In mathematics, the **history of differential equations** traces the development of "differential equations" from calculus, which itself was independently invented by English physicist Isaac Newton and German mathematician Gottfried Leibniz.



- The history of the subject of differential equations, in concise form, from a synopsis of the recent article "The History of Differential Equations, **1670-1950**"

  "**Differential equations** began with Leibniz, the Bernoulli brothers, and others from the **1680s**, not long after Newton's 'fluxional equations' in the **1670s**."

- Definition of DEs and some properties:

Differential Equations

Definition A differential equation is an equation involving derivatives of an unknown function and possibly the function itself as well as the independent variable.

Example $\boxed{y' = \sin(x), \quad (y')^4 - y^2 + 2xy - x^2 = 0, \quad y'' + y^3 + x = 0}$

1st order equations — 2nd order equation

Definition The order of a differential equation is the highest order of the derivatives of the unknown function appearing in the equation

In the simplest cases, equations may be solved by direct integration.

Examples $y' = \sin(x) \Rightarrow y = -\cos(x) + C$

$y'' = 6x + e^x \Rightarrow y' = 3x^2 + e^x + C_1 \Rightarrow y = x^3 + e^x + C_1 x + C_2$

Observe that the set of solutions to the above 1st order equation has 1 parameter, while the solutions to the above 2nd order equation depend on two parameters.

Derivatives

Derivatives

Ordinary Derivatives	Partial Derivatives
$\dfrac{dv}{dt}$	$\dfrac{\partial u}{\partial y}$
v is a function of one independent variable	u is a function of more than one independent variable

ODE PDE

dy/dx

Differential Equations

Differential Equations

Ordinary Differential Equations	**Partial Differential Equations**
$\dfrac{d^2v}{dt^2}+6tv=1$	$\dfrac{\partial^2 u}{\partial y^2}-\dfrac{\partial^2 u}{\partial x^2}=0$
involve one or more Ordinary derivatives of unknown functions	involve one or more partial derivatives of unknown functions

Ordinary Differential Equations (ODEs) involve one or more ordinary derivatives of unknown functions with respect to one independent variable

Examples :

$$\frac{dv(t)}{dt} - v(t) = e^t$$

$$\frac{d^2x(t)}{dt^2} - 5\frac{dx(t)}{dt} + 2x(t) = \cos(t)$$

x(t): unknown function

t: independent variable

Example 2:

$$\frac{dv}{dt} = 9.8 - \frac{c}{M}v$$

Ordinary differential equation

$$\frac{dv}{dt} = 9.8 - \frac{c}{M}v$$

(Dependent variable) unknown function to be determined

$$\frac{dv}{dt} = 9.8 - \frac{c}{M}v$$

(independent variable) the variable with respect to which other variables are differentiated

- Order of a Differential Equation:

The **order** of an ordinary differential equation is the order of the highest order derivative

Examples:

$$\frac{dx(t)}{dt} - x(t) = e^{t}$$ First order ODE

$$\frac{d^2x(t)}{dt^2} - 5\frac{dx(t)}{dt} + 2x(t) = \cos(t)$$ Second order ODE

$$\left(\frac{d^2x(t)}{dt^2}\right)^3 - \frac{dx(t)}{dt} + 2x^4(t) = 1$$ Second order ODE

- Linear ODE:

An **ODE** is linear if the unknown function and its derivatives appear to power one**.** No product of the unknown function and/or its derivatives.

Examples:

$$\frac{dx(t)}{dt} - x(t) = e^{t}$$ Linear ODE

$$\frac{d^2x(t)}{dt^2} - 5\frac{dx(t)}{dt} + 2t^2x(t) = \cos(t)$$ Linear ODE

$$\left(\frac{d^2x(t)}{dt^2}\right)^3 - \frac{dx(t)}{dt} + \sqrt{x(t)} = 1$$ Non-linear ODE

- Nonlinear ODE:

Examples of nonlinear ODE :

$$\frac{dx(t)}{dt} - \cos(x(t)) = 1, \qquad \frac{d^2x(t)}{dt^2} - 5\frac{dx(t)}{dt}x(t) = 2$$

$$\frac{d^2x(t)}{dt^2} - \left|\frac{dx(t)}{dt}\right| + x(t) = 1$$

Ordinary Differential Equations (ODEs)

Chapter1

- Auxiliary Conditions:

ODE

Auxiliary Conditions

Initial Conditions	Boundary Conditions
All conditions are at one point of the independent variable	The conditions are not at one point of the independent variable

IVP

-Boundary-Value and Initial value Problems:

Initial-Value Problems

- The auxiliary conditions are at one point of the independent variable

$$\ddot{x} + 2\dot{x} + x = e^{-2t}$$
$$x(0) = 1, \quad \dot{x}(0) = 2.5$$

same

Boundary-Value Problems

BVP

- The auxiliary conditions are not at one point of the independent variable
- More difficult to solve than initial value problems

$$\ddot{x} + 2\dot{x} + x = e^{-2t}$$
$$x(0) = 1, \quad x(2) = 1.5$$

different

- Classification of ODEs:

ODEs can be classified in different ways:

- Order
  - First order ODE
  - Second order ODE
  - N^{th} order ODE
- Linearity
  - Linear ODE
  - Nonlinear ODE
- Auxiliary conditions
  - Initial value problems
  - Boundary value problems

- Solutions of Ordinary Differential Equations

For example, this function

$x(t) = \cos(2t)$

is a solution to the ODE

$$\frac{d^2x(t)}{dt^2} + 4x(t) = 0$$

Is it unique?

All functions of the form $x(t) = \cos(2t + c)$ (where c is a real constant) are solutions.

- Uniqueness of a Solution

In order to uniquely specify a solution to an n^{th} order differential equation we need n conditions

$$\frac{d^2x(t)}{dt^2} + 4x(t) = 0 \qquad \text{Second order ODE}$$

$$\boxed{\begin{aligned} x(0) &= a \\ \dot{x}(0) &= b \end{aligned}}$$

Two conditions are needed to uniquely specify the solution

Classification of ODEs

ODEs can be classified in different ways:

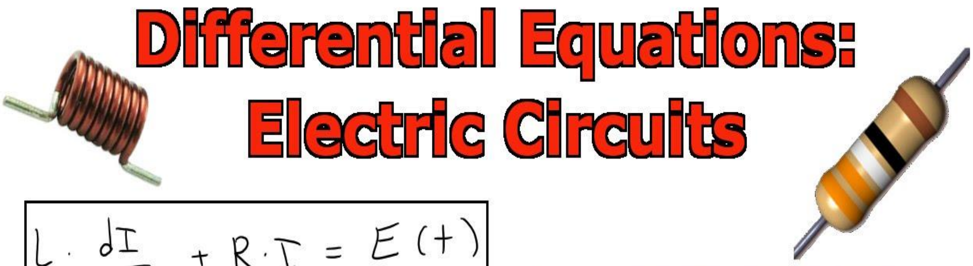
- **Order**
  - First order ODE
  - Second order ODE
  - N^{th} order ODE
- **Linearity**
  - Linear ODE
  - Nonlinear ODE
- **Auxiliary conditions**
  - Initial value problems
  - Boundary value problems

Ordinary Differential Equations (ODEs)

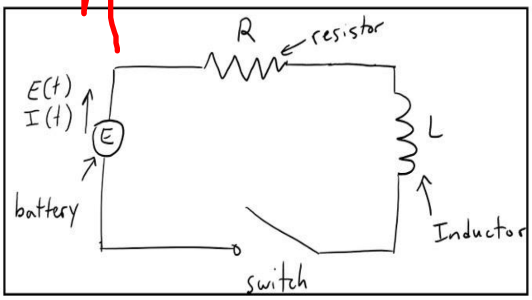
Chapter1

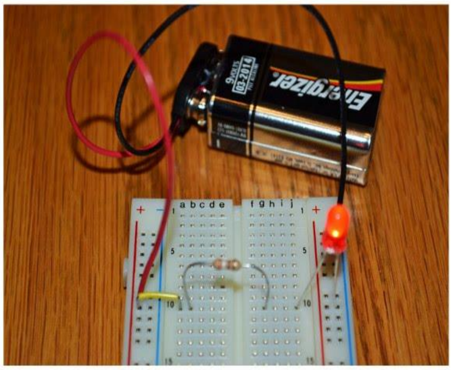
- Applications of Differential equations:

- **Electric Circuits:-**



$$L \cdot \frac{dI}{dt} + R \cdot I = E(t)$$

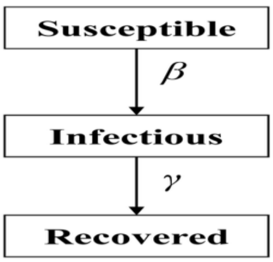




- **Biological Systems:-**

The ***SIR*** epidemic model is one of the simplest compartmental models, and many models are derivations of this basic form. The model consists of three compartments–*S* for the number susceptible, *I* for the number of infectious, and *R* for the number recovered (or immune).

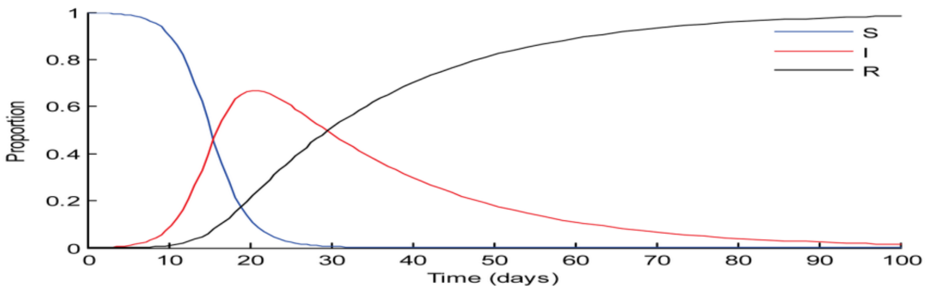
SIR



$$\frac{dS}{dt} = -\beta S I$$

$$\frac{dI}{dt} = \beta S I - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$



Chapter 2

First-Order Differential Equations

First-order differential equations

Separable Differential Equations

- *Objectives of Lesson*

- Differential Equation of first Order and first Degree
- Method of Solution: Separation of Variables
- Equations Reducible to Variable Separable Form
- Class Exercise

Definition of DE of first order ✓

A differential equation of the first order and first degree contains independent variable x, dependent variable y and its derivative

$$\frac{dy}{dx} \quad \text{i.e.} \quad \frac{dy}{dx} = f(x, y) \quad \text{or} \quad f\left(x, y, \frac{dy}{dx}\right) = 0$$

where f (x, y) is the function of x and y.

For example : $xy(y+1)\,dy = \left(x^2 + 1\right)dx, \quad \frac{dy}{dx} = \frac{x+y}{x-y},$

$$\frac{dy}{dx} + y = \sin x \quad \text{etc.}$$

Separation of Variables ✓

A **separable** differential equation is one that can be written so that the independent variable terms (along with its differential) are collected to one side of the equal sign, and the dependent variable terms (and its differential) to the other.

Example 1: $y' = xy^2$ is separable. It is first written as $\frac{dy}{dx} = xy^2$, then "separated":

$$\frac{dy}{y^2} = x\,dx.$$

This is separated as $\frac{dy}{y^2} = x\,dx$.

$\int \frac{dy}{y^2} = \int x\,dx$ Integrate both sides.

$-\frac{1}{y} = \frac{1}{2}x^2 + C$ Don't forget the constant of integration.

$\frac{1}{y} = C - \frac{1}{2}x^2$ Negate. The C "absorbs" the negative.

$y = \frac{1}{C - \frac{1}{2}x^2} = \frac{2}{C - x^2}$ Solve for y. Note that $2C$ is written as C.

Example 2: Solve the IVP: $y' = x + xy, y(0) = 3$.

Solution

Write y' as $\frac{dy}{dx}$: $\frac{dy}{dx} = x + xy$

Factor: $\frac{dy}{dx} = x(1+y)$

Separate: $\frac{dy}{1+y} = x\,dx$ $(y \neq -1)$

Integrate: $\int \frac{dy}{1+y} = \int x\,dx \rightarrow \ln|1+y| = \frac{1}{2}x^2 + C$

Isolate y: $|1+y| = e^{0.5x^2+C} \rightarrow |1+y| = Ce^{0.5x^2}$

$1 + y = \pm Ce^{0.5x^2} = Ce^{0.5x^2}$ $(\pm C = C)$

Thus, $y = Ce^{0.5x^2} - 1$ is the general solution of $y' = x + xy$.

- The constant of integration C is just a generic constant at this point. It absorbs all constants that come near it, so to speak. For example, $e^C = C$, $-C = C$, $2C = C$, $\frac{1}{C} = C$, and so on.
- The C can be determined with an initial condition. For example, suppose we have $y' = x + xy$ with $y(0) = 3$. The general solution is $y = C^{0.5x^2} - 1$. To find C, let $x = 0$ and $y = 3$:

$$3 = Ce^{0.5(0)^2} - 1 \quad \rightarrow \quad 3 = C - 1 \quad \rightarrow \quad C = 4.$$

Thus, the particular solution is $y = 4e^{0.5x^2} - 1$.

Example 3

- Solve the following first order nonlinear equation:

$$\frac{dy}{dx} = \frac{x^2 + 1}{y^2 - 1}$$

Solution

Separating variables, and using calculus, we obtain

$$(y^2 - 1)\, dy = (x^2 + 1)\, dx$$
$$\int (y^2 - 1)\, dy = \int (x^2 + 1)\, dx$$
$$\frac{1}{3} y^3 - y = \frac{1}{3} x^3 + x + C$$
$$y^3 - 3y = x^3 + 3x + C$$

Example 4:

- Solve the following first order nonlinear equation:

$$\frac{dy}{dx} = \frac{3x^2 + 4x + 2}{2(y-1)}, \textit{where } y(0) = -1$$

Solution

$$2(y-1)dy = (3x^2 + 4x + 2)dx$$
$$2\int (y-1)dy = \int (3x^2 + 4x + 2)dx$$
$$y^2 - 2y = x^3 + 2x^2 + 2x + C$$

- The equation above defines the solution *y* implicitly. An explicit expression for the solution can be found in this case:

$$y^2 - 2y - (x^3 + 2x^2 + 2x + C) = 0 \;\Rightarrow\; y = \frac{2 \pm \sqrt{4 + 4(x^3 + 2x^2 + 2x + C)}}{2}$$
$$y = 1 \pm \sqrt{x^3 + 2x^2 + 2x + C_1}$$

- Suppose we seek a solution satisfying *y*(0) = -1. Using the implicit expression of *y*, we obtain
- Thus the implicit equation defining *y* is

$$y^2 - 2y = x^3 + 2x^2 + 2x + C$$
$$(-1)^2 - 2(-1) = C \;\Rightarrow\; C = 3$$

- Using explicit expression of *y*,

$$y^2 - 2y = x^3 + 2x^2 + 2x + 3$$

Example 5

- Consider the following initial value problem:

$$y' = \frac{y\cos x}{1+3y^3}, \quad y(0)=1$$

Solution

Separating variables and using calculus, we obtain

$$\frac{1+3y^3}{y}dy = \cos x dx$$

$$\int\left(\frac{1}{y}+3y^2\right)dy = \int \cos x dx$$

$$\ln|y| + y^3 = \sin x + C$$

- Using the initial condition, it follows that

$$\ln y + y^3 = \sin x + 1$$

- Thus

$$y' = \frac{y\cos x}{1+3y^3}, \quad y(0)=1 \Rightarrow \ln y + y^3 = \sin x + 1$$

Example 6

Solve the differential equation $\frac{dy}{dx} = x - 1 + xy - y$.

Solution: The given differential equation is $\frac{dy}{dx} = x - 1 + xy - y$.

$$\Rightarrow \frac{dy}{dx} = (x-1) + y(x-1)$$

$$\Rightarrow \frac{dy}{y+1} = (x-1)dx \qquad \text{[Variable separable form]}$$

Integrating both sides, we get

$$\int \frac{dy}{y+1} = \int (x-1)dx$$

$$\Rightarrow \log_e|y+1| = \frac{x^2}{2} - x + C$$

Example 7

$$\tan y \frac{dy}{dx} = \frac{x\cos^4 y}{1-x^2}$$

$\longrightarrow$ Separation

$$\frac{\tan y}{\cos^2 y}dy = \frac{x}{1-x^2}dx \longrightarrow \frac{\sin y}{\cos^3 y}dy = \frac{x}{1-x^2}dx$$

$$\int \sin y \cos^{-3} y \, dy = \int \frac{x}{1-x^2} dx$$

$$-\int -\sin y \cos^{-3} y \, dy = \frac{-1}{2}\int \frac{-2 * x}{1-x^2} dx$$

$$\frac{-\cos^{-2} y}{-2} = \frac{-1}{2} \ln(1-x^2) + \frac{C}{2}$$

$$\cos^{-2} y = -\ln(1-x^2) + C$$

$$\boxed{\sec^2 y + \ln(1-x^2) = C}$$

(i) $(e^{2y}-y)\cosh x \frac{dy}{dx} = e^{y+x}\sinh 2x$

$$\int \frac{(e^{2y}-y)}{e^y} dy = \int \frac{\sinh 2x \, e^x}{\cosh x} dx$$

$$\int (e^y - ye^{-y}) dy = \int \frac{2\sinh x \cancel{\cosh x}\, e^x}{\cancel{\cosh x}} dx$$

$$e^y + ye^{-y} - \int e^{-y} dy = \int \cancel{2} e^x \frac{e^x - e^{-x}}{\cancel{2}} = \int e^{2x} - 1$$

$$\boxed{e^y + ye^{-y} + e^{-y} = \frac{1}{2}e^{2x} - x + C}$$

Reducible to Variable Separable Form

Differential equation of the form

$$\frac{dy}{dx} = f(ax + by + c)$$

Substitute ax + by + c = v to reducing variable separable form.

Example 8

Solve the differential equation: $(x + y)^2 \frac{dy}{dx} = 1$

Solution: We have $(x+y)^2 \frac{dy}{dx} = 1$ ...(i)

Putting $x + y = v$ and $1 + \frac{dy}{dx} = \frac{dv}{dx} \Rightarrow \frac{dy}{dx} = \frac{dv}{dx} - 1$

in (i), we get

$$v^2\left(\frac{dv}{dx} - 1\right) = 1 \Rightarrow v^2 \frac{dv}{dx} = 1 + v^2$$

$$\Rightarrow \frac{v^2}{1+v^2} dv = dx \quad \Rightarrow \frac{1+v^2-1}{1+v^2} dv = dx$$

$$\Rightarrow \left(1 - \frac{1}{1+v^2}\right) dv = dx$$

$$\Rightarrow \int\left(1 - \frac{1}{1+v^2}\right) dv = \int dx \Rightarrow v - \tan^{-1} v = x + C$$

$$\Rightarrow (x+y) - \tan^{-1}(x+y) = x + C \quad \Rightarrow y - \tan^{-1}(x+y) = C$$

Homogenous Differential Equations

Homogeneous Function:-

A function $f(x,y)$ is called a homogenous function of degree n if $f(tx,ty) = t^n f(x,y)$

Examples:

$g(x, y) = x^2 - xy + y^2$ is a homogenous function of degree 2

Since $g(tx,ty) = t^2x^2 - (tx)(ty) + t^2y^2 = t^2(x^2 - xy + y^2) = t^2 g(x,y)$

$Q(x,y) = x^3 sin\left(\frac{x}{y}\right) - \sqrt{x^2 - 4xy}$ is homogenous of degree 2

Method of Solution

(1) Write the differential equation in the form $\frac{dy}{dx} = \frac{f(x, y)}{g(x, y)}$

(2) Substitute $y = vx$ and $\frac{dy}{dx} = v + x\frac{dv}{dx}$ in the equation.

(3) The equation reduces to the form $v + x\frac{dv}{dx} = F(v)$

(4) Separate the variables of v and x.

(5) Integrate both sides to obtain the solution in terms of v and x.

(6) Replace v by $\frac{y}{x}$ to get the solution

Note: you could perform the same steps but with other assumption:

$x = u\,y$ and $\frac{dx}{dy} = u + y\frac{du}{dy}$

Example 1

Solve the differential equation $x\dfrac{dy}{dx} = x + y$.

Solution : The given differential equation is $x\dfrac{dy}{dx} = x + y$.

$$\Rightarrow \frac{dy}{dx} = \frac{x+y}{x} \quad \text{...(i)}$$

It is a homogeneous differential equation of degree 1.

Putting $y = vx$ and $\dfrac{dy}{dx} = v + x\dfrac{dv}{dx}$ ...(i), we get

$$v + x\frac{dv}{dx} = \frac{x+vx}{x}$$

$$\Rightarrow v + x\frac{dv}{dx} = 1 + v \quad \Rightarrow x\frac{dv}{dx} = 1 \quad \Rightarrow dv = \frac{1}{x}dx$$

Integrating both sides, we get

$$\int dv = \int \frac{1}{x}dx \quad \Rightarrow v = \log_e|x| + C$$

$$\Rightarrow \frac{y}{x} = \log_e|x| + C \quad [\because y = vx] \quad \Rightarrow y = x\log_e|x| + Cx$$

Example 2

$$\left(x^3 + y^2\sqrt{x^2+y^2}\right)dx - \left(xy\sqrt{x^2+y^2}\right)dy = 0$$

$\longrightarrow$ No Separation

$$\left(t^3x^3 + t^2y^2\sqrt{t^2x^2 + t^2y^2}\right)dx - \left(tx\,ty\sqrt{t^2x^2 + t^2y^2}\right)dy = 0$$

$$\left(t^3x^3 + t^2y^2\sqrt{t^2(x^2 + y^2)}\right)dx - \left(t^2xy\sqrt{t^2(x^2 + y^2)}\right)dy = 0$$

$$\left(t^3x^3 + t^2y^2\sqrt{t^2}\sqrt{x^2 + y^2}\right)dx - \left(t^2xy\sqrt{t^2}\sqrt{x^2 + y^2}\right)dy = 0$$

$$\left(t^3x^3 + t^2y^2t\sqrt{x^2 + y^2}\right)dx - \left(t^2xy\,t\sqrt{x^2 + y^2}\right)dy = 0$$

$$t^3\left(x^3 + y^2\sqrt{x^2 + y^2}\right)dx - t^3\left(xy\sqrt{x^2 + y^2}\right)dy = 0$$

$\longrightarrow$ Homogenous DE of 3rd order

$$\frac{dy}{dx} = \frac{x^3 + y^2\sqrt{x^2+y^2}}{xy\sqrt{x^2+y^2}}$$

Let $\quad y = u\,x \quad$ and $\quad \dfrac{dy}{dx} = u + x\,\dfrac{du}{dx}$

$$u + x\frac{du}{dx} = \frac{x^3 + u^2x^2\sqrt{x^2+u^2x^2}}{x^2u\sqrt{x^2+u^2x^2}} = \frac{x^3+u^2x^2\sqrt{x^2(1+u^2)}}{x^2u\sqrt{x^2(1+u^2)}}$$

$$= \frac{x^3+u^2x^2\sqrt{x^2}\sqrt{(1+u^2)}}{x^2u\sqrt{x^2}\sqrt{(1+u^2)}} = \frac{x^3+u^2x^3\sqrt{(1+u^2)}}{x^3u\sqrt{(1+u^2)}}$$

$$u + x\frac{du}{dx} = \frac{1+u^2\sqrt{(1+u^2)}}{u\sqrt{(1+u^2)}} \quad \rightarrow \quad x\frac{du}{dx} = \frac{1+u^2\sqrt{(1+u^2)}}{u\sqrt{(1+u^2)}} - u$$

$$x\frac{du}{dx} = \frac{1+u^2\sqrt{1+u^2}}{u\sqrt{1+u^2}} - u * \frac{u\sqrt{1+u^2}}{u\sqrt{1+u^2}}$$

$$x\frac{du}{dx} = \frac{1 + u^2\sqrt{1+u^2} - u^2\sqrt{1+u^2}}{u\sqrt{1+u^2}}$$

$$x\frac{du}{dx} = \frac{1}{u\sqrt{1+u^2}} \quad \rightarrow \quad \int u\sqrt{1+u^2}\,du = \int \frac{1}{x}\,dx$$

$$\frac{1}{2}\int 2 * u\,(1+u^2)^{0.5}\,du = \ln x + \ln C$$

$$\frac{1}{2}\,\frac{(1+u^2)^{1.5}}{1.5} = \ln x + \ln C$$

$$\boxed{\frac{1}{3}\left(1+\left(\frac{y}{x}\right)^2\right)^{1.5} = \ln C\,x}$$

Example 3

10. $(y^2 - x^2 e^{-\frac{y}{x}})\,dx - xy\,dy = 0$

$\longrightarrow$ No Separation

$$(t^2y^2 - t^2x^2 e^{-\frac{ty}{tx}})\,dx - t^2xy\,dy = 0$$

$$t^2(y^2 - x^2 e^{-\frac{y}{x}})\,dx - t^2xy\,dy = 0$$

$\longrightarrow$ Homogenous DE of 2nd order

$$\frac{dy}{dx} = \frac{y^2 - x^2 e^{-\frac{y}{x}}}{xy}$$

Let $y = ux$ and $\frac{dy}{dx} = u + x\frac{du}{dx}$

$$u + x\frac{du}{dx} = \frac{u^2x^2 - x^2 e^{-u}}{ux^2} = \frac{u^2 - e^{-u}}{u}$$

$$x\frac{du}{dx} = \frac{u^2 - e^{-u}}{u} - u * \frac{u}{u} = \frac{u^2 - e^{-u} - u^2}{u}$$

$$x\frac{du}{dx} = \frac{-e^{-u}}{u} \quad \rightarrow \quad \int u e^{u}\,du = \int \frac{-1}{x}\,dx$$

$$u e^{u} - \int e^{u}\,du = -\ln x + \ln C$$

$$u e^{u} - e^{u} = \ln\frac{C}{x}$$

$$\boxed{\frac{y}{x} e^{\frac{y}{x}} - e^{\frac{y}{x}} = \ln\frac{C}{x}}$$

Exact Differential Equations & Integrating Factors
Exact and Integrating factor

$$M(x, y)\,dx + N(x, y)\,dy = 0$$

An equation is said to be solved using exact method only if:

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

$$M_y = N_x$$

Steps to solve:

Let the solution of the differential equation, F, = C $\longrightarrow$ F = C

Then :

F $= \int M(x, y)\,dx$

$=$ ---------- $+ g(y)$

$$\frac{\partial F}{\partial y} \equiv N(x, y)$$

---------- $+ g'(y) \equiv$ $N(x, y)$

Then find the equivalent of $g'(y)$ in the right hand side thus,

$$g(y) = \int g'(y)\,dy$$

F $=$ ---------- $+ g(y) = C$

F $= \int N(x, y)\,dy$

$=$ ---------- $+ g(x)$

$$\frac{\partial F}{\partial x} \equiv M(x, y)$$

---------- $+ g'(x) \equiv$ $M(x, y)$

Then find the equivalent of $g'(x)$ in the right hand side thus,

$$g(x) = \int g'(x)\,dx$$

F $=$ ---------- $+ g(x) =$

Solve the differential equation $(2xy - 3x^2)\,dx + (x^2 - 2y)\,dy = 0$.

Solution The given differential equation is exact because

$$\frac{\partial M}{\partial y} = \frac{\partial}{\partial y}[2xy - 3x^2] = 2x = \frac{\partial N}{\partial x} = \frac{\partial}{\partial x}[x^2 - 2y].$$

The general solution, $f(x, y) = C$, is given by

$$\begin{aligned} f(x, y) &= \int M(x, y)\,dx \\ &= \int (2xy - 3x^2)\,dx = x^2y - x^3 + g(y). \end{aligned}$$

In Section 14.1, you determined $g(y)$ by integrating $N(x, y)$ with respect to y and reconciling the two expressions for $f(x, y)$. An alternative method is to partially differentiate this version of $f(x, y)$ with respect to y and compare the result with $N(x, y)$. In other words,

$$f_y(x, y) = \frac{\partial}{\partial y}[x^2y - x^3 + g(y)] = x^2 + g'(y) = \overbrace{x^2 - 2y}^{N(x,\,y)}.$$

$$\boxed{g'(y) = -2y}$$

Thus, $g'(y) = -2y$, and it follows that $g(y) = -y^2 + C_1$. Therefore,

$$f(x, y) = x^2y - x^3 - y^2 + C_1$$

and the general solution is $x^2y - x^3 - y^2 = C$. Figure 15.1 shows the solution curves that correspond to $C = 1, 10, 100$, and 1000.

Find the particular solution of

$$(\cos x - x\sin x + y^2)\,dx + 2xy\,dy = 0$$

that satisfies the initial condition $y = 1$ when $x = \pi$.

Solution The differential equation is exact because

$$\overbrace{\frac{\partial}{\partial y}[\cos x - x\sin x + y^2]}^{\frac{\partial M}{\partial y}} = 2y = \overbrace{\frac{\partial}{\partial x}[2xy]}^{\frac{\partial N}{\partial x}}.$$

Because $N(x, y)$ is simpler than $M(x, y)$, it is better to begin by integrating $N(x, y)$.

$$f(x, y) = \int N(x, y)\,dy = \int 2xy\,dy = xy^2 + g(x)$$

$$f_x(x, y) = \frac{\partial}{\partial x}[xy^2 + g(x)] = y^2 + g'(x) = \overbrace{\cos x - x\sin x + y^2}^{M(x,y)}$$

$$\boxed{g'(x) = \cos x - x\sin x}$$

Thus, $g'(x) = \cos x - x\sin x$ and

$$g(x) = \int (\cos x - x\sin x)\,dx$$
$$= x\cos x + C_1$$

which implies that $f(x, y) = xy^2 + x\cos x + C_1$, and the general solution is

$$xy^2 + x\cos x = C. \qquad \text{General solution}$$

Applying the given initial condition produces

$$\pi(1)^2 + \pi\cos\pi = C$$

which implies that $C = 0$. Hence, the particular solution is

$$xy^2 + x\cos x = 0. \qquad \text{Particular solution}$$

The graph of the particular solution is shown in Figure 15.3. Notice that the graph consists of two parts: the ovals are given by $y^2 + \cos x = 0$, and the y-axis is given by $x = 0$.

- $e^{x^2}(2x^2y - xy^2 + y)\,dx - e^{x^2}(y-x)dy = 0$

$\longrightarrow$ No separation

$\longrightarrow$ No Homogenous

$$e^{x^2}(2x^2y - xy^2 + y)\,dx + e^{x^2}(x-y)dy = 0$$

$$M_y = e^{x^2}(2x^2 - 2xy + 1)$$

$$N_x = e^{x^2} + 2x^2e^{x^2} - 2xye^{x^2}$$

$\longrightarrow M_y = N_x \longrightarrow$ *Exact*

F = C

$$F = \int N(x,y)\,dy = \int e^{x^2}(x-y)\,dy$$

$$F = xye^{x^2} - \frac{y^2}{2}e^{x^2} + g(x)$$

$$\frac{\partial F}{\partial x} \equiv M(x,y)$$

$$ye^{x^2} + 2x^2ye^{x^2} - xy^2e^{x^2} + g'(x) \equiv e^{x^2}(2x^2y - xy^2 + y)$$

$$\therefore\ g'(y) \equiv 0 \longrightarrow g(y) \equiv \int 0\,dy = C$$

$$F = xye^{x^2} - \frac{y^2}{2}e^{x^2} = C$$

$$\boxed{xye^{x^2} - \frac{y^2}{2}e^{x^2} = C}$$

$$\left(e^{2y} - y\cos(xy)\right)dx + \left(2x\,e^{2y} - x\cos(xy) + 2y\right)dy = 0 \quad , \; y(0) = 2$$

$\longrightarrow$ No Separation

$\longrightarrow$ No Homogenous

$M_y = 2\,e^{2y} - \cos(xy) + xy\sin(xy)$, $N_x = 2\,e^{2y} - \cos(xy) + xy\sin(xy)$

$\longrightarrow$ Exact

$F = C$

$$F = \int \left(e^{2y} - y\cos(xy)\right)dx = x\,e^{2y} - \sin(xy) + g(y)$$

$$\frac{\partial F}{\partial y} \equiv N(x, y)$$

$$2x\,e^{2y} - x\cos(xy) + g'(y) \equiv 2x\,e^{2y} - x\cos(xy) + 2y$$

$$\therefore\; g'(y) = 2y \quad\longrightarrow\quad g(y) = \int 2y\,dy = \frac{y^2}{2}$$

$$F = x\,e^{2y} - \sin(xy) + \frac{y^2}{2} = C$$

By substituting $y = 2$ and $x = 0$ $\longrightarrow$ $C = 2$

$$\boxed{x\,e^{2y} - \sin(xy) + \frac{y^2}{2} = 2}$$

Integrating Factors

Integrating Factor $\mu(x,y)$

If $M_y \neq N_x$ and there is a slight difference between them, you have to drive an integrating factor using either of the following ways:

If $\dfrac{M_y - N_x}{N} =$ $f(x)$ only

$$\mu(x) = e^{\int f(x)\,dx}$$

If $\dfrac{N_x - M_y}{M} =$ $f(y)$ only

$$\mu(y) = e^{\int f(y)\,dy}$$

Then multiply the original equation by μ getting a new equation in the form of :

$$\mu\, M(x,y)\,dx + \mu\, N(x,y)\,dy = 0$$

OR

$$M`(x,y)\,dx + N`(x,y)\,dy = 0$$

Then solve this new differential equation using the normal exact method.

Note: To check on your solutions, you must find that:

$$M_y` = N_x`$$

Examples:-

Solve the differential equation $(y^2 - x)\,dx + 2y\,dy = 0$.

Solution The given equation is not exact because $M_y(x, y) = 2y$ and $N_x(x, y) = 0$. However, because

$$\frac{M_y(x, y) - N_x(x, y)}{N(x, y)} = \frac{2y - 0}{2y} = 1 = h(x)$$

it follows that $e^{\int h(x)\,dx} = e^{\int dx} = e^x$ is an integrating factor. Multiplying the given differential equation by e^x produces the exact differential equation

$$(y^2e^x - xe^x)\,dx + 2ye^x\,dy = 0$$

whose solution is obtained as follows.

$$f(x, y) = \int N(x, y)\,dy = \int 2ye^x\,dy = y^2e^x + g(x)$$

$$f_x(x, y) = y^2e^x + g'(x) = \overbrace{y^2e^x - xe^x}^{M(x,y)}$$

$$\boxed{g'(x) = -xe^x}$$

Therefore, $g'(x) = -xe^x$ and $g(x) = -xe^x + e^x + C_1$, which implies that

$$f(x, y) = y^2e^x - xe^x + e^x + C_1.$$

The general solution is $y^2e^x - xe^x + e^x = C$, or $y^2 - x + 1 = Ce^{-x}$.

$$(2y^2x - yx^2 + x)dy + (y - x)dx = 0$$

$\longrightarrow$ No Separation

$\longrightarrow$ No Homogenous

$$(y - x)dx + (2y^2x - yx^2 + x)dy = 0$$

$M_y = 1$, $N_x = 2y^2 - 2xy + 1$ $\longrightarrow$ Not Exact

Find an integrating factor:

$$\frac{N_x - M_y}{M} = \frac{2y^2 - 2xy + 1 - 1}{y - x} = \frac{2y^2 - 2xy}{y - x} = \frac{2y(y - x)}{y - x} = 2y \longrightarrow f(y)\ only$$

$$\mu(y) = e^{\int f(y)\,dy} = e^{\int 2y\,dy} = e^{y^2}$$

Therefore, the new equation will be:

$$e^{y^2}(y - x)dx + e^{y^2}(2y^2x - yx^2 + x)dy = 0$$

$$\left(ye^{y^2} - xe^{y^2}\right)dx + \left(2y^2 x e^{y^2} - y x^2 e^{y^2} + x e^{y^2}\right)dy = 0$$

$F = C$

$$F = \int \left(ye^{y^2} - xe^{y^2}\right)dx = x y e^{y^2} - \frac{x^2}{2} e^{y^2} + g(y)$$

$$\frac{\partial F}{\partial y} \equiv N(x, y)$$

$$x\left(y e^{y^2} * 2y + e^{y^2}\right) - x^2 y e^{x^2} + g'(y) \equiv 2y^2 x e^{y^2} - y x^2 e^{y^2} + x e^{y^2}$$

$$\therefore g'(y) = 0 \longrightarrow g(y) = \int 0\, dy = C$$

$$\boxed{F = x y e^{y^2} - \frac{x^2}{2} e^{y^2} = C}$$

- $y' \tan x \sin 2y = \sin^2 x + \cos^2 y$

$\longrightarrow$ No separation

$\longrightarrow$ No Homogenous

$$\frac{dy}{dx} \tan x \sin 2y = \sin^2 x + \cos^2 y$$

$$(\sin^2 x + \cos^2 y)\, dx + (-\tan x \sin 2y)\, dy = 0$$

$M_y = -2 \cos y \sin y$

$= -\sin 2y$

$N_x = -\sec^2 x \sin 2y$

$\longrightarrow M_y \neq N_x \longrightarrow$ *Not Exact*

So try to find an integrating factor using either of the above cases:

$$\frac{M_y - N_x}{N} = \frac{-\sin 2y + \sec^2 x \sin 2y}{-\tan x \sin 2y} = \frac{-\sin 2y + \sec^2 x \sin 2y}{-\tan x \sin 2y}$$

$$= \frac{\sin 2y\,(-1 + \sec^2 x)}{-\tan x \sin 2y} = \frac{(-1 + \sec^2 x)}{-\tan x} = \frac{\tan^2 x}{-\tan x} = -\tan x$$

Since that the result is function in x only, then

$$\mu = e^{\int -\tan x \, dx} = e^{\int \frac{-\sin x}{\cos x} dx} = e^{\ln \cos x} = \cos x$$

Therefore, the new equation will be:

$$\cos x\,(\sin^2 x + \cos^2 y)\,dx + \cos x\,(-\tan x \sin 2y)\,dy = 0$$

$$(\cos x \sin^2 x + \cos x \cos^2 y)\,dx + (-\sin x \sin 2y)\,dy = 0$$

$$(\cos x \sin^2 x + \cos x \cos^2 y)\,dx + (-2 \sin x \sin y \cos y)\,dy = 0$$

F = C

$$F = \int N(x, y)\,dy = \int (-2 \sin x \sin y \cos y)\,dy$$

$$F = \sin x \cos^2 y + g(x)$$

$$\frac{\partial F}{\partial x} \equiv M(x, y)$$

$$\cos x \cos^2 y + g'(x) \equiv \cos x \sin^2 x + \cos x \cos^2 y$$

$$\therefore \quad g'(x) \equiv \cos x \sin^2 x \quad \longrightarrow \quad g(x) \equiv \int \cos x \sin^2 x \, dx$$

$$g(x) \equiv \frac{\sin^3 x}{3}$$

$$F = \sin x \cos^2 y + \frac{\sin^3 x}{3} = C$$

$$\boxed{\sin x \cos^2 y + \frac{\sin^3 x}{3} = C}$$

Linear & and Bernoulli Differential Equations

To say that a 1st order differential equation can be solved using Linear Method, the form of the equation must be either:

$$\frac{dy}{dx} + P(x)\,y = Q(x) \quad OR \quad \frac{dx}{dy} + P(y)\,x = Q(y)$$

Solving Criteria:

$$OR$$

$\frac{dy}{dx} + P(x)\,y = Q(x)$	$\frac{dx}{dy} + P(y)\,x = Q(y)$
Let	Let
$\mu\,(x) = e^{\int p(x)\,dx}$	$\mu\,(y) = e^{\int p(y)\,dy}$
and the solution will be	and the solution will be
$\mu\,(x) * y = \int \mu\,(x) * Q(x)\,dx$	$\mu\,(y) * x = \int \mu\,(y) * Q(y)\,dy$

Integration Factor:-

There is a process by which most first-order linear differential equations can be solved. This uses an **integration factor**, denoted $\mu(x)$ (Greek letter "mu").

The differential equation must be in the form

$$y' + f(x)y = g(x).$$

To find $\mu(x)$, we perform the following process (next slide)

Starting with $y' + f(x)y = g(x)$, multiply both sides by $\mu(x)$:

$$\mu(x)y' + \mu(x)f(x)y = \mu(x)g(x)$$

The left side is a product-rule derivative of $(\mu(x)y)$:

$$(\mu(x)y)' = \mu(x)y' + \mu'(x)y.$$

Thus, we have $\mu(x)y' + \mu'(x)y = \mu(x)y' + \mu(x)f(x)y$.

This forces $\mu'(x) = \mu(x)f(x)$. (*next slide*)

Now we find $\mu(x)$. From the last slide, we had $\mu'(x) = \mu(x)f(x)$.

This is a separable differential equation… so separate:

$$\frac{d\mu}{\mu(x)} = f(x)dx.$$

Integrating both sides, we have

$$\int \frac{d\mu}{\mu(x)} = \int f(x)dx\,.$$

(*next slide*)

After integration, we have

$$\ln \mu(x) = \int f(x)\, dx + C.$$

Here, we only need one form of the antiderivative, so we let $C = 0$. Taking base-e on both sides, we now know $\mu(x)$:

$$\mu(x) = e^{\int f(x)dx}.$$

Remark: you don't need to do all those steps each time. Just remember that if you have a differential equation of the form $y' + f(x)y = g(x)$, then find $\mu(x) = e^{\int f(x)dx}$.

Example 1

Find the general solution of the following ODE:

$$\frac{1}{x}\frac{dy}{dx} - \frac{2y}{x^2} = x\cos x$$

Solution

So $\frac{dy}{dx} + \frac{-2}{x}y = x^2\cos x$ is a linear ODE where $P(x) = -\frac{2}{x}$ and $Q(x) = x^2\cos x$

The integrating factor $\mu(x) = e^{\int P(x)dx} = e^{\int \frac{-2}{x}dx} = e^{-2\ln x} = x^{-2}$

The solution is $\mu(x)\, y = \int \mu(x)Q(x)dx$

$$x^{-2}\, y = \int \cos x\, dx = \sin x + C$$

$$y = x^2 \sin x + x^2 C$$

Example 2

A rock contains two radioactive isotopes, RA_1 and RA_2, that belong to the same radioactive series; that is, RA_1 decays into RA_2, which then decays into stable atoms. Assume that the rate at which RA_1 decays into RA_2 is $50e^{-t}$ kg/sec. Because the rate of decay of RA_2 is proportional to the mass $y(t)$ of RA_2 present, the rate of change in RA_2 is

$$\frac{dy}{dt} = \text{Rate of concentration} - \text{Rate of decay}$$

$$\frac{dy}{dt} = 50e^{-t} - ky$$

Where $k > 0$ is the decay constant. If $k = 2$, $y(0) = 40\, kg$ then find the mass $y(t)$ of RA_2 for $t \geq 0$.

Solution

$$\frac{dy}{dt} + 2y = 50e^{-t} \text{ is linear ODE where } P(t) = 2 \text{ and } Q(t) = 50e^{-t}$$

The integrating factor $\mu(t) = e^{\int P(t)dt} = e^{\int 2\,dt} = e^{2t}$

The solution is $\mu(t)\, y = \int \mu(t) Q(t) dt$

$$e^{2t} y = 50 \int e^{-8t} dt = \frac{-25}{4} e^{-8t} + C$$

$$y = \frac{-25}{4} e^{-8t} + e^{-2t} C$$

Since $y(0) = 40$ then $C = \frac{185}{4}$ so the solution will be

$$y = \frac{-25}{4} e^{-8t} + \frac{185}{4} e^{-2t}$$

Example: Find the general solution of
$y' + \frac{2}{x}y = x.$

(Note that $f(x) = \frac{2}{x}$ and $g(x) = x$ and that $x \neq 0$)

So $\mu(x) = e^{\int \left(\frac{2}{x}\right)dx} = e^{2\ln x} = e^{\ln x^2} = x^2.$

Now, use the formula $y = \frac{\int \mu(x)g(x)dx + C}{\mu(x)}$:

Solve $y' + \frac{2}{x}y = x$. From previous slide, we know that $\mu(x) = x^2$.

Now we use the formula $y = \frac{\int \mu(x)g(x)dx + C}{\mu(x)}$:

$$y = \frac{\int x^2 x\, dx + C}{x^2} = \frac{\int x^3 dx + C}{x^2} = \frac{\frac{1}{4}x^4 + C}{x^2} = \frac{1}{4}x^2 + Cx^{-2}.$$

Thus, $y = \frac{1}{4}x^2 + Cx^{-2}$ is the general solution of
$y' + \frac{2}{x}y = x.$

Check that $y = \frac{1}{4}x^2 + Cx^{-2}$ is the general solution of $y' + \frac{2}{x}y = x.$
First, differentiate y:

$$y' = \frac{1}{2}x - 2Cx^{-3}$$

Now insert y' and y into the differential equation and simplify:

$$\left(\frac{1}{2}x - 2Cx^{-3}\right) + \frac{2}{x}\left(\frac{1}{4}x^2 + Cx^{-2}\right) = x$$
$$\left(\frac{1}{2}x - 2Cx^{-3}\right) + \left(\frac{1}{2}x + 2Cx^{-3}\right) = x$$
$$\left(\frac{1}{2}x + \frac{1}{2}x\right) + (-2Cx^{-3} + 2Cx^{-3}) = x$$
$$x + 0 = x.$$

Ordinary Differential Equations (ODEs)

Chapter 2

Bernoulli Differential equations

Exercise 1.

$$\frac{dy}{dx} + P(x)y = Q(x)y^n$$

DIVIDE by y^n: $\frac{1}{y^n}\frac{dy}{dx} + P(x)y^{1-n} = Q(x)$

SET $z = y^{1-n}$: i.e. $\frac{dz}{dx} = (1-n)y^{(1-n-1)}\frac{dy}{dx}$

i.e. $\frac{1}{(1-n)}\frac{dz}{dx} = \frac{1}{y^n}\frac{dy}{dx}$

SUBSTITUTE $\frac{1}{(1-n)}\frac{dz}{dx} + P(x)z = Q(x)$

i.e. $\boxed{\frac{dz}{dx} + P_1(x)z = Q_1(x)}$ linear in z

where $P_1(x) = (1-n)P(x)$
$Q_1(x) = (1-n)Q(x)$.

OR

$$\frac{dy}{dx} + P(x)\, y = Q(x) * y^n$$

Divide both sides by y^n

Thus the equation will be:

$$y^{-n}\frac{dy}{dx} + P(x)\, y^{1-n} = Q(x)$$

Then substitute:

Let $u = y^{1-n}$ *and*

$$\frac{du}{dx} = (1-n)\, y^{-n}\frac{dy}{dx}$$

$$\therefore\ y^{-n}\frac{dy}{dx} = \frac{1}{(1-n)}\frac{du}{dx}$$

So the equation will be :

$$\frac{1}{(1-n)}\frac{du}{dx} + P(x)\, u = Q(x)$$

Then multiply both sides by (1 − n)

$$\frac{du}{dx} + (1-n)P(x)\, u = (1-n)Q(x)$$

Let $\mu(x) = e^{\int (1-n)\, p(x)\, dx}$

and the solution will be

$$\mu(x) * u = \int \mu(x) * (1-n)\, Q(x)\, dx$$
$$\mu(x) * y^{-n} = \int \mu(x) * (1-n)Q(x)dx$$

$$\frac{dx}{dy} + P(y)\, x = Q(y) * x^n$$

Divide both sides by x^n

Thus the equation will be:

$$x^{-n}\frac{dx}{dy} + P(y)\, x^{1-n} = Q(y)$$

Then substitute:

Let $u = x^{1-n}$ *and*

$$\frac{du}{dy} = (1-n)\ x^{-n}\frac{dx}{dy}$$

$$\therefore\ x^{-n}\frac{dx}{dy} = \frac{1}{(1-n)}\frac{du}{dy}$$

So the equation will be :

$$\frac{1}{(1-n)}\frac{du}{dy} + P(y)\, u = Q(y)$$

Then multiply both sides by (1 − n)

$$\frac{du}{dy} + (1-n)P(y)\, u = (1-n)Q(y)$$

Let $\mu(y) = e^{\int (1-n)\, p(y)\, dy}$

and the solution will be

$$\mu(y) * u = \int \mu(y) * (1-n)Q(y)\, dy$$
$$\mu(y) * x^{-n} = \int \mu(y) * (1-n)Q(y)dy$$

Example

Find the general solution of $2\dfrac{dy}{dx} + \tan x \cdot y = \dfrac{(4x+5)^2}{\cos x} y^3$

Solution

Divide by 2 to get standard form:

$$\frac{dy}{dx} + \frac{1}{2}\tan x \cdot y = \frac{(4x+5)^2}{2\cos x} y^3$$

This is of the form $\dfrac{dy}{dx} + P(x)y = Q(x)y^n$

$$\begin{aligned} \text{where} \qquad P(x) &= \frac{1}{2}\tan x \\ Q(x) &= \frac{(4x+5)^2}{2\cos x} \\ \text{and} \qquad n &= 3 \end{aligned}$$

DIVIDE by y^n: i.e. $\dfrac{1}{y^3}\dfrac{dy}{dx} + \dfrac{1}{2}\tan x \cdot y^{-2} = \dfrac{(4x+5)^2}{2\cos x}$

SET $z = y^{1-n} = y^{-2}$: i.e. $\dfrac{dz}{dx} = -2y^{-3}\dfrac{dy}{dx} = -\dfrac{2}{y^3}\dfrac{dy}{dx}$

$$\therefore \quad -\frac{1}{2}\frac{dz}{dx} + \frac{1}{2}\tan x \cdot z = \frac{(4x+5)^2}{2\cos x}$$

i.e. $$\frac{dz}{dx} - \tan x \cdot z = \frac{(4x+5)^2}{\cos x}$$

Integrating factor, $\mathrm{IF} = e^{\int -\tan x \cdot dx} = e^{\int -\frac{\sin x}{\cos x}\,dx} \left[\equiv e^{\int \frac{f'(x)}{f(x)}\,dx}\right]$

$$= e^{\ln\cos x} = \cos x$$

$$\therefore \quad \cos x \frac{dz}{dx} - \cos x \tan x \cdot z = \cos x \frac{(4x+5)^2}{\cos x}$$

$$\text{i.e.} \quad \cos x \frac{dz}{dx} - \sin x \cdot z = (4x+5)^2$$

$$\text{i.e.} \quad \cos x \cdot z = \int (4x+5)^2 dx$$

$$\text{i.e.} \quad \cos x \cdot z = \left(\frac{1}{4}\right) \cdot \frac{1}{3}(4x+5)^3 + C$$

Use $z = \frac{1}{y^2}$:

$$\frac{\cos x}{y^2} = \frac{1}{12}(4x+5)^3 + C$$

$$\text{i.e.} \quad \frac{1}{y^2} = \frac{1}{12\cos x}(4x+5)^3 + \frac{C}{\cos x}\ .$$

Example

Find the general solution of $x\frac{dy}{dx} + y = y^2x^2 \ln x$

Solution

Standard form: $\frac{dy}{dx} + \left(\frac{1}{x}\right) y = (x \ln x) y^2$

i.e. $P(x) = \frac{1}{x}$, $Q(x) = x \ln x$, $n = 2$

DIVIDE by y^2: $\frac{1}{y^2}\frac{dy}{dx} + \left(\frac{1}{x}\right) y^{-1} = x \ln x$

SET $z = y^{-1}$: $\frac{dz}{dx} = -y^{-2}\frac{dy}{dx} = -\frac{1}{y^2}\frac{dy}{dx}$

$$\therefore \quad -\frac{dz}{dx} + \left(\frac{1}{x}\right) z = x \ln x$$

$$\text{i.e.} \quad \frac{dz}{dx} - \frac{1}{x} \cdot z = -x \ln x$$

Integrating factor: $\quad \text{IF} = e^{-\int \frac{dx}{x}} = e^{-\ln x} = e^{\ln x^{-1}} = \dfrac{1}{x}$

$$\therefore \quad \frac{1}{x}\frac{dz}{dx} - \frac{1}{x^2}z = -\ln x$$

$$\text{i.e.} \quad \tfrac{1}{x}z = -\int \ln x\,dx + C'$$

[Use integration by parts: $\int u\frac{dv}{dx}dx = uv - \int v\frac{du}{dx}dx$,

with $u = \ln x$, $\frac{dv}{dx} = 1$]

$$\text{i.e.} \quad \tfrac{1}{x}z = -\left[x\ln x - \int x \cdot \tfrac{1}{x}dx\right] + C$$

Use $z = \frac{1}{y}$: $\quad \frac{1}{xy} = x(1 - \ln x) + C$.

Example

Find the general solution of $\dfrac{dy}{dx} = y\cot x + y^3 \operatorname{cosec} x$

Solution

Standard form: $\quad \frac{dy}{dx} - (\cot x)\cdot y = (\operatorname{cosec} x)\, y^3$

DIVIDE by y^3: $\quad \frac{1}{y^3}\frac{dy}{dx} - (\cot x)\cdot y^{-2} = \operatorname{cosec} x$

SET $z = y^{-2}$: $\quad \frac{dz}{dx} = -2y^{-3}\frac{dy}{dx} = -2\cdot\frac{1}{y^3}\frac{dy}{dx}$

$$\therefore \quad -\frac{1}{2}\frac{dz}{dx} - \cot x \cdot z = \operatorname{cosec} x$$

$$\text{i.e.} \quad \frac{dz}{dx} + 2\cot x \cdot z = -2\operatorname{cosec} x$$

Integrating factor: $\mathrm{IF} = e^{2\int \frac{\cos x}{\sin x}dx} \equiv e^{2\int \frac{f'(x)}{f(x)}dx} = e^{2\ln(\sin x)} = \sin^2 x.$

$$\therefore \quad \sin^2 x \cdot \frac{dz}{dx} + 2\sin x \cdot \cos x \cdot z = -2\sin x$$

i.e. $\quad \frac{d}{dx}\left[\sin^2 x \cdot z\right] = -2\sin x$

i.e. $\quad z\sin^2 x = (-2)\cdot(-\cos x) + C$

Use $z = \frac{1}{y^2}$: $\quad \frac{\sin^2 x}{y^2} = 2\cos x + C$

i.e. $\quad y^2 = \frac{\sin^2 x}{2\cos x + C}$.

Example

A 30-volt battery is applied to R-L series circuit with R=50 ohm and L=0.1 henry. Find the current $i(t) if\ i(0) = 0$. Determine the time at which $i = 0.25\, i_{ss}$.

Solution

$$L\frac{di(t)}{dt} + Ri(t) = E(t)$$

$$\frac{di(t)}{dt} + \frac{R}{L}i(t) = \frac{E(t)}{L}$$

$$\frac{di(t)}{dt} + \frac{50}{0.1}i(t) = \frac{30}{0.1}$$

$$\frac{di(t)}{dt} + 500\, i(t) = 300$$

$$\mu = e^{\int 500dt} = e^{500t}$$

$$e^{500t}\, i(t) = \int 300\, e^{500t}\, dt = \frac{3}{5}e^{500t} + C$$

$$i(t) = \frac{3}{5} + Ce^{-500t}$$

$$since\ i(0) = 0, \text{then} \quad C = \frac{-3}{5}$$

$$i(t) = \frac{3}{5} - \frac{3}{5}e^{-500t}$$

At $i(t) = 0.25\, i_{ss} = 0.25\, \left(\frac{3}{5}\right) = \frac{3}{20}$then $\frac{3}{20} = \frac{3}{5}\ (1 - e^{-500t})$

Then $t = 5.75 \times 10^{-4}\ sec$

Equations with Linear Coefficients

Equations with linear coefficients that is, equations of the form

$$\textbf{(13)} \qquad (a_1x + b_1y + c_1)dx + (a_2x + b_2y + c_2)dy = 0 \ ,$$

where the a_i's, b_i's, and c_i's are constants. We leave it as an exercise to show that when $a_1b_2 = a_2b_1$, equation (13) can be put in the form $dy/dx = G(ax + by)$, which we solved via the substitution $z = ax + by$.

Before considering the general case when $a_1b_2 \neq a_2b_1$, let's first look at the special situation when $c_1 = c_2 = 0$. Equation (13) then becomes

$$(a_1x + b_1y)dx + (a_2x + b_2y)dy = 0 \ ,$$

which can be rewritten in the form

$$\frac{dy}{dx} = -\frac{a_1x + b_1y}{a_2x + b_2y} = -\frac{a_1 + b_1(y/x)}{a_2 + b_2(y/x)} \ .$$

This equation is homogeneous, so we can solve it using the method discussed earlier in this section.

The above discussion suggests the following procedure for solving (13). If $a_1b_2 \neq a_2b_1$, then we seek a translation of axes of the form

$$x = u + h \qquad \text{and} \qquad y = v + k \ ,$$

where h and k are constants, that will change $a_1x + b_1y + c_1$ into $a_1u + b_1v$ and change $a_2x + b_2y + c_2$ into $a_2u + b_2v$. Some elementary algebra shows that such a transformation exists if the system of equations

$$\textbf{(14)} \qquad \begin{aligned} a_1h + b_1k + c_1 &= 0 \ , \\ a_2h + b_2k + c_2 &= 0 \end{aligned}$$

has a solution. This is ensured by the assumption $a_1b_2 \neq a_2b_1$, which is geometrically equivalent to assuming that the two lines described by the system (14) intersect. Now if (h, k) satisfies (14), then the substitutions $x = u + h$ and $y = v + k$ transform equation (13) into the homogeneous equation

$$\textbf{(15)} \qquad \frac{dv}{du} = -\frac{a_1u + b_1v}{a_2u + b_2v} = -\frac{a_1 + b_1(v/u)}{a_2 + b_2(v/u)} \ ,$$

which we know how to solve.

Example

Solve $\quad (-3x + y + 6)dx + (x + y + 2)dy = 0$. (16)

Solution Since $a_1b_2 = (-3)(1) \neq (1)(1) = a_2b_1$, we will use the translation of axes $x = u + h$, $y = v + k$, where h and k satisfy the system

$$-3h + k + 6 = 0 \ ,$$
$$h + k + 2 = 0 \ .$$

Solving the above system for h and k gives $h = 1, k = -3$. Hence, we let $x = u + 1$ and $y = v - 3$. Because $dy = dv$ and $dx = du$, substituting in equation (16) for x and y yields

$$(-3u + v)du + (u + v)dv = 0$$

The last equation is homogeneous, so we let $z = v/u$. Then $dv/du = z + u(dz/du)$, and, substituting for v/u, we obtain

$$z + u\frac{dz}{du} = \frac{3 - z}{1 + z} \ .$$

Separating variables gives

$$\int \frac{z + 1}{z^2 + 2z - 3}dz = -\int \frac{1}{u}du \ ,$$

$$\frac{1}{2}\ln|z^2 + 2z - 3| = -\ln|u| + C_1 \ ,$$

from which it follows that

$$z^2 + 2z - 3 = Cu^{-2} \ .$$

When we substitute back in for z, u, and v, we find

$$(v/u)^2 + 2(v/u) - 3 = Cu^{-2} \ ,$$
$$v^2 + 2uv - 3u^2 = C \ ,$$
$$(y + 3)^2 + 2(x - 1)(y + 3) - 3(x - 1)^2 = C \ .$$

This last equation gives an implicit solution to (16).

Riccati Differential Equation

Riccati Equation. An equation of the form

$$\frac{dy}{dx} = P(x)y^2 + Q(x)y + R(x) \tag{18}$$

is called a generalized Riccati equation.

(a) If one solution—say, $u(x)$—of (18) is known, show that the substitution $y = u + 1/v$ reduces (18) to a linear equation in v.

(b) Given that $u(x) = x$ is a solution to

$$\frac{dy}{dx} = x^3(y - x)^2 + \frac{y}{x} ,$$

use the result of part (a) to find all the other solutions to this equation.

Solution:

Ordinary Differential Equations (ODEs)

Chapter 2

Chapter Summary:

In this chapter we have discussed various types of first-order differential equations. The most important were the separable, linear, and exact equations. Their principal features and method of solution are outlined below.

Separable Equations: $dy/dx = g(x)p(y)$. Separate the variables and integrate.

Linear Equations: $dy/dx + P(x)y = Q(x)$. The integrating factor $\mu = \exp\left[\int P(x)dx\right]$ reduces the equation to $d(\mu y)/dx = \mu Q$, so that $\mu y = \int \mu Q\,dx + C$.

Exact Equations: $dF(x, y) = 0$. Solutions are given implicitly by $F(x, y) = C$. If $\partial M/\partial y = \partial N/\partial x$, then $M\,dx + N\,dy = 0$ is exact and F is given by

$$F = \int M\,dx + g(y)\ , \quad \text{where} \quad g'(y) = N - \frac{\partial}{\partial y}\int M\,dx$$

or

$$F = \int N\,dy + h(x)\ , \quad \text{where} \quad h'(x) = M - \frac{\partial}{\partial x}\int N\,dy\ .$$

When an equation is not separable, linear, or exact, it may be possible to find an integrating factor or perform a substitution that will enable us to solve the equation.

Special Integrating Factors: $\mu M\,dx + \mu N\,dy = 0$ is exact. If $(\partial M/\partial y - \partial N/\partial x)/N$ depends only on x, then

$$\mu(x) = \exp\left[\int\left(\frac{\partial M/\partial y - \partial N/\partial x}{N}\right)dx\right]$$

is an integrating factor. If $(\partial N/\partial x - \partial M/\partial y)/M$ depends only on y, then

$$\mu(y) = \exp\left[\int\left(\frac{\partial N/\partial x - \partial M/\partial y}{M}\right)dy\right]$$

is an integrating factor.

Homogeneous Equations: $dy/dx = G(y/x)$. Let $v = y/x$. Then $dy/dx = v + x(dv/dx)$, and the transformed equation in the variables v and x is separable.

Equations of the Form: $dy/dx = G(ax + by)$. Let $z = ax + by$. Then $dz/dx = a + b(dy/dx)$, and the transformed equation in the variables z and x is separable.

Bernoulli Equations: $dy/dx + P(x)y = Q(x)y^n$. For $n \neq 0$ or 1, let $v = y^{1-n}$. Then $dv/dx = (1 - n)y^{-n}(dy/dx)$, and the transformed equation in the variables v and x is linear.

Linear Coefficients: $(a_1x + b_1y + c_1)dx + (a_2x + b_2y + c_2)dy = 0$. For $a_1b_2 \neq a_2b_1$, let $x = u + h$ and $y = v + k$, where h and k satisfy

$$a_1h + b_1k + c_1 = 0\ ,$$
$$a_2h + b_2k + c_2 = 0\ .$$

Then the transformed equation in the variables u and v is homogeneous.

Chapter 3

FIRST-ORDER DIFFERENTIAL EQUATIONS OF HIGHER DEGREE

CHAPTER 3

FIRST-ORDER DIFFERENTIAL EQUATIONS OF HIGHER DEGREE

3.1 Equations of the First-order and not of First Degree

3.2 First-Order Equations of Higher Degree Solvable for Derivative $\frac{dy}{dx} = p$

3.3 Equations Solvable for y

3.4 Equations Solvable for x

3.5 Equations of the First Degree in x and y - Lagrange and Clairant

3.6 Exercises

3.1 Equations of the first-Order and not of First Degree

In this Chapter we discuss briefly basic properties of differential equations of first-order and higher degree. In general such equations may not have solutions. We confine ourselves to those cases in which solutions exist.

The most general form of a differential equation of the first order and of higher degree say of nth degree can be written as

$$\left(\frac{dy}{dx}\right)^n + a_1(x,y)\left(\frac{dy}{dx}\right)^{n-1} + a_2(x,y)\left(\frac{dy}{dx}\right)^{n-2} + \dots \dots$$

$$\dots \dots \dots + a_{n-1}(x,y)\frac{dy}{dx} + a_n(x,y) = 0$$

or
$$p^n+a_1p^{n-1}+a_2p^{n-2}+ \dots\dots +a_{n-1}\,p+a_n=0 \tag{3.1}$$

where $p = \dfrac{dy}{dx}$ and $a_1, a_2, \ldots, a_n$ are functions of x and y.

(3.1) can be written as

$$F(x, y, p) = 0 \tag{3.2}$$

3.2 First-Order Equations of Higher Degree Solvable for p

Let (3.2) can be solved for p and can be written as

$$(p-q_1(x,y))\ (p-q_2(x,y))\ \ldots\ldots\ (p-q_n(x,y)) = 0$$

Equating each factor to zero we get equations of the first order and first degree. One can find solutions of these equations by the methods discussed in the previous chapter. Let their solution be given as:

$$f_i(x,y,c_i)=0,\ i=1,2,3\ \ldots\ldots n \tag{3.3}$$

Therefore the general solution of (3.1) can be expressed in the form

$$f_1(x,y,c)\ f_2(x,y,c)\ldots\ldots f_n(x,y,c) = 0 \tag{3.4}$$

where c in any arbitrary constant.

It can be checked that the sets of solutions represented by (3.3) and (3.4) are identical because the validity of (3.4) in equivalent to the validity of (3.3) for at least one i with a suitable value of c, namely $c=c_i$

Example 3.1 Solve $xy\left(\overset{p}{\dfrac{dy}{dx}}\right)^2 + (x^2+y^2)\overset{p}{\dfrac{dy}{dx}} + xy = 0$ (3.5)

Solution: This is first-order differential equation of degree 2. Let $p = \dfrac{dy}{dx}$

$xyp^2 + (x^2+y^2)p + xy = 0$

$(xp + y)(yp + x) = 0$ ✓

Equation (3.5) can be written as

$$xy\,p^2+(x^2+y^2)\,p+xy=0 \tag{3.6}$$

$$(xp+y)(yp+x)=0$$

This implies that

$xp+y=0 \Rightarrow$ $\quad x\frac{dy}{dx}+y=0$ $\quad x\,dy+y\,dx=0$ $\quad x\,dy=-y\,dx$

$$xp+y=0,\; yp+x=0 \tag{3.7}$$

By solving equations in (3.7) we get

$\frac{dy}{y}=-\frac{dx}{x}$ $\quad xy=c_1$ and $x^2+y^2=c_2$, respectively

$$x\frac{dy}{dx}+y=0 \;\text{ or }\; \frac{dy}{dx}+\frac{1}{x}y=0,$$ Integrating factor

$$I(x)=e^{\int\frac{1}{x}dx}=e^{\log x}.$$

This gives

$$y.x=\int o.x\,dx+c_1 \text{ or } xy=c_1$$

$$y\frac{dy}{dx}+x=0, \quad \text{or} \quad ydy+xdx=0$$

By integration we get $\frac{1}{2}y^2+\frac{1}{2}x^2=c$

or $\quad x^2+y^2=c_2,\; c_2>0,\; -\sqrt{c_2}\le x\le\sqrt{c_2}$

The general solution can be written in the form

$$(x^2+y^2-c_2)\,(xy-c_1)=0 \tag{3.8}$$

It can be seen that none of the nontrivial solutions belonging to $xy=c_1$ or $x^2+y^2=c_2$ is valid on the whole real line.

3.3 Equations Solvable for y

Let the differential equation given by (3.2) be solvable for y. Then y can be expressed as a function x and p, that is,

$$y = f(x,p) \tag{3.9}$$

Differentiating (3.9) with respect to x we get

$$\frac{dy}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial p}\cdot\frac{dp}{dx} \tag{3.10}$$

(3.10) is a first order differential equation of first degree in x and p. It may be solved by the methods of Chapter 2. Let solution be expressed in the form

$$\varphi(x,p,c) = 0 \tag{3.11}$$

The solution of equation (3.9) is obtained by eliminating p between (3.9) and (3.11). If elimination of p is not possible then (3.9) and (3.11) together may be considered parametric equations of the solutions of (3.9) with p as a parameter.

Example 3.2: Solve $y^2-1-p^2=o$

Solution: It is clear that the equation is solvable for y, that is

$$y = \sqrt{1+p^2} \tag{3.12}$$

By differentiating (3.12) with respect to x we get

$$\frac{dy}{dx} = \frac{1}{2}\frac{1}{\sqrt{1+p^2}}.2p\frac{dp}{dx} \qquad \text{or} \qquad p = \frac{p}{\sqrt{1+p^2}}\frac{dp}{dx}$$

or $$p\left[1-\frac{1}{\sqrt{1+p^2}}\frac{dp}{dx}\right] = 0 \qquad (3.13)$$

(3.13) gives p=o or $1-\frac{p}{\sqrt{1+p^2}}\frac{dp}{dx} = 0$

By solving p=0 in (3.12) we get y=1

By $$1-\frac{1}{\sqrt{1+p^2}}\frac{dp}{dx} = 0$$

we get a separable equation in variables p and x.

$$\frac{dp}{dx} = \sqrt{1+p^2}$$

By solving this we get

$$p=\sinh(x+c) \qquad (3.14)$$

By eliminating p from (3.12) and (3.14) we obtain

$$y=\cos h\,(x+c) \qquad (3.15)$$

(3.15) is a general solution.

Solution y=1 of the given equation is a singular solution as it cannot be obtained by giving a particular value to c in (3.15).

3.4 Equations Solvable for x

Let equation (3.2) be solvable for x, that is

$$x=f(y,p) \tag{3.16}$$

Then as argued in the previous section for y we get a function Ψ such that

$$\Psi(y, p, c) = 0 \tag{3.17}$$

By eliminating p from (3.16) and (3.17) we get a general solution of (3.2). If elimination of p with the help of (3.16) and (3.17) is cumbersome then these equations may be considered parametric equations of the solutions of (3.16) with p as a parameter.

Example 3.3 Solve $x\left(\frac{dy}{dx}\right)^3 - 12\frac{dy}{dx} - 8 = 0$

Solution: Let $p = \frac{dy}{dx}$, *then* $xp^3-12p-8=0$

It is solvable for x, that is,

$$x = \frac{12p+8}{p3} = \frac{12}{p^2} + \frac{8}{p^3} \tag{3.18}$$

Differentiating (3.18) with respect to y, we get

$$\frac{dx}{dy} = -2\frac{12}{p^3}\frac{dp}{dy} - 3\frac{8}{p^4}\frac{dp}{dy} \quad \text{or} \quad \frac{1}{p} = -\frac{24}{p^3}\frac{dp}{dy} - \frac{24}{p^4}\frac{dp}{dy} \quad or \quad dy = \left(-\frac{24}{p^2} - \frac{24}{p^3}\right)dp$$

$$or \quad y = +\frac{24}{p} + \frac{12}{p^2} + c \tag{3.19}$$

(3.18) and (3.19) constitute parametric equations of solution of the given differential equation.

3.5 Equations of the First Degree in x and y – Lagrange's and Clairaut's Equation.

Let Equation (3.2) be of the first degree in x and y, then

$$y = x\varphi_1(p) + \varphi_2(p) \tag{3.20}$$

Equation (3.20) is known as Lagrange's equation. If $\varphi_1(p) = p$ then the equation

$$y = xp + \varphi_2(p) \tag{3.21}$$

is known as Clairaut's equation. By differentiating (3.20) with respect to x, we get

$$\frac{dy}{dx} = \phi_1(p) + x\phi_1'(p)\frac{dp}{dx} + \phi_2'(p)\frac{dp}{dx}$$

$$\text{or} \quad p - \phi_1(p) = (x\phi_1'(p) + \phi_2'(p))\frac{dp}{dx} \tag{3.22}$$

From (3.22) we get

$$(x + \varphi_2'(p))\frac{dp}{dx} = 0 \quad \text{for} \quad \varphi_1(p) = p$$

This gives

$$\frac{dp}{dx} = 0 \text{ or } x + \varphi_2'(p) = 0$$

$$\frac{dp}{dx} = 0 \text{ gives } p = c \text{ and}$$

by putting this value in (3.21) we get

$$y = cx + \varphi_2(c)$$

This is a general solution of Clairaut's equation. The elimination of p between $x+\varphi_2'(p) = 0$ and (3.21) gives a singular solution. If $\varphi_1(p) \neq p$ for any p, then we observe from (3.22) that $\frac{dp}{dx} \neq 0$ everywhere. Division by

$[p-\varphi_1(p)]\frac{dp}{dx}$ in (3.22) gives $\frac{dx}{dp} - \frac{\varphi_1'}{p-\varphi_1(p)}x = \frac{\varphi_2'(p)}{p-\varphi_1(p)}$

which is a linear equation of first order in x and thus can be solved for x as a function of p, which together with (3.20) will form a parametric representation of the general solution of (3.20).

Example 3.4 Solve $\left(\frac{dy}{dx}-1\right)\left(y-x\frac{dy}{dx}\right)=\frac{dy}{dx}$

Solution: Let $p=\frac{dy}{dx}$ then, $(p-1)(y-xp)=p$

This equation can be written as

$$y = xp + \frac{p}{p-1}$$

Differentiating both sides with respect to x we get

$$\frac{dp}{dx}\left[x-\frac{1}{(p-1)^2}\right]=0$$

Thus either $\frac{dp}{dx}=0$ or $x-\frac{1}{(p-1)^2}=0$

$\frac{dp}{dx} = 0$ gives p=c .

Putting p=c in the equation we get

$$y = cx + \frac{c}{c-1} \Rightarrow (y-cx)(c-1)=c$$

which is the required solution.

3.6 Exercises

Solve the following differential equations

1. $\left(\frac{dy}{dx}\right)^3 = \frac{dy}{dx}e^{2x}$

2. $y(y-2)p^2 - (y-2x+xy)p+x=0$

3. $-\left(\frac{dy}{dx}\right)^2 + 4y - x^2 = 0$

4. $\left(\frac{dy}{dx} + y + x\right)\left(x\frac{dy}{dx} + y + x\right)\left(\frac{dy}{dx} + 2x\right) = 0$

5. $y + x\frac{dy}{dx} - x^4\left(\frac{dy}{dx}\right)^2 = 0$

6. $\left(x\frac{dy}{dx} - y\right)\left(y\frac{dy}{dx} + x\right) = h^2\frac{dy}{dx}$

7. $y\left(\frac{dy}{dx}\right)^2 + (x-y)\frac{dy}{dx} = x$

8. $x\left(\frac{dy}{dx}\right)^2 - 2y\frac{dy}{dx} + ax = 0$

Probability

Probability

CHAPTER ONE

Probability

1.1 Random experiments

A random experiment is an experiment whose outcomes cannot be predicted with certainty. However, in most cases the collection of every possible outcome of a random experiment can be listed.

1.2 Sample space

In statistics, the set of all possible outcomes of an experiment is called the sample space of the experiment, because it usually consists of the things that can happen when one takes a sample. Sample spaces are usually denoted by the letter S.

Each outcome in a sample space is called an element or a member of the sample space, or simply a sample point.

Example 1.1

Experiment: Tossing a fair coin

Sample space, S={H, T}

Where: H= Head, T= Tail

Example 1.2

Consider the experiment of rolling a six-sided dice and observing the number which appears on the uppermost face of the dice. The result can be any of the numbers 1, 2, 3, … , 6. This is a random experiment since the outcome is uncertain.

Example 1.3

A coin is tossed twice.

Sample space, S={HH, HT, TH, TT}

Example 1.4

Consider the experiment of rolling a red dice and a green dice and observing the number which appears on the uppermost face of each dice. The sample space of the experiment consists of the following array of 36 outcomes.

The first coordinate of each point is the number which appears on the red dice, while the second coordinate is the number which appears on the green dice.

	Green die					
Red die	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
	.	.	.	.	.	.
	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Using the standard notation for sets, we can express this sample space as follows:
S = {(i, j): i = 1, 2, 3, 4, 5, 6; j = 1, 2, 3, 4, 5, 6}.

1.3 The **Event**

A subset of a sample space is called an ***event***. The empty set, ϕ, is a subset of S and S is also a subset of S. ϕ and S are therefore events. We call ϕ the ***impossible event*** and S the ***certain event***. A subset of S containing one element of S is called a **simple event**.

Example 1.5

Suppose that we toss a fair coin three times and record the outcomes.

Find:

1- Sample space.

2- The events A and B.

Where A: we observe two heads

B: we observe at least one tail.

C: we observe at least one head.

The sample space is:

S={HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}
A={HHT, HTH, THH}
B={ HHT, HTH, THH, HTT, THT, TTH, TTT}
C={HHH, HHT, HTH, THH, HTT, THT, TTH }

Example 1.6

Suppose that we toss two dice.

Find:

1- Sample space.
2- Describe the event A that the total numbers of points rolled with the pair of dice is 7.

The sample space is:

S={(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6)
(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)
(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)
(5,1), (5,2), (5,3), (5,4), (5,5), (5,6)
(6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}

A={(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}

1.4 Operations on events

Since an event is a subset of a sample space, we can combine events to form new events, using the various set operations. The sample space is considered as the universal set. If A and B are two events defined on the same sample space, then:

(1) **Union:** $A \cup B = \{x : x \in A \text{ or } x \in B\}$
$A \cup B$ denotes the event "A or B or both". Thus the event $A \cup B$ occurs if either A occurs or B occurs or both A and B occur.

(2) **Intersection:** $A \cap B = \{x : x \in A \text{ and } x \in B\}$
$A \cap B$ denotes the event "both A and B". Thus the event $A \cap B$ occurs if both A and B occur.

(3) **Complement:** $A^c = \{x : x \in S \text{ but } x \notin A\}$
A^c or A' denotes the event which occurs if and only if A does not occur.

(4) **Difference:** $A - B = \{x : x \in A \text{ and } x \notin B\}$

(5) **Symmetric Difference**: $A \Delta B = (A - B) \cup (B - A)$

(6) Distributive law:

i- : $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

ii- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(7) **Demorgan's law:**

i- $(A \cup B)^c = A^c \cap B^c$

ii- $(A \cap B)^c = A^c \cup B^c$

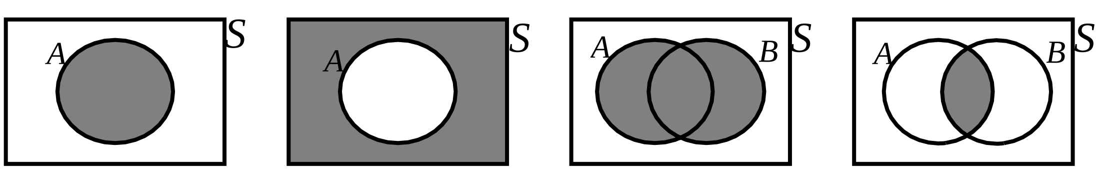
Example 1.7

Let P be the event that an employee selected at random from an oil drilling company smokes cigarettes. Let Q be the event that an employee selected drinks alcoholic beverages. Then $P \cup Q$ is the event that an employee selected either drinks or smokes, or drinks and smokes. $P \cap Q$ is the event that an employee selected drinks and smokes.

Example 1.8

If $A = \{x : 3 < x < 9\}$ and $B = \{x : 5 \leq x < 12\}$, then $A \cup B = \{x : 3 < x < 12\}$, and
$A \cap B = \{x : 5 \leq x < 9\}$.

Sample spaces and events, particularly relationships among events, are often depicted by means of Venn diagrams like those of Fig. 1.2. In each case, the sample space is represented by a rectangle, whereas events are represented by regions within the rectangle, usually by circles or parts of circles. The shaded regions of the four diagrams of Fig. 1.2 represent event A, the complement of A, the union of events A and B, and the intersection of events A and B.



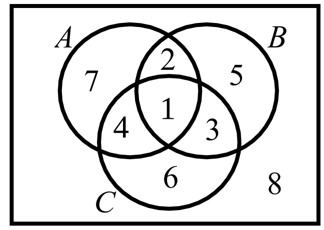
A $\quad$ $A \cup B$ $\quad$ $A \cap B$

A

′

***Fig. 1.2**: Venn diagrams showing the complement, union and intersection*

When we deal with three events, we draw the circles as in Fig. 1.3. In this diagram, the circles divide the sample space into eight regions, numbered 1 through 8, and it is easy to determine whether the



corresponding events are parts of A or and C or C'.

Mutually exclusive (or disjoint) events

Any two events that cannot occur simultaneously, so that their intersection is the impossible event, are said to be ***mutually exclusive*** (or disjoint). Thus two events A and B are mutually exclusive if and only if $A \cap B = \phi$. In general, a collection of events $A_1, A_2, \ldots, A_n$, is said to be mutually exclusive if there is no overlap among any of them. That is, if $A_i \cap A_j = \phi$ $(i \neq j,\ i, j = 1, 2, \ldots n)$. Fig. 1.4 shows several mutually exclusive events

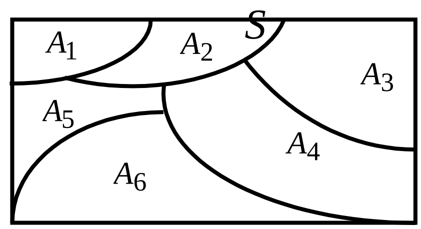


Fig. 1.4*: Mutually exclusive events*

De Morgan's Laws

Venn diagrams are often used to verify relationships among sets, thus making it unnecessary to give formal proofs based on the algebra of sets. To illustrate, let us show that $(A \cup B)^c = A^c \cap B^c$, which expresses the fact that the complement of the union of two sets equals the intersection of their complements. To begin, note that in Figures 1.5 and 1.6, A and B are events defined on the same sample space S. In Fig. 1.5, the shaded area represents the event $(A \cup B)^c$.

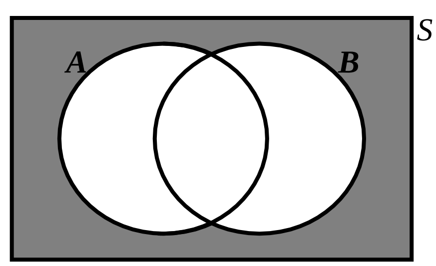


Fig. 1.5*: $(A \cup B)^c$ is shaded*

In Fig. 1.6, the area shaded vertically represents the event A^c while the area shaded horizontally represents the event B^c. It follows that the cross-shaded area represents the event $A^c \cap B^c$. But the total shaded area in Fig. 1.5 is identical zto the cross-shaded area in Fig. 1.6. We can therefore state the following theorem.

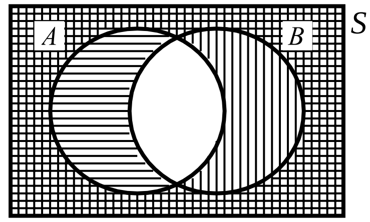


Fig. 1.6. A^c and B^c *are shaded*

1

$(A \cup B)^c = A^c \cap B^c$.

Similarly, we can use Venn diagrams to verify the following two theorems.

2

$(A \cap B)^c = A^c \cup B^c$.

3

If A, B and C are events defined on the same sample space, then

$(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$, $(A \cap B \cap C)^c = A^c \cup B^c \cup C^c$, and $(A^c \cup B^c \cup C^c)^c = A \cap B \cap C'$.

The results given in 1, 2 and 3 are called ***de Morgan's laws***.

Other useful facts concerning operations on events

The following results can be verified by means of Venn diagrams.

1. **Commutative law**

   $A \cup B = B \cup A$; $A \cap B = B \cap A$.

2. **Associative law**

   $A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$.

   Because of these two statements, we can use the simpler notations

   $A \cup B \cup C$ and $A \cap B \cap C$

   without fear of ambiguity when referring to $A \cup (B \cup C)$ and $A \cap (B \cap C)$, respectively.

3. **Distributive law**

   $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

4. Other useful results

(a) $S^c = \phi$, $\phi^c = S$, $(A^c)^c = A$

(b) $A \cup S = S$, $A \cap \phi = \phi$

(c) $A \cup A^c = S$, $A \cap A^c = \phi$

1.5 Counting sample points

In this section, we discuss some techniques for determining, without direct enumeration, the number of possible outcomes of an experiment. Such techniques are useful in finding probabilities of some complex events.

The multiplication principle

We begin with the following basic principle:

Theorem 1.1 (The multiplication theorem)

If an operation can be performed in n_1 ways and after it is performed in any one of these ways, a second operation can be performed in n_2 ways and, after it is performed in any one of these ways, a third operation can be performed in n_3 ways, and so on for k operations, then the k operations can be performed together in $n_1 n_2 \ldots n_k$ ways.

Example 1. 10

How many lunches consisting of a soup, sandwich, dessert, and a drink are possible if we can select 4 soups, 3 kinds of sandwiches, 5 desserts and 4 drinks?

Solution

Here, $n_1 = 4$, $n_2 = 3$, $n_3 = 5$ and $n_4 = 4$. Hence there are

$$n_1 \times n_2 \times n_3 \times n_4 = 4 \times 3 \times 5 \times 4 = 240$$

different ways to choose a lunch.

Theorem 1.2

> The number of combinations of n distinct objects taken r at a time is
> $$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}.$$

Example 1.11

From 4 chemists and 3 physicists, find the number of committees that can be formed consisting of 2 chemists and 1 physicist.

Solution

The number of ways of selecting 2 chemists from 4 is:

$$\binom{4}{2} = \frac{4!}{2!2!} = 6$$

The number of ways of selecting 1 physicist from 3 is:

$$\binom{3}{1} = \frac{3!}{1!2!} = 3$$

Using the multiplication theorem (see Theorem 1.4 on page 8) with $n_1 = 6$ and $n_2 = 3$, it can be seen that we can form $n_1 \times n_2 = 6\times3 = 18$ committees with 2 chemists and 1 physicist.

1.3 The probability of an event

1.3.1 Introduction

It is frequently useful to quantify the likelihood, or chance, that an outcome of a random experiment will occur. For example, we may hear a physician say that a patient has 50-50 chance of surviving a certain operation. Another physician may say that she is 95% certain that a patient has a particular disease. A public health nurse may say that 80% of certain clients will break an appointment. As these examples suggest, most people express probabilities in terms of percentages. However, in dealing with probabilities mathematically, it is convenient to express probabilities as fractions. Thus, we measure the probability of the occurrence of some event by a number between zero and one. The more likely the event, the closer the number is to one; and the more unlikely the event, the closer the number is to zero.

1.3.2 Classical probability

The classical treatment of probability dates back to the 17th century and the work of two mathematicians, Pascal and Fermat [see Todhunter (1931) and David (1962)]. Much of this theory developed out of attempts to solve problems related to games of chance, such as those involving the rolling of dice [see Jeffreys (1939), Ore (1960), and Keynes (1921)]. We can calculate the probability of an event in the classical sense as follows.

Definition 1.1

If a trial of an experiment can result in m mutually exclusive and equally likely outcomes, and if exactly h of these outcomes correspond to an event A, then the probability of event A is given by

$$P(A) = \frac{h}{m} = \frac{\text{number of ways that } A \text{ can occur}}{\text{number of ways the sample space } S \text{ can occur}}.$$

Thus, if all the simple events in S are equally likely, then

$$P(A) = \frac{n(A)}{n(S)} \; for\ all\ \ A \subset S$$

where $n(A)$ denotes the number of elements in A. We emphasize that the above expression for $P(A)$ is applicable only when all the simple events in S are equally likely.

It is important to realize that here, we are using the same symbol A to represent two different things. In the expression n(A), A represents a set (for example, the set of even integers less than 7) whereas when we write P(A), A represents an event (for example, the score on a die is even).

Example 1.12

A mixture of candies contains 6 mints, 4 toffees, and 3 chocolates. If a person makes a random selection of one of these candies, find the probability of getting
(a) a mint, (b) a toffee or a chocolate.

Solution

Let M, T, and C represent the events that the person selects, respectively, a mint, toffee, or chocolate candy. The total number of candies is 13, all of which are equally likely to be selected.

(a) Since 6 of the 13 candies are mints,

$$P(M) = \frac{6}{13}.$$

(b) $P(T \cup C) = \dfrac{n(T \cup C)}{n(S)} = \dfrac{7}{13}.$

Example 1.13

The following table shows 100 patients classified according to blood group and gender.

	Blood group		
	A	*B*	*O*
Male	30	20	17
Female	15	10	8

If a patient is selected at random from this group, find the probability that the patient selected:
(a) is a male and has blood group B, (b) is a female and has blood group A.

Solution

There are 100 ways in which we can select a patient from the 100 patients. Since the patient is selected at random, all the 100 ways of selecting a patient are equally likely.

(a) There are 20 males with blood group *B*. Therefore the probability that the patient selected is a male and has a blood group B is $\frac{20}{100}$=0.2

(b) There are 15 females with blood group *A*. Therefore the probability that the patient selected is a female and has blood group *A* is $\frac{15}{100}$ =0.15

One advantage of the classical definition of probability is that it does not require experimentation. Furthermore, if the outcomes are truely equally likely, then the probability assigned to an event is not an approximation. It is an accurate description of the frequency with which the event will occur.

1.3.1 Relative frequency probability

The relative frequency approach was developed by Fisher (1921) and Von Mises (1941), and depends on the repeatability of some process and the ability to count the number of repetitions, as well as the number of times that some event of interest occurs. In this context, we may define the probability of observing some characteristic, *A*, of an event as follows:

Definition 1.2

If some process is repeated a large number of times *n*, and if some resulting event with the characteristic *A* occurs *m* times, the relative frequency of occurrence of *A*, *m n* , will be approximately equal to the probability of *A*. Thus,

$$P(A) = \lim_{n \to \infty} \frac{n(A)}{n}.$$

The disadvantage in this approach is that the experiment must be repeatable. Remember that any probability obtained this way is an approximation. It is a value based on *n* trials. Further testing might result in a different approximate value.

Example 1.14

The following table gives the frequency distribution of the heights of 150 students. If a student is selected at random from this group, find the probability that the student selected is taller than the modal height of the students.

Height(cm)	130	140	150	160	170	180	190
Frequency	8	16	28	44	33	17	4

Solution

The modal height of the students is 160 cm. This is the height with the highest frequency. The number of students who are taller than 160 cm is (33 + 17 + 4) = 54. An estimate of the required probability is the relative frequency

$$\frac{54}{150} = 0.36.$$

1.3.2 Subjective probability

In the early 1950s, Savage (1972) gave considerable impetus to what is called **subjective** concept of probability. This view holds that probability measures the confidence that a particular individual has in the truth of a particular proposition. This concept does not depend on the repeatability of any process. By applying this concept of probability, one can calculate the probability of an event that can only

happen once, for example, the probability that a cure for HIV/AIDS will be discovered within the next 8 years.

Although the subjective view of probability has enjoyed increased attention over the years, it has not been fully accepted by statisticians who have traditional orientations.

1.4 Some probability laws

In the last section, we considered how to interpret probabilities. In this section, we consider some laws that govern their behaviour.

1.4.1 Axioms of probability

In 1933, the axiomatic approach to probability was formalized by the Russian mathematician A. N. Kolmogorov (1964). The basis of this approach is embodied in three axioms from which a whole system of probability theory is constructed through the use of mathematical logic. The three axioms are as follows.

Axioms of probability

Let S be the sample space of an experiment and P be a set function which assigns a number $P(A)$ to every $A \subset S$. Then the function $P(A)$ is said to be a probability function if it satisfies the following three axioms:

Axiom 1: $P(S) = 1$.

Axiom 2: $0 \leq P(A) \leq 1$ for every event A.

Axiom 3: If A and B are mutually exclusive events, then: $\boldsymbol{P(A \cup B) = P(A) + P(B)}$.

Theorem 1.3

> If Φ is a empty set (that is an impossible event), then $P(\phi) = 0$.

Proof:

For any event A, $A \cap \phi = \phi$, Then A and ϕ are two exclusive events, and $A \cup \phi = A$,

$P(A \cup \phi) = P(A)$, by axiom 3, we get

$P(A \cup \phi) = P(A) = P(A) + P(\phi)$, then

$P(\phi) = 0$.

And the proof of the theorem is complete.

This theorem says that the probability of an impossible event is zero.

Theorem 1.4

> If A be any event of the sample space S, then $P(A^c) = 1 - P(A)$.

Where A^c denotes the complement of A with respect to S.

Proof:

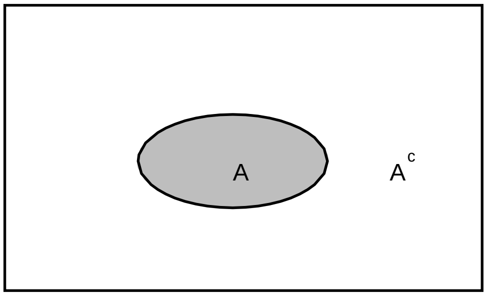
Let A be any subset of S. Then $S = A \cup A^c$. Further A and A^c are mutually exclusive. Hence, by Axiom 3, we get

$$1 = P(S) = P(A \cup A^c)$$
$$= P(A) + P(A^c).$$

Hence, we see that

This completes the proof

$$P(A^c) = 1 - P(A).$$



Theorem 1.5

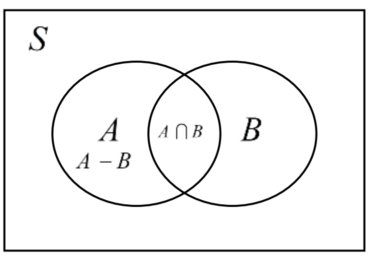
> If A and B are events of the sample space S, then: $P(A-B)=P(A)-P(A\cap B)$.

Proof:

We have $A=(A-B)\cup(A\cap B)$

Where $(A-B),(A\cap B)$ are exclusive events, $(A-B)\cap(A\cap B)=\phi$.

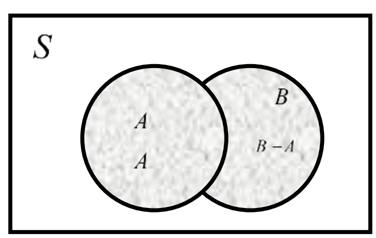
So
$$P(A)=P\big((A-B)\cup(A\cap B)\big)=P(A-B)+P(A\cap B),$$
$$\therefore P(A-B)=P(A)-P(A\cap B).$$



Theorem 1.6

> If A and B are events of the sample space S, then: P (A ∪ B) = P (A) + P (B) − P (A ∩ B).

Proof:



It is easy to see that

$A\cup B=A\cup(B-A)$

We have $A,(B-A)$ are exclusive events

$A\cap(B-A)=\phi$,

Applying Axiom 3, we obtain

$$p(A \cup B) = p[A \cup (B - A)] = p(A) + p(B - A)$$

From Theorem 1.5

$$p(B - A) = p(B) - p(A \cap B)$$

We get:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

Corollary 1.1

If the events A and B are mutually exclusive, then $A \cap B = \phi$ and so by Theorem 1.8, $P(A \cap B) = 0$. Theorem 1.10 then becomes

$$P(A \cup B) = P(A) + P(B).$$

Corollary 1.1 can be extended by mathematical induction to the following corollary.

Corollary 1.2

If the events $A_1, A_2, \ldots, A_n$ are mutually exclusive, then;

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

The following corollary gives an extension of Theorem 1.10 to 3 events.

Corollary 1.3

If A, B, and C are three events defined on the same sample space, then;

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Example 1.15

The probability that Sara passes Mathematics is $\frac{2}{3}$, and the probability that she passes English is $\frac{4}{9}$.

If the probability that she passes both courses is $\frac{1}{4}$, what is the probability that she passes at least one of the two courses?

Solution

Let M denote the event "Sara passes Mathematics" and E the event "Sara passes English". We wish to find $P(M \cup E)$. By the addition rule of probability, (see Theorem 1.10 on page 17),

$$P(M \cup E) = P(M) + P(E) - P(M \cap E)$$

$$= \frac{2}{3} + \frac{4}{9} - \frac{1}{4} = \frac{31}{36}.$$

Example 1.16

Refer to Example 1.20 on page 14. If a patient is selected at random from the 100 patients, find the probability that the patient selected:

(a) Is a male or has blood group A,

(b) Does not have blood group A,

(c) is a female or does not have blood group B.

Solution

There are 100 ways in which we can select a patient from the 100 patients. Since the patient is selected at random, all the 100 ways of selecting a patient are equally likely.

Let M denote the event "a patient is a male" and A the event "a patient has blood group A". We

wish to find $P(M \cup A)$. By the addition rule of probability,

(a) Let M denote the event "a patient is a male" and A the event "a patient has blood group A". We wish to find $P(M \cup A)$. By the addition rule of probability,

$$\begin{aligned} P(M \cup A) &= P(M) + P(A) - P(M \cap A) \\ &= \frac{n(M)}{100} + \frac{n(A)}{100} - \frac{n(M \cap A)}{100} \\ &= \frac{67}{100} + \frac{45}{100} - \frac{30}{100} = \frac{82}{100} = 0.82. \end{aligned}$$

(b) We wish to find $P(A')$. By Theorem 1.9,

$$P(A') = 1 - P(A) = 1 - \frac{n(A)}{100} = 1 - \frac{45}{100} = 0.55.$$

(c) Let F denote the event "a patient selected is a female" and B the event "a patient selected has blood group B". We wish to find $P(F \cup B')$. By the addition rule of probability,

$$\begin{aligned} P(F \cup B') &= P(F) + P(B') - P(F \cap B') \\ &= P(F) + \{1 - P(B)\} - P(F \cap B') \\ &= \frac{33}{100} + \left(1 - \frac{30}{100}\right) - \frac{23}{100} = 1 - \frac{20}{100} = 0.8. \end{aligned}$$

Example 1.17

Of 200 students in a certain Senior High School, 60 study Mathematics, 40 study Biology, 30 study Chemistry, 10 study Mathematics and Biology, 5 study Mathematics and Chemistry, 3 study Biology and Chemistry and 1 studies the three subjects. If a student is selected at random from this group, find the probability that the student selected studies at least one of the three subjects.

Solution

Let S = {the 200 students}, M = {those who study Mathematics}, B = {those who study Biology} and C = {those who study Chemistry}. Then, $n(M) = 60$, $n(B) = 40$, $n(C) = 30$, $n(M \cap B) = 10$, $n(M \cap C) = 5$, $n(B \cap C) = 3$ and $n(M \cap B \cap C) = 1$. We are required to calculate $P(M \cup B \cup C)$. By

Corollary 1.3,

$$\begin{aligned} P(M \cup B \cup C) &= P(M) + P(B) + P(C) - P(M \cap B) \\ &\quad - P(M \cap C) - P(B \cap C) + P(M \cap B \cap C) \\ &= \frac{60}{200} + \frac{40}{200} + \frac{30}{200} - \frac{10}{200} - \frac{5}{200} - \frac{3}{200} + \frac{1}{200} = \frac{113}{200}. \end{aligned}$$

Example 1.18

If the probabilities are, respectively, 0.08, 0.14, 0.22, and 0.24 that a person buying a new car will choose the colour green, white, red, or black, calculate the probability that a given buyer will purchase a new car that comes in one of these colours.

Solution

Let G, W, R, and B be the events that a buyer selects, respectively, a green, white, red or black car. Since the four colours are mutually exclusive, the required probability is

$$\begin{aligned} P(G \cup W \cup R \cup B) &= P(G) + P(W) + P(R) + P(B) \\ &= 0.08 + 0.14 + 0.22 + 0.24 = 0.68. \end{aligned}$$

1.3.1 Two-set problems

If A and B are any two events defined on a sample space S, then we can draw Fig. 1.8. It can be seen that S can be split into the following four mutually exclusive events:
$A \cap B, A^c \cap B, A \cap B^c$ and $A^c \cap B^c$.

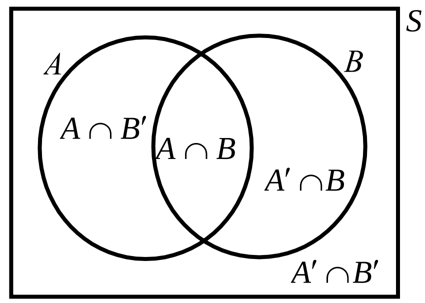


Fig. 1.8: *Two-set problems*

Notice that:

$$A = (A \cap B^c) \cup (A \cap B)$$

Since $A \cap B^c$ and $A \cap B$ are mutually exclusive,

$$P(A) = P(A \cap B^c) + P(A \cap B) \quad (1.4)$$

Similarly,

$$P(B) = P(A^c \cap B) + P(A \cap B) \quad (1.5)$$

Moreover,

$$(A \cap B^c) \cup (A \cap B) \cup (A' \cap B) \cup (A^c \cap B^c) = S$$

and since the four events are mutually exclusive,

$$P(A \cap B^c) + P(A \cap B) + P(A^c \cap B) + P(A^c \cap B^c) = P(S) = 1.$$

Example 1.19

The probability that a new airport will get prize for its design is 0.04, the probability that it will get prize for the efficient use of materials is 0.2 and the probability that it will get both prizes is 0.03. Find the probability that it will get:

(a) At least one of the two prizes,

(b) Only one of the two prizes,

(c) None of the two prizes.

Solution

Let D denote the event "the airport will get an award for its design", and E the event "the airport will get an award for the efficient use of materials".

We are given that $P(D) = 0.04$, $P(E) = 0.2$ and $P(D \cap E) = 0.03$. We can therefore draw Fig. 1.9. Notice that, since $P(D) = 0.04$, and $P(D \cap E) = 0.03$, $P(D \cap E^c) = 0.04 - 0.03 = 0.01$.

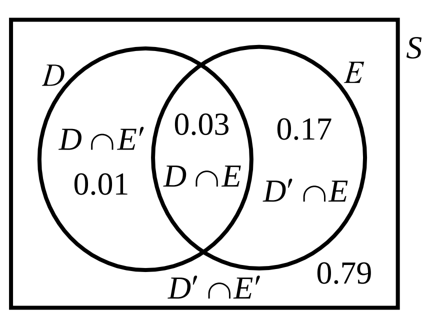


Fig. 1.9: A *Venn diagram*

(a) We wish to find $P(D \cup E)$. From Fig. 1.9,

$$P(D \cup E) = 0.17 + 0.03 + 0.01 = 0.21$$

Alternatively,

$$P(D \cup E) = P(D) + P(E) - P(D \cap E)$$
$$= 0.04 + 0.2 - 0.03 = 0.21.$$

(b) The probability that it will get only one of the awards is

$$P(D \cap E^c) + P(D^c \cap E) = 0.01 + 0.17 = 0.18.$$

(c) We wish to find $P(D^c \cap E^c)$. From Fig. 1.9,

$$P(D^c \cap E^c) = 0.79.$$

Alternatively, using Theorem 1.6, we obtain

$$P(D^c \cap E^c) = 1 - P\left[(D^c \cap E^c)^c\right]$$
$$= 1 - P(D \cup E) = 1 - 0.21, \quad \text{(from part (a))}$$
$$= 0.79.$$

Example 1.20

The events A, B and C have probabilities $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ and $P(C) = \frac{1}{4}$. Furthermore, $A \cap C = \phi$, $B \cap C = \phi$ and $P(A \cap B) = \frac{1}{6}$. Find:

(a) $P[(A \cap B)^c]$ (b) $P(A \cap B^c)$, (c) $P[(A \cup B)^c]$,

(d) $P(A^c \cap B^c)$, (e) (e) $P(A \cup B \cup C)$.

Solution

(a) Applying Theorem 1.4, we obtain

$$P[(A \cap B)^c] = 1 - P(A \cap B) = 1 - \frac{1}{6} = \frac{5}{6}.$$

(b) $P(A \cap B^c) = P(A) - P(A \cap B)$, (see Equation (1.4) on page 20)

$$= \frac{1}{2} - \frac{1}{6} = \frac{1}{3}.$$

(c) Applying Theorem 1.6, we obtain

$$P[(A \cup B)^c] = 1 - P(A \cup B) = 1 - \{P(A) + P(B) - P(A \cap B)\}$$
$$= 1 - \left(\frac{1}{2} + \frac{1}{3} - \frac{1}{6}\right) = 1 - \frac{2}{3} = \frac{1}{3}.$$

(d) $P(A^c \cap B^c) = P[(A \cup B)^c]$, (see Theorem 1.1 on page 6)

$$= \frac{1}{3}, \quad \text{(from part (c))}.$$

(e) Applying Corollary 1.3, we obtain

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$
$$= \frac{1}{2} + \frac{1}{3} + \frac{1}{4} - \frac{1}{6} - 0 - 0 + 0 = \frac{11}{12}.$$

1.3 Conditional probability

A box contains n white balls and m red balls. All the balls are of the same size. Without looking, a person takes a ball from the box. Without replacing the ball in the box, he then takes a second ball. Let B denote the event "the first ball drawn is white" and A the event "the second ball drawn is red". For the first draw, any of the $(m + n)$ balls in the box is equally likely to be drawn, while n of these are white. Therefore,

$$P(B) = \frac{n}{m+n}.$$

If the first ball drawn was white, then the probability that the second ball drawn is red is

$\frac{m}{m+n-1}$, since only m of the remaining $(m + n - 1)$ balls are red.

We introduce a new notation to describe this probability. We call it the ***conditional probability*** of event A given B. We denote it by $P(A|B)$. This is usually read as "the probability of A, given B". in the above example, Thus,

$$P(A \mid B) = \frac{m}{m+n-1}.$$

We now give a formal definition of conditional probability.

Definition 1.3

If A and B are any two events defined on the same sample space S, the conditional probability of A given B, is defined by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0 \text{ .... } (1.6)$$

and if $P(B) = 0$, then $P(A \mid B)$ is undefined.

In particular, if S is a finite, equiprobable sample space (see page 3), then:

$$P(A \cap B) = \frac{n(A \cap B)}{n(S)}, \quad P(B) = \frac{n(B)}{n(S)}, \text{ and so}$$

$$P(A \mid B) = \left\{\frac{n(A \cap B)}{n(S)}\right\} \Big/ \left\{\frac{n(B)}{n(S)}\right\} = \frac{n(A \cap B)}{n(B)} \text{ ....}$$

Example 1.21

Consider the data given in Example 1.13. If a patient is chosen at random from the 100 patients, find the probability that the patient chosen has blood group *A* given that he is a male.

Solution

Let *M* denote the event "the patient chosen is a male" and *A* the event "the patient chosen has blood group *A*". We wish to find $P(A \mid M)$.

$$P(A \mid M) = \frac{P(A \cap M)}{P(M)} = \frac{30/100}{67/100} = \frac{30}{67}.$$

Alternatively, since the sample space is finite and equiprobable,

Example 1.22

The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is $P(D \cap A) = 0.78$.

Find the probability that a plane:

(a) arrives on time given that it departed on time,

(b) departed on time given that it has arrived on time,

(c) arrives on time, given that it did not depart on time.

Solution

(a) The probability that a plane arrives on time given that it departed on time is

$$P(A|D) = \frac{P(D \cap A)}{P(D)} = \frac{0.78}{0.83} = 0.94.$$

(b) The probability that a plane departed on time given that it has arrived on time is

$$P(D|A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

(c) We wish to find $P(A|D')$. Now,

$$P(A|D') = \frac{P(A \cap D')}{P(D')} = \frac{P(A) - P(A \cap D)}{1 - P(D)} = \frac{0.82 - 0.78}{1 - 0.83} = 0.24.$$

The notion of conditional probability provides the capability of re-evaluating the idea of probability of an event in the light of additional information, that is, when it is known that another event has occurred. The probability $P(A|B)$ is an updating of $P(A)$ based on the knowledge that event B has occurred.

1.6 The multiplication rule

If we multiply each side of Equation (1.6) by $P(B)$, we obtain the following multiplication rule, which enables us to calculate the probability that two events will both occur. If in an experiment, the events A and B can both occur, then

$P(A \cap B) = P(B)P(A\ B)$……………………………………………………….(1.8)

Since the events $A \cap B$ and $B \cap A$ are equivalent, it follows from Equation (1.8) that we can also write

$P(A \cap B) = P(B \cap\ A) = P(A)P(B\ )$………………………………………….(1.9)

In other words, it does not matter which event is referred to as A and which event is referred to as B.

Example 1.23

Suppose that we have a fuse box containing 20 fuses, of which 5 are defective. If 2 fuses are selected at random and removed from the box in succession without replacing the first, what is the probability that both fuses are defective?

Solution

Let A denote the event that the first fuse is defective and B the event that the second fuse is defective.

We wish to find $P(A \cap B)$. The probability of first removing a defective fuse is $\frac{5}{20}$. If the first fuse is defective, then the probability of removing a second defective fuse from the remaining 4 is $\frac{4}{19}$. By the multiplication rule,

$$P(A \cap B) = P(A)P(B|A) = \left(\frac{5}{20}\right)\left(\frac{4}{19}\right) = \frac{1}{19}.$$

Example 1.24

Bag 1 contains 4 white balls and 3 green balls, and bag 2 contains 3 white balls and 5 green balls. A ball is drawn from bag 1 and placed unseen in bag 2. Find the probability that a ball now drawn from bag 2 is

(a) green,

(b) (b) white.

Solution

Let G_1, G_2 , W_1 and W_2 represent, respectively, the events of drawing a green ball from bag 1, a green ball from bag 2, a white ball from bag 1 and a white ball from bag 2.

(a) We wish to find $P(G_2)$.

We can express G_2 in the form $G_2 = (G_1 \cap G_2) \cup (W_1 \cap G_2)$.

The events $(G_1 \cap G_2)$ and $(W_1 \cap G_2)$ are mutually exclusive and so

$$
\begin{aligned}
P(G_2) &= P(G_1 \cap G_2) + P(W_1 \cap G_2) \\
&= P(G_1)P(G_2 | G_1) + P(W_1)P(G_2 | W_1).
\end{aligned}
$$

Using the tree diagram in Fig. 1.10, we obtain

$$P(G_2) = \left(\tfrac{3}{7}\right)\left(\tfrac{6}{9}\right) + \left(\tfrac{4}{7}\right)\left(\tfrac{5}{9}\right) = \frac{38}{63}.$$

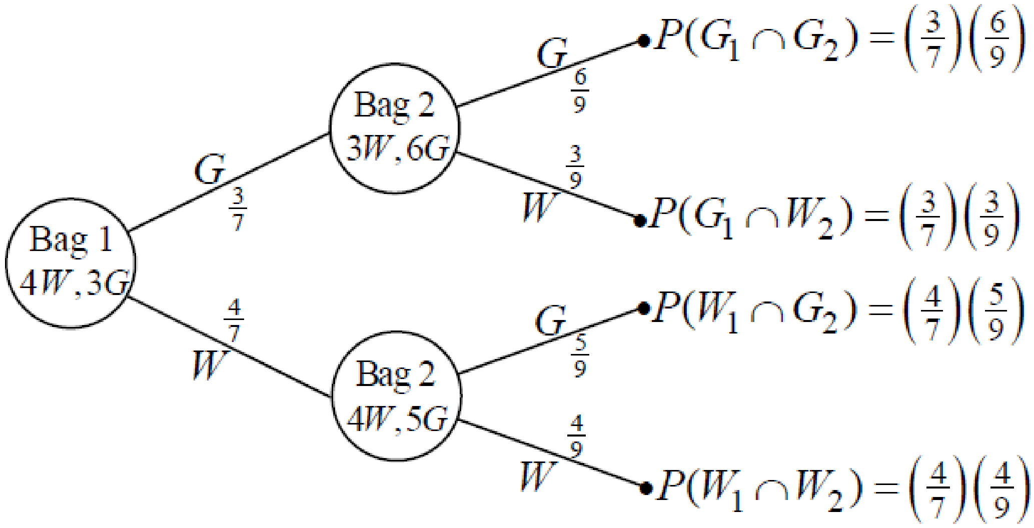


Fig. 1.10: *Tree diagram for Example 1.31*

(b) We wish to find $P(W_2)$. We can express W_2 in the form $W_2 = (G_1 \cap W_2) \cup (W_1 \cap W_2)$. The events $G_1 \cap W_2$ and $W_1 \cap W_2$ are mutually exclusive and so

$$
\begin{aligned}
P(W_2) &= P(G_1 \cap W_2) + P(W_1 \cap W_2) \\
&= P(G_1)P(W_2 | G_1) + P(W_1)P(W_2 | W_1) \\
&= \frac{3}{7} \times \frac{3}{9} + \frac{4}{7} \times \frac{4}{9} = \frac{25}{63}.
\end{aligned}
$$

The multiplication rule can be applied to two or more events. For three events A, B and C, the multiplication rule takes the following form.

Theorem 1.7

$$P(A \cap B \cap C) = P(A)\,P(B|A)\,P(C|\,A \cap B)\text{, where } P(A) \neq 0 \text{ and } P(A \cap B) \neq 0.$$

Proof

By the associative law,

$A \cap B \cap C = (A \cap B) \cap C$. Therefore,

$$\begin{aligned} P(A \cap B \cap C) &= P[(A \cap B) \cap C] \\ &= P(A \cap B)\,P(C|A \cap B) \\ &= P(A)\,P(B|A)\,P(C|A \cap B). \end{aligned}$$

Theorem 1.7 can be extended by mathematical induction to the following theorem.

Theorem 1.8

For any events $A_1, A_2, \ldots, A_n \quad (n > 2)$

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\ldots P(A_n|A_1 \cap A_2 \ldots \cap A_{n-1}).$$

Example 1.25

A box contains 5 red, 4 white and 3 blue balls. If three balls are drawn successively from the box, find the probability that they are drawn in the order red, white and blue if each ball is not replaced.

Solution

Let R be the event "red on first draw", W the event "white on second draw" and B the event "blue on third draw". We wish to find $P(R \cap W \cap B)$. Since there are 5 balls out of 12 balls, $P(R) = \frac{5}{12}$ If the first ball drawn is red, then there are 4 white balls out of the 11 balls remaining in the box. Hence $P(W|R) = \frac{4}{11}$. If the first ball is red and the second ball is white, then there are 3 blue balls out of the

10 balls remaining in the box. It follows that $P(B|R\cap W)=\frac{3}{10}$. Hence by Theorem 1.11,

$$P(R\cap W\cap B)=\left(\frac{5}{12}\right)\left(\frac{4}{11}\right)\left(\frac{3}{10}\right)=\frac{1}{22}.$$

1.4 Independent events

If the events A and B are independent, then the multiplication theorem becomes $P(A\cap B)=P(A)\,P(B)$. This result illustrates the following general result.

> Two events with nonzero probabilities are independent if and only if, any one of the following equivalent statements is true.
>
> (a) $P(A\mid B)=P(A)$, (b) $P(B\mid A)=P(B)$, (c) $P(A\cap B)=P(A)P(B)$.

Definition 1.4

Two events that are not independent are said to be dependent. Usually, physical conditions under which an experiment is performed will enable us to decide whether or not two or more events are independent. In particular, the outcomes of unrelated parts of an experiment can be treated as independent.

Example 1.26

A small town has one fire engine and one ambulance available for emergencies. The probability that a fire engine is available when needed is 0.96, and the probability that the ambulance is available when called is 0.90. In the event of an injury resulting from a burning building, find the probability that both the ambulance and the fire engine will be available.

Solution

Let A and B represent the respective events that the fire engine and the ambulance are available. The two events are independent and so

$$P(A \cap B) = P(A)P(B) = (0.96)(0.90) = 0.864.$$

Example 1.27

A pair of fair dice is thrown twice. Find the probability of getting totals of 7 and 11.

Solution

Let A_1, A_2, B_1 and B_2 be the respective events that a total of 7 occurs on the first throw, a total of 7 occurs on the second throw, a total of 11 occurs on the first throw and a total of 11 occurs on the second throw. We are interested in the probability of the event $(A_1 \cap B_2) \cup (A_2 \cap B_1)$. It is clear that the events A_1, A_2, B_1 and B_2 are independent. Moreover, $A_1 \cap B_2$ and $A_2 \cap B_1$ are mutually exclusive events. Hence,

$$\begin{aligned} P[(A_1 \cap B_2) \cup (A_2 \cap B_1)] &= P(A_1 \cap B_2) + P(A_2 \cap B_1) \\ &= P(A_1)P(B_2) + P(A_2)P(B_1) \\ &= \left(\frac{6}{36}\right)\left(\frac{2}{36}\right) + \left(\frac{6}{36}\right)\left(\frac{2}{36}\right) = \frac{1}{54}. \end{aligned}$$

Definition 1.6

The three events A_1, A_2, and A_3 are independent if and only if:

$$P(A_1 \cap A_2) = P(A_1)P(A_2), \qquad P(A_1 \cap A_3) = P(A_1)P(A_3),$$
$$P(A_2 \cap A_3) = P(A_2)P(A_3), \qquad P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3).$$

That is, three events A_1, A_2, and A_3 are independent if and only if they are pairwise independent, and $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$. It can be proved that if the above equations are satisfied, so is any equation we obtain by replacing an event by its complement on both sides of one of the original

Example 28:

If A and B are independent, Prove that:

1- A^c and B^c are independent.
2- A^c and B are independent.
3- A and B^c are independent.

Solution:

A and B are independent:

$$P(A \cap B) = P(A).P(B)$$

$$P(A^c \cap B^c) = P(A \cup B)^c = 1 - P(A \cup B)$$
$$= 1 - P(A) - P(B) + P(A \cap B)$$
$$= 1 - P(A) - P(B)(1 - P(A))$$
$$= (1 - P(A))(1 - P(B))$$
$$= P(A^c).P(B^c)$$

A^c and B^c are independent.

$$2 - P(A^c \cap B) = P(B - A) = P(B) - P(A \cap B)$$
$$= P(B) - P(A).P(B)$$
$$= P(B)(1 - P(A))$$
$$= P(B).P(A^c)$$

$\therefore A^c \text{ are } B \text{ independent}$

$$3 - P(A \cap B^c) = P(A - B) = P(A) - P(A \cap B)$$
$$= P(A) - P(A).P(B)$$
$$= P(A)(1 - P(B))$$
$$= P(A).P(B^c)$$

$\therefore A \text{ are } B^c \text{ independent}$

1.3 Bayes' theorem

1.3.1 The total probability rule

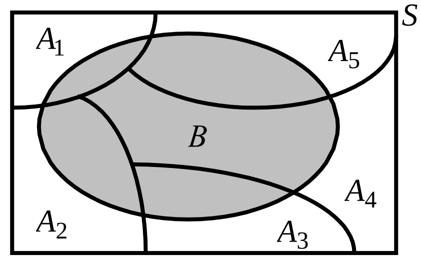


Fig. 1.11: *A partition of a sample space*

In Fig. 1.11, the events A_1, A_2, A_3, A_4, A_5 are mutually exclusive and $S = A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$. These events are said to form a **partition of the sample space** S. By a partition of S, we mean a collection of mutually exclusive events whose union is S. In general, the events A_1, A_2, …, A_n form a partition of the sample space S if

(a) $A_i \neq \phi \ (i = 1, 2, \ldots, n)$, (b) $A_i \cap A_j = \phi \ (i \neq j,\ i, j = 1, 2, \ldots, n)$, (c) $S = \bigcup_{i=1}^{n} A_i$.

In Fig. 1.11, it can be seen that if B is an event defined on the sample space S such that $P(B) > 0$, then

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_5 \cap B).$$

Since $(A_1 \cap B)$, $(A_2 \cap B)$, ..., $(A_5 \cap B)$ are mutually exclusive events,

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \ldots + P(A_5 \cap B) \\ &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \ldots + P(A_5)\ P(B/A_5). \end{aligned}$$

The following theorem gives the general result.

Theorem 1.9

If $A_1, A_2, \ldots, A_n$ form a partition of a sample space S, then for any event B defined on S such that $P(B) > 0$,

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i).$$

This result is called the **total probability rule**.

Example 1.29

In a certain assembly plant, three machines A, B, and C make 30%, 45% and 25%, respectively, of the products. It is known from past experience that 2% of the products made by machine A, 3% of the products made by machine B and 2% of the products made by machine C are defective. If a finished product is selected at random, what is the probability that it is defective?

Solution

Let A_1 denote the event "the finished product was made by machine A",
A_2 denote the event "the finished product was made by machine B",
A_3 denote the event "the finished product was made by machine C",
and let D denote the event "the finished product is defective".
We wish to find $P(D)$.

We are given that:

$$P(A_1)=0.3,\ P(A_2)=0.45,\ P(A_3)=0.25,\ P(D|A_1)=0.02,\ P(D|A_2)=0.03 \text{ and } P(D|A_3)=0.02.$$

A_1, A_2 and A_3 form a partition of the sample space. Hence,

$$\begin{aligned} P(D) &= P(A_1)P(D|A_1)+P(A_2)P(D|A_2)+P(A_3)P(D|A_3) \\ &= 0.3\times 0.02 + 0.45\times 0.03 + 0.25\times 0.02 \\ &= 0.0245. \end{aligned}$$

The probability that a finished product selected at random is defective is 0.0245.

1.9.2 Bayes' theorem

Consider the following example.

Example 1.30

In Example 1.29, if a finished product is found to be defective, what is the probability that it was made by machine A_1?

Solution

We wish to find $P(A_1|D)$. By the multiplication rule,

$$P\left(A_1|D\right) = \frac{P(A_1 \cap D)}{P(D)} = \frac{P(A_1)P(D|A_1)}{P(D)} = \frac{(0.3)(0.02)}{0.0245} = 0.245.$$

The probability that a defective finished product was made by machine A_1 is 0.245.

Example 1.37 was solved by using **Bayes' theorem**. We now state the theorem.

Theorem 1.10 (Bayes' theorem)

Let A_1, A_2, …, A_n be a collection of events which partition a sample space S. Let B be an event defined on S such that $P(B) \neq 0$. Then for any of the events A_j, $(j = 1, 2, \ldots, n)$

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}, \quad \text{for } j = 1, 2, \ldots, n.$$

Proof

By the definition of conditional probability,

$$P\left(A_j|B\right) = \frac{P(A_j \cap B)}{P(B)}.$$

Using Theorem 1.13 in the denominator, we obtain

$$P\left(A_j|B\right) = \frac{P(A_j \cap B)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)} = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)},$$

which completes the proof.

Bayes' theorem was named after the English philosopher and theologian, Reverend Thomas Bayes (1702 – 1761). The theorem is applicable in situations where quantities of the form $P(B|A_i)$ and $P(A_i)$ are known and we wish to determine $P(A_i|B)$. The following example illustrates an application of Bayes' theorem.

Example 1.31

A consulting firm rents cars from three agencies: 30% from agency *A*, 20% from agency *B* and 50% from agency *C*. 15% of the cars from *A*, 10% of the cars from *B* and 6% of the cars from *C* have bad tyres. If a car rented by the firm has bad tyres, find the probability that it came from agency *C*.

Solution

Let A_1 denote the event "the car came from agency *A*",
A_2 denote the event "the car came from agency *B*",
A_3 denote the event "the car came from agency *C*",

and let *T* denote the event "a car rented by the firm has bad tyres". We wish to find $P(A3 \mid T)$. We are

given $P(A_1) = 0.3$, $P(A_2) = 0.2$, $P(A_3) = 0.5$, $P(T|A_1) = 0.15$, $P(T|A_2) = 0.1$ and $P(T|A_3) = 0.06$. A_1, A_2 and A_3 are mutually exclusive and $P(A_1) + P(A_2) + P(A_3) = 1$, and so A_1, A_2 and A_3 form a partition of the sample space. Hence, by Bayes' theorem,

$$
\begin{aligned}
P(A_3|T) &= \frac{P(A_3)P(T|A_3)}{P(A_1)P(T|A_1) + P(A_2)P(T|A_2) + P(A_3)P(T|A_3)} \\
&= \frac{0.5 \times 0.06}{0.3 \times 0.15 + 0.2 \times 0.1 + 0.5 \times 0.06} \\
&= 0.3158.
\end{aligned}
$$

Exercise 1

1. Use Venn diagrams to verify that:

   (a) $A \cup (A \cap B) = A,$ (b) $(A \cap B) \cup (A \cap B') = A,$ (c) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$

2. A experiment involves tossing a red and a green dice, and recording the numbers that come up
   (a) List the elements corresponding to the event, A, that the sum is greater than 8.
   (b) List the elements corresponding to the event, B, that a 2 occurs on either dice.
   (c) List the elements corresponding to the event, C, that a number greater than 4 comes up on the green die.
   (d) List the elements corresponding to the following events
   (i) $A \cap C$, (ii) $A \cap B$, (iii) $B \cap C$.

3. Consider the experiment of rolling two dice.
   (a) Let F = "the sum of the two numbers which appear on the dice is 8". List the sample points in F.
   (b) Let I = "the sum of the two numbers which appear on the dice is even". List the sample points in I.
   (c) Let J = "the number on the red dice is 3". List the sample points in J.
   (d) Let $G = \{(1, 1), (1, 2), (2, 1)\}$,
   $H = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\}$.
   Are the events G and H mutually exclusive?
   (e) List the sample points in $G \cap H$.
   (f) Let $L = \{(2, 2)\}$. Are the events G and L mutually exclusive?

4. In how many ways can 3 of 20 laboratory assistants be chosen to assist with an experiment?

5. Out of 6 mathematicians and 8 physicists, a committee consisting of 3 mathematicians and 4 physicists is to be formed. In how many ways can this be done?

6. Let A and B be events with $P(A) = 0.25$, $P(B) = 0.40$ and $P(A \cap B) = 0.15$.

   Find: (i) $P(A^c \cap B^c)$ (ii) $P(A \cap B^c)$, (iii) $P(A^c \cap B)$.

7. Samples of a cast aluminum part are classified on the basis of surface finish (in microinches) and length measurements. The results of 100 parts are summarized below.

		length	
		excellent	good
surface finish	excellent	75	7
	good	10	8

Let A denote the event that a sample has excellent surface finish, and let B denote the event that a sample has excellent length. Find
(a) $P(A)$, (b) $P(B)$, (c) $P(A^c)$, (d) $P(A \cap B)$, (e) $P(A \cup B)$.

8. In a space lot, the primary computer system is backed up by two secondary systems. They operate independently of one another and each is 90% reliable. Find the probability that all three systems will be operable at the time of the launch.

9. Show that if A_1 and A_2 are independent, then A_1 and A^c_2 are also independent.

10. Kofi feels that the probability that he will get an A in the first Physics test is 0.5 and the probability that he will get A's in the first and second Physics tests is 1/3 .

    If Kofi is correct, what is the conditional probability that he will get an A in the second test, given that he gets an A in the first test?

11. In rolling 2 balanced dice, if the sum of the two values is 7, what is the probability that one of the values is 1?

12. A random sample of 200 adults are classified below by sex and their level of education attained.

Education	Male	Female
Elementary	38	45
High School	28	50
University	22	17

If a person is chosen at random from this group, find the probability that:

(a) the person is a male, given that the person has High School education,

(b) the person does not have a university degree, given that the person is a female.

13. In an experiment to study the relationship of hypertension and smoking habits,

the following data were collected for 180 individuals.

	Non-smokers	Moderate smokers	Heavy smokers
Hypertension	21	36	30
No hypertension	48	26	19

If one of these individuals is selected at random, find the probability that the person is

(a) experiencing hypertension, given that he/she is a heavy smoker;

(b) a non-smoker, given that he/she is experiencing no hypertension.

14. The probability that a married man watches a certain television show is 0.4 and the probability that a married woman watches the show is 0.5. The probability that a man watches the show, given that his wife does is 0.7. Find the probability that
(a) a married couple watches the show;
(b) a wife watches the show given that her husband does;
(c) at least 1 person of a married couple will watch the show.

15. A town has 2 fire engines operating independently. The probability that a specific engine is available when needed is 0.96.
(a) What is the probability that neither is available when needed?
(b) What is the probability that exactly one fire engine is available when needed?

16. A factory employs three machine operators, George, Andrew and Eric, to produce its brand of goods. George works 45% of the time, Andrew works 30% of the time and Eric works 25% of the time. Each operator is prone to produce defective items. George produces defective items 2% of the time, Andrew produces defective items 4% of the time while Eric produces defective items 6% of the time. If a defective item is produced, what is the probability that it was produced by Andrew?

17. In a certain assembly plant, three machines B1, B2 and B3, make 30%, 45% and 25%, respectively, of the products. It is known from past experience that 2%, 3%, and 2% of the products made by B1, B2 and B3, respectively, are defective.

(a) If a finished product is selected at random, what is the probability that it is defective?

(b) If a finished product is found to be defective, what is the probability that it was produced by B3?

18. A large industrial firm uses local hotels A, B and C to provide overnight accommodation for its clients. From past experience, it is known that 20% of the clients are assigned rooms at hotel A, 50% at hotel B, and 30% at hotel C. If the plumbing is faulty in 5% of the rooms at hotel A, in 4% of the rooms at hotel B, and in 8% of the rooms at hotel C, what is the probability that:

(a) a client will be assigned a room with faulty plumbing?

(b) a person with a room having faulty plumbing was assigned accommodation at hotel B?

19. Suppose that at a certain accounting office, 30%, 25% and 45% of the statements are prepared by Mr. George, Mr. Charles and Mrs. Joyce, respectively. These employees are very reliable. Nevertheless, they are in error some of the time. Suppose that 0.01%, 0.005% and 0.003% of the statements prepared by Mr. George, Mr. Charles and Mrs. Joyce, respectively, are in error. If a statement from the accounting office is in

20. error, what is the probability that it was prepared (caused) by Mr. George?

21. A certain construction company buys 20%, 30%, and 50% of their nails from hardware suppliers A, B, and C, respectively. Suppose it is known that 0.05%, 0.02% and 0.01% of the nails from A, B, and C, respectively, are defective.

(a) What percentage of the nails purchased by the construction company are defective?

(b) If a nail purchased by the construction company is defective, what is the probability that it came from the supplier C?

CHAPTER TWO

Random Variables and Probability Distributions

2.1 The concept of a random variable

Results of random experiments are often summarized in terms of numerical values. Consider, for example, the experiment of testing two electronic components. When an electronic component is tested, it is either defective or non-defective. The sample space of the experiment may therefore be written as $S = \{NN, DN, ND, DD\}$, where N denotes non-defective and D denotes defective.

Let X denote the number of electronic components which are defective. One is naturally interested in the possible values of X. Can X take the value 3? What about the value 1.5? The values X can take are 0, 1 and 2. Notice that X takes the value 0 at the sample point NN and the value 1 at the sample points DN and ND. What value does X take at the sample point DD?

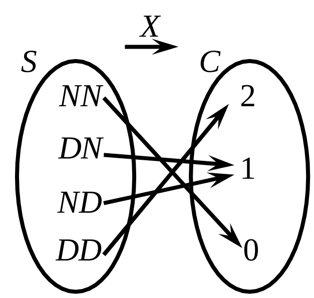


Fig. 2.1

It can be seen that X assigns a unique real number $X(s)$ to each sample point s of S (see Fig. 2.1). X is therefore a function with domain S and co-domain $C = \{0, 1, 2\}$. Such a function is called a **random variable.**

Definition 2.1:

> A random variable is a function that assigns a real number to each element in the sample space of a random experiment.

A random variable is denoted by an uppercase letter, such as X, and a corresponding lowercase letter, such as x, is used to denote a possible value of X. We refer to the set of possible values of a random variable X as ***the range of X.***

Example 2.1

Three balls are drawn in succession without replacement from a box containing 5 white and 4 green balls. Let Y denote the number of white balls selected. The possible outcomes and the values y of Y are:

outcome	*GGG*	*GGW*	*GWG*	*WGG*	*GWW*	*WGW*	*WWG*	*WWW*
y	0	1	1	1	2	2	2	3

where G denotes "green" and W denotes "white" and the i^{th} letter in a triple, denotes the colour of the i^{th} ball drawn (i = 1, 2, 3). For example, *GWG* means the first ball drawn is green, the second ball drawn is white and the third ball drawn is green.

Example 2.2

A coin is tossed three times, X is the number of heads.

S={HHH, HTH, HHT, THH, HTT, THT, TTH, TTT}

X(HHH)=3, X(HTH)=2, …, X(TTT)=0

$$\mathrm{P}(3) = P_r(\mathrm{X} = 3) = P_r(\mathrm{HHH}) = \frac{1}{8}$$

$$\mathrm{P}(2) = P_r(\mathrm{X} = 2) = P_r(\mathrm{HTH, HHT, THH}) = \frac{3}{8}$$

$$\mathrm{P}(1) = P_r(\mathrm{X} = 1) = P_r(\mathrm{HTT, THT, TTH}) = \frac{3}{8}$$

$$\mathrm{P}(0) = P_r(\mathrm{X} = 0) = P_r(\mathrm{TTT}) = \frac{1}{8}$$

X	0	1	2	3	Sum
$P(X=x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\sum P(x)=1$

$$P(x)=\begin{cases} \frac{1}{8} & , x=0,3 \\ \frac{3}{8} & , x=1,2 \\ 0 & o.w \end{cases}$$

Definition 2.2 (Discrete random variable)

A random variable is discrete if it can assume a finite or a countably infinite set of values.

Definition 2.3 (Continuous random variable)

If the range of a random variable X contains an interval (either finite or infinite) of real numbers, then X is a ***continuous random variable***.

In most practical problems, continuous random variables represent measured data, such as heights, weights, temperatures, distances, or life periods, whereas discrete random variables represent count data, such as the number of defectives in a sample of n items or the number of road traffic accidents in Accra in a week.

2.2 Discrete probability distributions

When dealing with a random variable, it is not enough just to determine what values are possible. We also need to determine what is probable. Consider the following example.

Example 2.4

Fifty taxi drivers were asked of the number of road traffic accidents they have had in a year. The results are given in Table 2.1.

Table 2.1: Number of accidents per year of 50 taxi drivers

Number of accidents	0	1	2	3	4	5
Frequency	15	12	9	7	5	2

Suppose we select a taxi driver from this group and X is the number of road traffic accidents the person selected had in a year. What values can X take? X can take the values 0, 1, 2, 4 and 5. What is the probability that $X = 0$? Out of the 50 equally likely ways of selecting a taxi driver, there are 15 ways in which the taxi driver selected had no accident in a year. Hence the probability that $X = 0$ is

$\frac{15}{50} = 0.30.$

This is written as

$$P(X = 0) = 0.30$$

Similarly,

$$P(X = 1) = \frac{12}{50} = 0.30$$

Table 2.2 gives the Possible values x, of X and their probabilities. Table 2.2 is called the **probability distribution** of X. Note that the values of X exhaust all possible cases and hence the probabilities add up to 1.

Table 2.2:* The probability distribution of *X

x	0	1	2	3	4	5
$P(X = x)$	0.30	0.24	0.18	0.14	0.10	0.04

Example 2.5

A shipment of 8 similar microcomputers to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution of the number of defectives.

Solution

Let X be the number of defective computers purchased by the school. X can take the values 0, 1, and 2. Now,

$$P(X=0)=\frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}}=\frac{10}{28},\qquad P(X=1)=\frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}}=\frac{15}{28},\qquad P(X=2)=\frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}}=\frac{3}{28}.$$

The probability distribution of X is given in the following table.

x	0	1	2
$P(X=x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

Representation of the probability distribution of a discrete random variable

The probability distribution of a discrete random variable X can be represented by a table, a formula or a graph.

Tabular form

For a random variable that can assume a small number of values, it is simplest to present its probability distribution in the form of a table having two rows: the upper row contains the possible values the random variable assumes and the lower row contains the corresponding probabilities of the values (see Table 2.3).

Table 2.3: **The probability distribution of X**

x	x_1	x_2	…	x_n
$P(X=x)$	$p(x_1)$	$p(x_2)$	…	$p(x_n)$

Formula

Frequently, it is convenient to represent the probability distribution of a discrete random variable by a formula. For example,

$$f(x)=\frac{1}{7},\qquad x=1,2,\ldots,7$$

defines a probability distribution of a discrete random variable.

The values of a discrete random variable are often called **mass points**; and $f(x_j)$ denotes the **mass** associated with the mass point x_j. The function $f(x)=P(X=x)$ is therefore called the **probability mass function** of the random variable X. Other terms used are **frequency function** and **probability function**. Also, the notation $p(x)$ is sometimes used instead of $f(x)$ for probability mass functions.

Definition 2.4 Probability mass function

A function $f(x)$ is the probability mass function of a discrete random variable X if it has the following two properties:

$$(1)\quad p(x) \geq 0 \text{ for all } x, \qquad (2)\quad \sum_{\text{all } x} p(x) = 1.$$

Graphical form

The probability distribution of a discrete random variable can also be represented graphically, as shown in Fig. 2.2. Such a graph is called a **probability graph** (see Fig. 2.2).

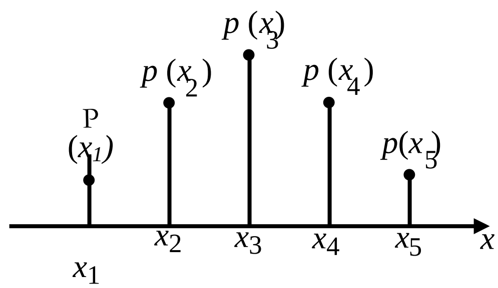


Fig. 2.2: *Probability graph*

The probability distribution of a discrete random variable can also be represented by a **probability histogram** (see Fig. 2.3). Similar to the probability graph, the height of each rectangle of a probability histogram is equal to the probability that the random variable takes on the value which corresponds to the mid-point of the base.

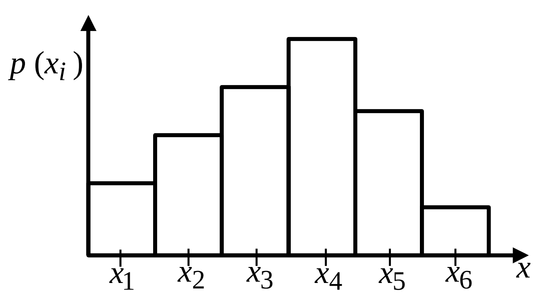


Fig. 2.3: *Probability histogram*

Random variables are so important in random experiments that sometimes we essentially ignore the original sample space of the experiment and focus on the probability distribution of the random variable. In this manner, a random variable can simplify the description and analysis of a random experiment.

Example 2.6

Determine whether each of the following can serve as a probability mass function of a discrete random variable:

(a) $f(x) = \frac{1}{2}(x-2)$, $x = 1, 2, 3, 4.$ (b) $g(x) = \frac{1}{10}(x+1)$, $x = 0, 1, 2, 3.$

(c) $h(x) = \frac{1}{20}x^2$, $x = 0, 1, 2, 3, 4.$

Solution

(a) $f(1) = \frac{1}{2}(1-2) = -\frac{1}{2}$. $f(1)$ is negative and so $f(x)$ cannot serve as a probability mass function of a discrete random variable.

(b) $g(x) \geq 0$ for all values of x, and $\sum_{x=0}^{3} g(x) = \frac{1}{10}(1+2+3+4) = 1$. $g(x)$ is therefore the probability mass function of a discrete random variable.

(a)(c) $h(x) \geq 0$ for all values of x and $\sum_{x=0}^{4} h(x) = \frac{1}{20}(0+1+4+9+16) = 1.5$. $\sum_{x=0}^{4} h(x) \neq 1$ and so $h(x)$ cannot serve as a probability mass function of a discrete random variable.

Example 2.7

The following table gives the probability distribution of a random variable X.

x	1	2	3	4
$p(x)$	$\frac{1}{4}$	c	$\frac{1}{2}$	$\frac{1}{8}$

(a) Find the value of the constant c.

(b) Represent the probability distribution by:

(i) a probability graph, (ii) a probability histogram.

(c) Find (i) $P(X > 1)$, (ii) $P(0 < X < 2)$, (iii) $P(X \geq 2)$.

Solution

(a) $\sum_{x=0}^{3} p(x) = 1 \quad \Rightarrow \quad \frac{1}{4} + c + \frac{1}{2} + \frac{1}{8} = 1 \quad \Rightarrow \quad c = \frac{1}{8}$.

(b) (i) Fig. 2.4 is the required probability graph.
(ii) Fig. 2.5 is the required probability histogram.

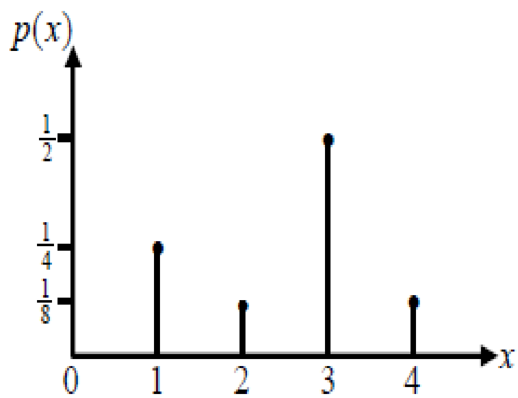


Fig. 2.4: Probability graph for Example 2.7

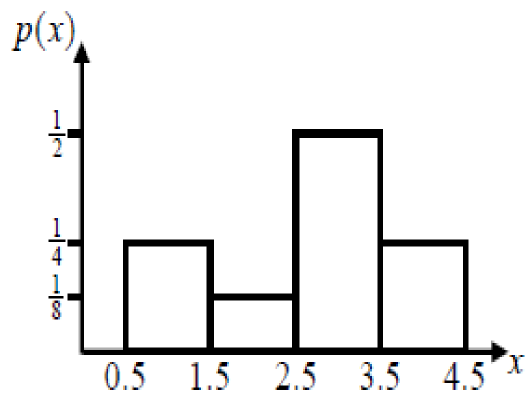


Fig. 2.5: Probability histogram for Example 2.7

(c) (i) $P(X > 1) = 1 - P(X \le 1) = 1 - P(X = 1) = 1 - \frac{1}{4} = \frac{3}{4}$.

Alternatively,

$P(X > 1) = P(X = 2) + P(X = 3) + P(X = 4) = \frac{1}{8} + \frac{1}{2} + \frac{1}{8} = \frac{3}{4}$.

(ii) $P(0 < X < 2) = P(X = 1) = \frac{1}{4}$.

(iii) $P(X \ge 2) = P(X = 2) + P(X = 3) + P(X = 4) = \frac{1}{8} + \frac{1}{2} + \frac{1}{8} = \frac{3}{4}$.

Alternatively,

$P(X \ge 2) = 1 - P(X < 2) = 1 - P(X = 1) = 1 - \frac{1}{4} = \frac{3}{4}$.

2.3 Continuous probability distributions

In Section 2.2, we learnt that for a complete characterization of a discrete random variable, it is necessary and sufficient to know the probability mass function of the random variable. Corresponding to every continuous random variable X, there is a function f, called the **probability density function** (p.d.f.) of X such that

(a) $f(x) \geq 0$, (b) $\int_{-\infty}^{\infty} f(x)dx = 1$, (c) $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$.

For a complete characterization of a continuous random variable, it is necessary and sufficient to know the p.d.f. of the random variable.

Geometrically, relation (c) means the following: the probability that a continuous random variable X takes values in the interval (a, b) is equal to the area of the region defined by the p.d.f. of X, the straight lines $x = a$ and $x = b$, and the x-axis (see Fig. 2.6).

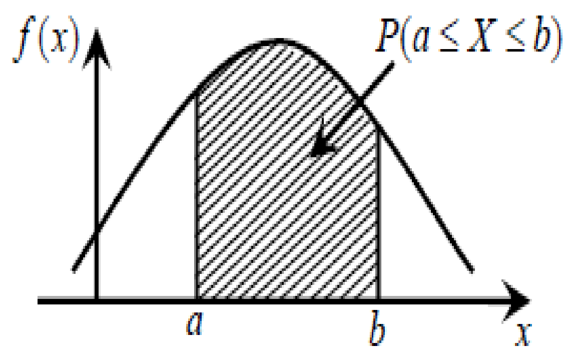


Fig. 2.6: Probability as an area under a p.d.f.

A consequence of X being a continuous random variable is that for any value in the range of X, say x,

$$P(X = x) = \int_{x}^{x} f(t)dt = 0 \quad \text{......(2.1)}$$

Thus a continuous random variable has a probability of zero of assuming **exactly** any of its values. At first this may seem startling, but it becomes more plausible when we consider an example. Consider a random variable whose values are the heights of all people over twenty years of age. Between any two values, say 162.99 and 163.01 centimetres, there are infinite number of heights, one of which is 163 centimetres. The probability of selecting a person at random who is **exactly** 163 centimetres tall and not one of the infinitely large set of heights so close to 163 centimetres that you cannot humanly measure the difference, is remote, and thus we assign a probability of zero to the event.

As an immediate consequence of Equation (2.1), if X is a continuous random variable, then for any numbers a and b, with $a \le b$,

$$P(a \le X \le b) = P(a \le X < b) = P(a < X \le b) = P(a < X < b). \quad \text{........(2.2)}$$

That is, it does not matter whether we include an endpoint of the interval or not. This is not true, though, when X is discrete.

Example 2.8

(a) Show that

$$f(x) = \begin{cases} \frac{1}{9}x^2, & 0 < x < 3, \\ 0, & \text{elsewhere.} \end{cases}$$

is the p.d.f. of a continuous random variable X.

(b) Sketch the graph of $f(x)$.

Solution

(a) We have to show that $f(x) \ge 0$ for all x and $\int_{-\infty}^{\infty} f(x)dx = 1$. It is clear that $f(x) \ge 0$ for all x.

Moreover,
$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{0} f(x)dx + \int_{0}^{3} f(x)dx + \int_{3}^{\infty} f(x)dx. \\ &= \int_{-\infty}^{0} 0\,dx + \int_{0}^{3} \frac{1}{9}x^2 dx + \int_{3}^{\infty} 0\,dx \\ &= 0 + \left[\frac{1}{27}x^3\right]_0^3 + 0 = 1. \end{aligned}$$

Hence, $f(x)$ is the p.d.f. of a continuous random variable.

(b) Fig. 2.7 shows a sketch of the graph of $f(x)$.

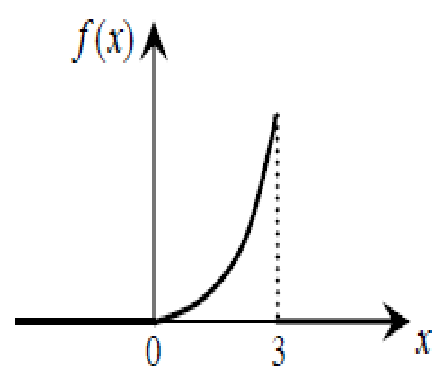


Fig. 2.7: The p.d.f. of X

Example 2.9

Refer to Example 2.8. Find:

(a) $P(X \le 2)$, (b) $P(X > 1)$, (c) $P(2 \le X \le 3)$.

Solution

$$\text{(a)} \quad P(X \le 2) = \int_{-\infty}^{2} f(x)\,dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{2} f(x)dx$$

$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{2} \tfrac{1}{9}x^2 dx$$

$$= 0 + \left[\tfrac{1}{27}x^3\right]_0^2 = \tfrac{8}{27}.$$

$$\text{(b)} \quad P(X > 1) = \int_{1}^{\infty} f(x)\,dx = \int_{1}^{3} f(x)dx + \int_{3}^{\infty} f(x)dx$$

$$= \int_{1}^{3} \frac{1}{9}x^2 dx + \int_{3}^{\infty} 0\,dx$$

$$= \left[\frac{1}{27}x^3\right]_1^3 + 0 = \frac{1}{27}(27-1) = \frac{26}{27}.$$

$$\text{(c)} \quad P(2 \le X \le 3) = \int_{2}^{3} \frac{1}{9}x^2 dx = \left[\frac{1}{27}x^3\right]_2^3 = \frac{1}{27}(27-8) = \frac{19}{27}.$$

Example 2.10

A random variable X has the p.d.f. given by

$$f(x) = \begin{cases} c\sqrt{x}, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the value of the constant c.

(b) Calculate $P\left(X < \frac{1}{4}\right)$.

Solution

(a) The value of c is given by the equation

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} f(x)dx + \int_{0}^{1} f(x)dx + \int_{1}^{\infty} f(x)dx$$

$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{1} c\sqrt{x}\,dx + \int_{1}^{\infty} 0\,dx = 0 + \left[\frac{cx^{\frac{3}{2}}}{3/2}\right]_0^1 + 0 = \tfrac{2}{3}c$$

$$\Rightarrow c = 1\tfrac{1}{2}.$$

Thus, $f(x) = \begin{cases} \frac{3}{2}\sqrt{x}, & 0 < x < 1, \\ 0, & \text{elsewhere.} \end{cases}$

(b) $$P\left(X < \tfrac{1}{4}\right) = \int_{-\infty}^{\frac{1}{4}} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\frac{1}{4}} f(x)\,dx$$

$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{\frac{1}{4}} \tfrac{3}{2}\sqrt{x}\,dx = 0 + \left[x^{\frac{3}{2}}\right]_0^{\frac{1}{4}} = \tfrac{1}{8}.$$

The following example illustrates how conditional probability can be applied to random variables.

Example 2.11

A random variable X has p.d.f. given by

$$f(x)=\begin{cases}\frac{1}{10}, & 0<x<10,\\ 0, & \text{elsewhere.}\end{cases}$$

Find: (a) $P(X>8 | X>5)$, (b) $P(X>7 | X<9)$.

Solution

(a)
$$\begin{aligned} P(X>8 | X>5) &= \frac{P(X>8,\, X>5)}{P(X>5)}=\frac{P(X>8)}{P(X>5)} \\ &= \left(\int_8^{10}\tfrac{1}{10}\,dx\right)\Big/\left(\int_5^{10}\tfrac{1}{10}\,dx\right) \\ &= \left[\tfrac{1}{10}x\right]_8^{10}\Big/\left[\tfrac{1}{10}x\right]_5^{10} = (10-8)/(10-5) = \tfrac{2}{5}. \end{aligned}$$

(b)
$$\begin{aligned} P(X>7 | X<9) &= \frac{P(X>7,\, X<9)}{P(X<9)}=\frac{P(7<X<9)}{P(X<9)} \\ &= \left(\int_7^{9}\tfrac{1}{10}\,dx\right)\Big/\left(\int_0^{9}\tfrac{1}{10}\,dx\right) \\ &= \left[\tfrac{1}{10}x\right]_7^{9}\Big/\left[\tfrac{1}{10}x\right]_0^{9} = (9-7)/(9-0) = \tfrac{2}{9}. \end{aligned}$$

2.4 The cumulative distribution function

There are many problems where we may wish to compute the probability that the observed value of a random variable X will be less than or equal to some real number x. For example, what are the chances that a certain candidate will get no more than 30% of the votes? What are the chances that the prices of gold will remain at or below $800 per ounce? Writing $F(x) = P(X \le x)$ for every real number x, we define $F(x)$ to be the cumulative distribution function of X, or more simply, the distribution function of the random variable X.

2.4.1 The cumulative distribution function of a discrete random variable

Definition 2.5 (Cumulative distribution function)

> The cumulative distribution function $F(x)$ of a discrete random variable X with probability mass function $f(x)$ is defined by
>
> $$F(x) = P(X \le x) = \sum_{x_i \le x} f(x_i).$$

If X takes on only a finite number of values $x_1, x_2, \ldots, x_n$, then the cumulative distribution function of X is given by

$$F(x) = \begin{cases} 0, & -\infty < x < x_1, \\ f(x_1), & x_1 \le x < x_2, \\ f(x_1) + f(x_2), & x_2 \le x < x_3, \\ \vdots & \vdots \\ f(x_1) + f(x_2) + \ldots + f(x_n) = 1, & x_n \le x < \infty. \end{cases}$$

Fig. 2.8 depicts the graph of $F(x)$. It can be seen that $F(x)$ is discontinuous at the points $x_1, x_2, \ldots, x_n$. At these points, $F(x)$ is continuous from the right but discontinuous from the left. Because of the shape of its graph, the cumulative distribution function of a discrete random variable is called a **staircase function** or a **step function**. Notice that $F(x)$ has a jump of height $f(x_i)$ at the point x_i and is constant in the interval (x_i, x_{i+1}).

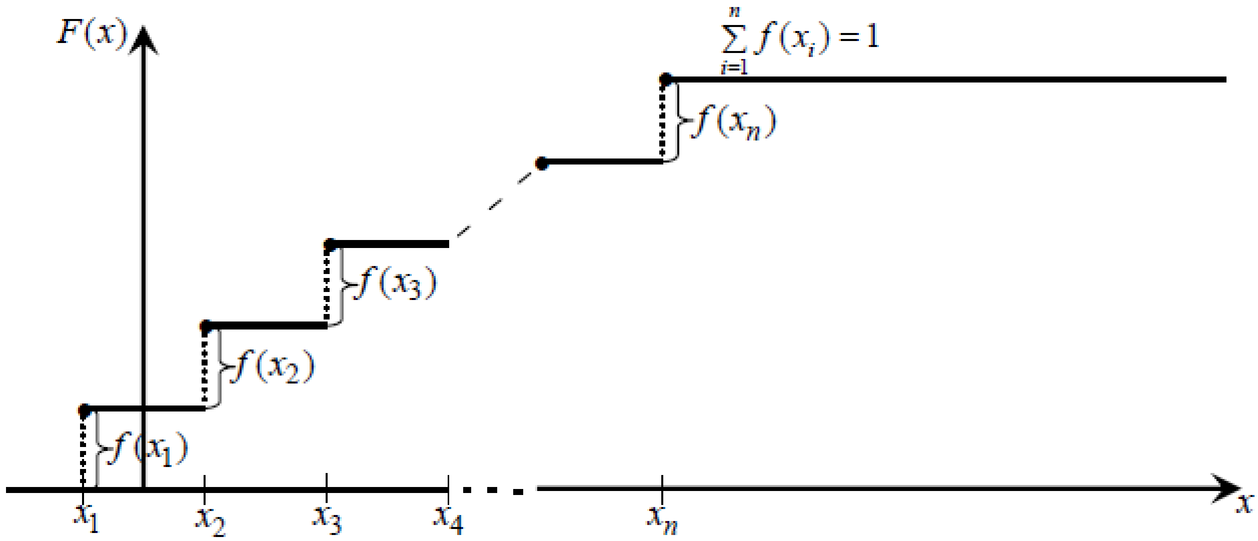


***Fig. 2.8**: A typical cumulative distribution function of a discrete random variable*

Properties of $F(x)$

The cumulative distribution function of a discrete random variable has the following properties:

(1) $0 \le F(x) \le 1$ for all x.

(2) If $x \le y$, then $F(x) \le F(y)$. This means that F is a non-decreasing function.

(3) $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.

(4) F is right-continuous. That is, $\lim_{x \to a+} F(x) = F(a)$ for all points a.

Notice that $F(x)$ is not continuous from the left, that is, $\lim_{x \to a-} F(x) \ne F(a)$ for all points a. This is because we have defined $F(x)$ by $F(x) = P(X \le x)$. If $F(x)$ is defined by $F(x) = P(X < x)$, it will be continuous from the left, but not from the right.

Example 2.12

The following table gives the probability mass function of X. Find the cumulative distribution function of X and sketch its graph.

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Solution

If $x < 0$, $F(x) = P(X \le x) = 0$.

If $0 \le x < 1$, $F(x) = f(0) = \frac{1}{16}$.

If $1 \le x < 2$, $F(x) = f(0) + f(1) = \frac{1}{16} + \frac{1}{4} = \frac{5}{16}$.

If $2 \le x < 3$, $F(x) = f(0) + f(1) + f(2) = \frac{1}{16} + \frac{1}{4} + \frac{3}{8} = \frac{11}{16}$.

If $3 \le x < 4$, $F(x) = f(0) + f(1) + f(2) + f(3) = \frac{1}{16} + \frac{1}{4} + \frac{3}{8} + \frac{1}{4} = \frac{15}{16}$.

If $x \ge 4$, $F(x) = f(0)+f(1)+f(2)+f(3)+f(4) = \frac{1}{16} + \frac{1}{4} + \frac{3}{8} + \frac{1}{4} + \frac{1}{16} = 1$.

The cumulative distribution function of X is therefore given by

$$F(x) = \begin{cases} 0, & x < 0, \\ \frac{1}{16}, & 0 \le x < 1, \\ \frac{5}{16}, & 1 \le x < 2, \\ \frac{11}{16}, & 2 \le x < 3, \\ \frac{15}{16}, & 3 \le x < 4, \\ 1, & x \ge 4. \end{cases}$$

Fig. 2.9 shows the graph of $F(x)$.

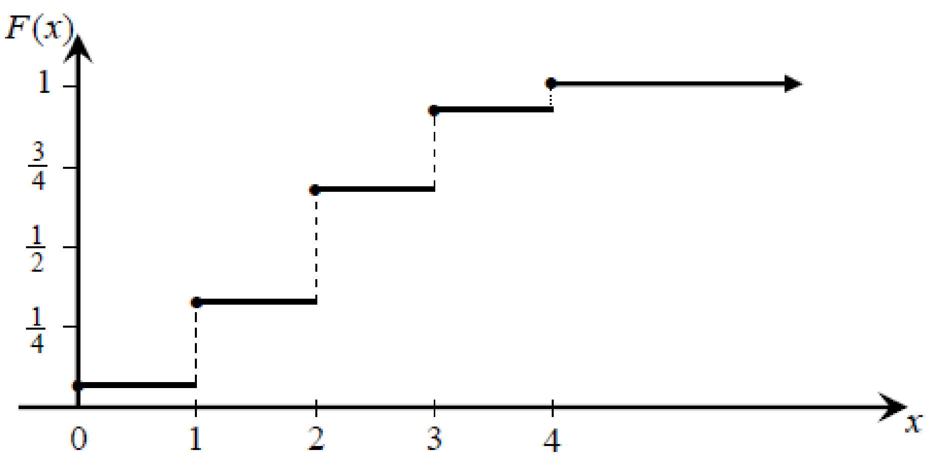


***Fig. 2.9**: The cumulative distribution function for Example 2.12*

Notice that even if the random variable X can assume only integers, the cumulative distribution function of X can be defined for non-integers. For example, in Example 2.12, $F(1.5) = \frac{5}{16}$, $F(2.5) = \frac{11}{16}$

To find the probability mass function, $f(x)$, corresponding to a given cumulative distribution function, $F(x)$, we first find the points where $F(x)$ is discontinuous. We then find the magnitudes of the jumps at these points. If $F(x)$ has a jump at the point x_i, then $f(x_i)$ is equal to the magnitude of this jump. Example 2.13 illustrates the procedure.

Example 2.13

Suppose the cumulative distribution function of X is

$$F(x) = \begin{cases} 0, & x < -2. \\ 0.2, & -2 \le x < 0. \\ 0.7, & 0 \le x < 2. \\ 1, & 2 \le x. \end{cases}$$

Determine the probability mass function of X.

Solution

$F(x)$ is discontinuous at the points –2, 0 and 2. The jump at the point –2 is 0.2 – 0 = 0.2 and so $f(-2)$ = 0.2. The jump at the point 0 is 0.7 – 0.2 = 0.5, and so $f(0) = 0.5$. Similarly $f(2) = 1 - 0.7 = 0.3$. The following table gives the probability mass function of X.

x	–2	0	2
$f(x)$	0.2	0.5	0.3

2.4.2 The cumulative distribution function of a continuous random variable

As in the case of discrete random variables, we are often interested in the probability that the value of a particular continuous random variable will be less than or equal to a given number. Again, as in the case of discrete random variables, the mathematical function used to designate a probability of this type is called a **cumulative distribution function**. The formal definition follows.

Definition 2.6 (Cumulative distribution function)

Let X be a continuous random variable with probability density function f. The function $F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$ for $-\infty < x < \infty$, is called the cumulative distribution function or the distribution function of the random variable X.

As an immediate consequence of Definition 2.6, we can write the following two results

$$P(a < X < b) = \int_{-\infty}^{b} f(t)\,dt - \int_{-\infty}^{a} f(t)\,dt = F(b) - F(a) \quad \text{.......... (2.3)}$$

and $$f(x) = \frac{d}{dx}F(x) \quad \text{.......... (2.4)}$$

Example 2.14

A continuous random variable X has the p.d.f.

$$f(x) = \begin{cases} \frac{1}{3}x^2, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the distribution function of X and sketch its graph.

(b) Find $P(0 < X \le 1)$.

Solution

(a) If $x \le -1$, then

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0.$$

If $-1 \le x \le 2$, then

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{x} \tfrac{1}{3}t^2\,dt = \left[\tfrac{1}{9}t^3\right]_{-1}^{x} = \tfrac{1}{9}(x^3+1)$$

If $x \ge 2$, then

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-1} f(t)\,dt + \int_{-1}^{2} f(t)\,dt + \int_{2}^{x} f(t)\,dt \\ &= \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{2} \tfrac{1}{3}t^2\,dt + \int_{2}^{x} 0\,dt \\ &= 0 + \left[\tfrac{1}{9}t^3\right]_{-1}^{2} + 0 = \tfrac{1}{9}(8+1) = 1. \end{aligned}$$

Therefore,

$$F(x) = \begin{cases} 0, & x \le -1, \\ \frac{1}{9}(x^3+1), & -1 \le x \le 2, \\ 1, & x \ge 2. \end{cases}$$

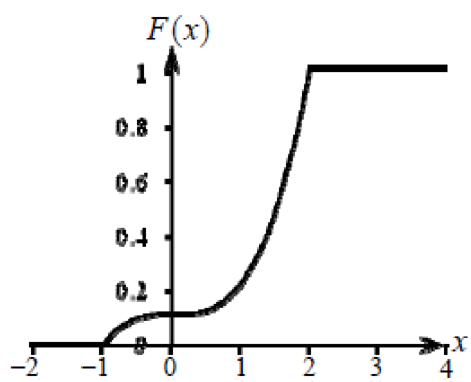


Fig. 2.10: The distribution function for Example 2.14

Fig. 2.10 shows the graph of $F(x)$. Notice that $F(x)$ is a continuous, non-decreasing function.

(b) $P(0 < X \le 1) = F(1) - F(0) = \tfrac{1}{9}(1^2 + 1) - \tfrac{1}{9}(0 + 1) = \tfrac{1}{9}.$

Example 2.15

Let X be a continuous random variable with cumulative distribution function given by

$$F(x) = \begin{cases} 0, & x \le 0, \\ \frac{1}{4}x^2, & 0 \le x \le 2, \\ 1, & x \ge 2. \end{cases}$$

Find the p.d.f. of X.

Solution

By applying Equation (2.4) on page 51, we find that the p.d.f. of X is given by

$$f(x) = \frac{d\,F(x)}{dx} = \begin{cases} \frac{1}{2}x, & 0 \le x \le 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Properties of $F(x)$

Let $F(x)$ be the distribution function of a continuous random variable. $F(x)$ has the following properties:

1. $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0, \qquad F(+\infty) = \lim\limits_{x \to \infty} F(x) = 1.$
2. The function $F(x)$ is the probability of an event, and so $0 \le F(x) \le 1, \quad -\infty < x < \infty.$
3. $F(x)$ is continuous everywhere.
4. $\frac{d}{dx}F(x)$ exists at all points x.
5. $F(x)$ is a non-decreasing function of x, that is, if $x_1 \le x_2$, then $F(x_1) \le F(x_2)$.

2.5 Expected Value of Random Variables

A random variable X is characterized by its probability density function, which defines the relative likelihood of assuming one value over the others. In Chapter 3, we have seen that given a probability density function f of a random variable X, one can construct the distribution function F of it through summation or integration. Conversely, the density function f(x) can be obtained as the marginal value or derivative of F (x). The density function can be used to infer a number of characteristics of the underlying random variable. The two most important attributes are measures of location and dispersion. In this section, we treat the measure of location and treat the other measure in the next section.

Definition 2.5.1 Let X be a random variable with space R_X and probability density function f(x). The mean μ_X $(E(X))$ of the random variable X is defined as

$$\mu_X = \begin{cases} \sum\limits_{x \in R_X} x\, f(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x\, f(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

if the right hand side exists.

The mean of a random variable is a composite of its values weighted by the corresponding probabilities. The mean is a measure of central tendency: the value that the random variable takes "on average." The mean is also called the expected value of the random variable X and is denoted by E(X). The symbol E is called the expectation operator. The expected value of a random variable may or may not exist.

Properties of Expected Values

1- E(c) = c

$$E(c) = \begin{cases} \sum\limits_{-\infty}^{\infty} c.P(x) = c.\sum\limits_{-\infty}^{\infty} P(x) = c \\ \int\limits_{-\infty}^{\infty} c.f(x)dx = c.\int\limits_{-\infty}^{\infty} f(x)dx = c \end{cases}$$

2- E(ax+b) = a E(x)+b

$$E(ax+b)=\begin{cases}\sum_{-\infty}^{\infty}(ax+b).P(x)=a\sum_{-\infty}^{\infty}x.P(x)+\sum_{-\infty}^{\infty}b.P(x)=aE(x)+b\\ \int_{-\infty}^{\infty}(ax+b).f(x)dx=a.\int_{-\infty}^{\infty}x.f(x)dx+\int_{-\infty}^{\infty}b.f(x)dx=aE(x)+b\end{cases}$$

3- $y=g(x)$

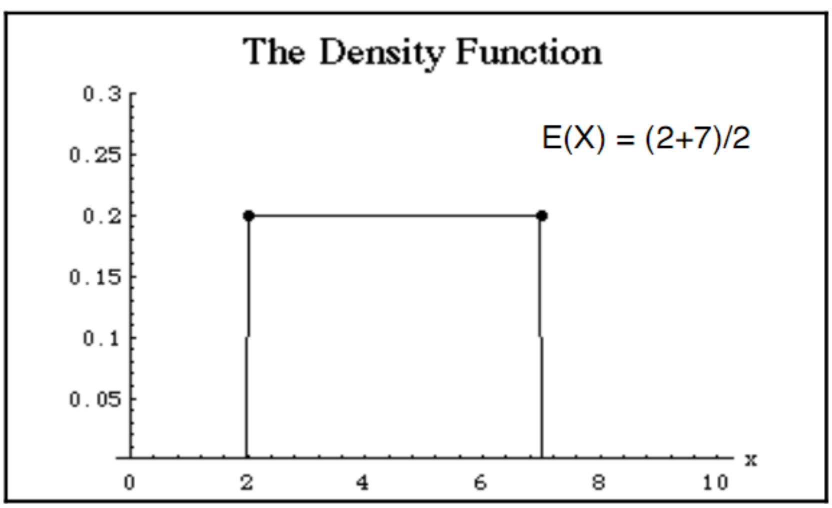
$$E(y)=E(g(x))=\begin{cases}\sum_{-\infty}^{\infty}g(x).P(x)\\ \int_{-\infty}^{\infty}g(x).f(x)dx\end{cases}$$

4- $E(g_1(x)+g_2(x))=E(g_1(x))+E(g_2(x))$

$$5-E(a+bx)^n=\sum_{i=0}^{n}\binom{n}{i}a^{n-i}b^iE(x^i)$$

$$(a+bx)^n=\sum_{i=0}^{n}\binom{n}{i}a^{n-i}b^ix^i$$

Example 2.16. If X is a uniform random variable on the interval (2,7), then what is the mean of X?



Answer: The density function of X is

$$f(x) = \begin{cases} \frac{1}{5} & \text{if } 2 < x < 7 \\ 0 & \text{otherwise.} \end{cases}$$

Thus the mean or the expected value of X is

$$\mu_x = E(x) = \int_{-\infty}^{\infty} f(x)\,dx$$

$$= \int_2^7 \frac{x}{5}\,dx = \frac{x}{10}\Big|_2^7 = \frac{9}{2}$$

In general, if X ~ UNIF (a, b), then $E(X) = \dfrac{a+b}{2}$

Example 2.17. If the probability density of X is given by :

$$f(x) = \begin{cases} \dfrac{4}{\pi\left(1+x^2\right)}, & 0 < x < 1 \\ 0 & o.w \end{cases}$$

1- Find the expected value of X

2- Find E(y) where $y=(3+2x)^2$

Solution:

$$E(x) = \int x . f(x) dx$$

$$= \int_0^1 \frac{4x}{\pi\left(1+x^2\right)} dx = \frac{2}{\pi}\int_0^1 \frac{2x}{\left(1+x^2\right)} dx$$

$$= \frac{2}{\pi}\ln\left(1+x^2\right)|_0^1 = \frac{2}{\pi}\left(\ln(2) - \ln(1)\right)$$

$$= \frac{2}{\pi}\ln(2)$$

2- $y=(3+2x)^2$

$$E(y)=E(3+2x)^2=\sum_{i=0}^{2}\binom{2}{i}3^{2-i}\ 2^i E(x^i)$$

$$=\binom{2}{0}3^2\ 2^0E(x^0)+\binom{2}{1}3^1\ 2^1E(x^1)+\binom{2}{2}3^0\ 2^2E(x^2)$$

$$E(x)=\frac{2}{\pi}\ln 2$$

$$E(x^2)=\int x^2.f(x)dx$$

$$=\int_0^1\frac{4x^2}{\pi\left(1+x^2\right)}dx=\frac{4}{\pi}\int_0^1\frac{1+x^2-1}{\left(1+x^2\right)}dx=\frac{4}{\pi}\int_0^1 1-\frac{1}{1+x^2}dx$$

$$=\frac{4}{\pi}\int_0^1 dx-\frac{4}{\pi}\int_0^1\frac{1}{1+x^2}dx=\frac{4}{\pi}\left(x\ |_0^1-\tan^{-1}x\ |_0^1\right)$$

$$=\frac{4}{\pi}\left(1-\tan^{-1}1\right)=\frac{4}{\pi}\left(1-\frac{\pi}{4}\right)$$

$$=\frac{4}{\pi}-1$$

$$E(y)=\binom{2}{0}3^2\ 2^0E(x^0)+\binom{2}{1}3^1\ 2^1E(x^1)+\binom{2}{2}3^0\ 2^2E(x^2)$$

$$=\binom{2}{0}3^2\ 2^0+\binom{2}{1}3^1\ 2^1\frac{\ln 4}{\pi}+\binom{2}{2}3^0\ 2^2(\frac{4}{\pi}-1)$$

Example 2.18. If the probability mass function of X is given by :

$$P(x) = \frac{\binom{2}{x}\binom{10}{3-x}}{\binom{12}{3}}, \quad x = 0,1,2$$

1- Find the expected value of this random variable X?

2- Find E(y) where $y=(3+2x)^3$.

Solution:

$$P(0) = \frac{\binom{2}{0}\binom{10}{3}}{\binom{12}{3}} = \frac{\frac{10!}{3!7!}}{\frac{12!}{3!9!}} = \frac{6}{11}$$

$$P(1) = \frac{\binom{2}{1}\binom{10}{2}}{\binom{12}{3}} = \frac{\frac{10!}{2!8!}}{\frac{12!}{3!9!}} = \frac{9}{22}$$

$$P(2) = \frac{\binom{2}{2}\binom{10}{1}}{\binom{12}{3}} = \frac{10}{\frac{12!}{3!9!}} = \frac{1}{22}$$

$$E(x) = \sum_{x=0}^{2} x.P(x)$$

$$= 0.P(0) + P(1) + 2.P(2)$$

$$= \frac{9}{22} + 2.\frac{1}{22} = \frac{11}{22} = \frac{1}{2}$$

2-E(y) where $y=(3+2x)^3$.

$$E(y)=E(3+2x)^3=\sum_{i=0}^{3}\binom{3}{i}\,3^{3-i}\;2^i\,E(x^i)$$
$$=\binom{3}{0}3^3\;2^0E(x^0)+\binom{3}{1}3^2\;2^1E(x^1)+\binom{3}{2}3^1\;2^2E(x^2)+\binom{3}{3}3^0\;2^3E(x^3)$$

$$E(x^0)=1,\quad E(x)=\frac{1}{2},$$

$$E(x^2)=\sum_{x=0}^{2}x^2.P(x)$$
$$=0.P(0)+P(1)+4.P(2)$$
$$=\frac{9}{22}+4.\frac{1}{22}=\frac{13}{22}.$$

$$E(x^3)=\sum_{x=0}^{2}x^3.P(x)$$
$$=0.P(0)+P(1)+8.P(2)$$
$$=\frac{9}{22}+8.\frac{1}{22}=\frac{17}{22}.$$

$$E(y)=E(3+2x)^3=$$
$$=\binom{3}{0}3^3\;2^0E(x^0)+\binom{3}{1}3^2\;2^1E(x^1)+\binom{3}{2}3^1\;2^2E(x^2)+\binom{3}{3}3^0\;2^3E(x^3)$$

$$=3^3+3^3\;2E(x)+3^2\;2^2E(x^2)+2^3E(x^3)$$

$$=3^3+3^3\,2\left(\frac{1}{2}\right)+3^2\;2^2\left(\frac{13}{22}\right)+2^3\left(\frac{17}{22}\right)$$

2.6 Variance of Random Variables

The spread of the distribution of a random variable X is its variance.

Definition 4.4. Let X be a random variable with mean μ_x . The variance of X, denoted by V (X), is defined as

$$V(X) = E\,[X - \mu_x\,]^2.$$

It is also denoted by σ_x^2 . The positive square root of the variance is called the standard deviation of the random variable X. Like variance, the standard deviation also measures the spread. The following theorem tells us how to compute the variance in an alternative way.

Theorem 2.6.1. If X is a random variable with mean μ_x and variance σ_x^2 then

$$\sigma_x^2 = E\,(x^2) - \mu_x^2$$

Proof:

$$\begin{aligned} V(x) &= E\,[\,X - \mu x\,]^2 \\ &= E\,(\,X^2) - 2\,\mu_X X\,E(X) + \mu_X^2 \\ &= E(X^2) - 2\,\mu x\,E(X) + (\,\mu x\,)^2 \\ &= E(X^2) - 2\,\mu x\,\mu x + (\,\mu x\,) \\ &= E(X^2) - (\,\mu x\,)^2. \end{aligned}$$

Theorem 2.6.2.

If X x is a random variable with mean μ_x and variance σ_x^2 , then

$$V\,(a\,X + b) = a^2\,V\,(X),$$

where a and b are arbitrary real constants.

Proof:

$$
\begin{aligned}
V(aX + b) &= E[(aX + b) - \mu_{aX+b}]^2 \\
&= E[aX + b - E(aX + b)]^2 \\
&= E[aX + b - a\mu_{X+} - b]^2 \\
&= E\, a^2 [X - \mu_X]^2 \\
&= a^2 E[X - \mu_X]^2 \\
&= a^2 V(X).
\end{aligned}
$$

Example 2.19. Let X have the density function

$$
f(x) = \begin{cases} \frac{2x}{k^2} & \text{for } 0 \leq x \leq k \\ 0 & \text{otherwise.} \end{cases}
$$

For what value of k is the variance of X equal to 2? **Answer:** The expected value of X is

$$
\begin{aligned}
E(X) &= \int_0^k x\, f(x)\, dx \\
&= \int_0^k x\, \frac{2x}{k^2}\, dx \\
&= \frac{2}{3} k.
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \int_0^k x^2\, f(x)\, dx \\
&= \int_0^k x^2\, \frac{2x}{k^2}\, dx \\
&= \frac{2}{4} k^2.
\end{aligned}
$$

Hence, the variance is given by

$$V(X) = E(X^2) - (\mu_X)^2$$
$$= (2/4)\, k^2 - (4/9)\, k^2$$

Since this variance is given to be 2, we get

$$\frac{1}{18} k^2 = 2$$

and this implies that $k = \pm 6$. But k is given to be greater than 0, hence k must be equal to 6.

Example 2.20. If the probability density function of the random variable is

$$f(x) = \begin{cases} 1 - |x| & \text{for } |x| < 1 \\ 0 & \text{otherwise,} \end{cases}$$

then what is the variance of X?

Answer:

Since $V(X) = E(X^2) - \mu_x^2$, we need to find the first and second moments of X. The first moment of X is given by

$$\mu_x = E(X)$$

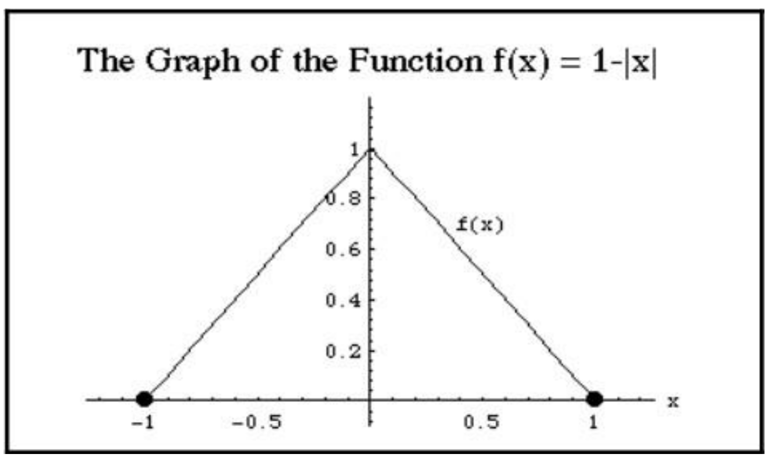
$$\begin{aligned}
&= \int_{-\infty}^{\infty} x\, f(x)\, dx \\
&= \int_{-1}^{1} x\, (1 - |x|)\, dx \\
&= \int_{-1}^{0} x\, (1 + x)\, dx + \int_{0}^{1} x\, (1 - x)\, dx \\
&= \int_{-1}^{0} (x + x^2)\, dx + \int_{0}^{1} (x - x^2)\, dx \\
&= \frac{1}{3} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} \\
&= 0.
\end{aligned}$$

The second moment $E(X^2)$ of X is given by

$$
\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2\, f(x)\, dx \\
&= \int_{-1}^{1} x^2\,(1-|x|)\, dx \\
&= \int_{-1}^{0} x^2\,(1+x)\, dx + \int_{0}^{1} x^2\,(1-x)\, dx \\
&= \int_{-1}^{0} (x^2+x^3)\, dx + \int_{0}^{1} (x^2-x^3)\, dx \\
&= \frac{1}{3} - \frac{1}{4} + \frac{1}{3} - \frac{1}{4} \\
&= \frac{1}{6}.
\end{aligned}
$$

Thus, the variance of X is given by

$$V(x) = E(x^2) - \mu_x^2 = \frac{1}{6} - 0 = \frac{1}{6}$$



Properties of the Variance

Let X be random variable and a and b two real numbers then

$$V(ax+b)=a^2V(x)$$

Proof:

$$V(ax+b)=E(ax+b-E(ax+b))^2=E(ax+b-E(ax)-b)^2$$
$$=E(ax-aE(x))^2=a^2E(x-E(x))^2$$
$$=a^2V(x)$$

Theorem 2.6.3:

If X is random variable has mean μ and variance σ^2 ,then

$$y=\frac{x-\mu}{\sigma},\ E(y)=0, V(y)=1$$

Proof:

$$E(y)=E(\frac{x-\mu}{\sigma})=\frac{1}{\sigma}E(x-\mu)=\frac{1}{\sigma}\left(E(x)-\mu\right)=0$$

$$V(y)=V(\frac{x-\mu}{\sigma})=\frac{1}{\sigma^2}V(x-\mu)=\frac{1}{\sigma^2}V(x)=\frac{\sigma^2}{\sigma^2}=1$$

Standard deviation:

$$\sigma=\sqrt{V(x)}$$

Example 2.21 :

If the probability mass function of X is given by :

$$P(x)=\frac{1}{6}, \qquad x=1,2,\ 3,\ 4,\ 5,\ 6$$

1- Find the variance of X and the standard deviation of X?
2- Find the variance of Y where $Y=5x-3$

Solution:

$$E(x)=\sum_{x=1}^{6}x\,P(x)=\frac{1}{6}(1+2+3+4+5+6)=\frac{21}{6}=\frac{7}{2}$$

$$E(x^2)=\sum_{x=1}^{6}x^2\,P(x)=\frac{1}{6}(1+4+9+16+25+36)=\frac{91}{6}$$

$$V(x)=E(x^2)-\left(E(x)\right)^2=\frac{91}{6}-\frac{49}{4}=\frac{35}{12}$$

$$\sigma=\sqrt{V(x)}=\sqrt{\frac{35}{12}}$$

$$V(y)=V(5x-3)=25V(x)=25.\left(\frac{35}{12}\right)$$

Example 2.22 :

If the probability density function of X is given by :

$$f(x)=\frac{8}{3}x^3e^{-2x}, \qquad x\geq 0$$

1- Find the variance of X and the standard deviation of X?
2- Find the variance of Y where $Y=5x-3$

Solution:

$$\Gamma(n) = (n-1)! = \int_0^\infty x^{n-1} e^{-x}\,dx$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$E(x) = \int_0^\infty x\, f(x)dx = \frac{8}{3}\int_0^\infty x^4 e^{-2x}dx = \frac{8}{3}\int_0^\infty \frac{y^4}{16} e^{-y}\,\frac{dy}{2} = \frac{1}{12}\Gamma(5)$$

$$= \frac{1}{12}4! = \frac{24}{12} = 2$$

$$E(x^2) = \int_0^\infty x^2 f(x)dx = \frac{8}{3}\int_0^\infty x^5 e^{-2x}dx = \frac{8}{3}\int_0^\infty \frac{y^5}{32} e^{-y}\,\frac{dy}{2} = \frac{1}{24}\Gamma(6)$$

$$= \frac{1}{24}5! = \frac{120}{24} = 5$$

$$V(x) = E(x^2) - \left(E(x)\right)^2 = 5 - 4 = 1$$

$$\sigma = \sqrt{V(x)} = \sqrt{1} = 1$$

$$V(y) = V(5x-3) = 25V(x) = 25.\left(1\right) = 25$$

Exercise 2

1. A discrete random variable X has a probability mass function given by:
   $$f(x) = c(x+1), \quad x = 0, 1, 2, 3.$$

(a) Find the value of the constant c.
(b) Draw (i) a probability graph,
(ii) a probability histogram, to represent $f(x)$.
(c) Find: (i) $P(0 \leq X < 2)$, (ii) $P(X > 1)$.

2. Determine whether each of the following functions can serve as a probability mass function of a discrete random variable:

(a) $f(x) = (x-1), \quad x = 0, 1, 2, 3.$
(b) $g(x) = x, \quad x = 1, 2, 3, 4.$
(c) $h(x) = x, \quad x = -1, 0, 1, 2.$

3. Let X be a random variable whose probability mass function is defined by the values

   $f(-2) = \frac{1}{10}$, $f(0) = \frac{2}{10}$, $f(4) = \frac{4}{10}$, $f(11) = \frac{3}{10}$.

   Find:
   (a) $P(-2 \leq X < 4)$, (b) $P(X > 0)$, (c) $P(X \leq 4)$.

4. A Check whether the following functions satisfy the conditions of a probability mass function.

(a) $f(x) = \frac{1}{4}, \quad x = -3, 0, 1, 4.$
(b) $f(x) = \frac{1}{x}, \quad x = 1, 2, 3, 4.$
(c) $f(x) = 1 - x, \quad x = 0, \frac{1}{2}, \frac{3}{2}.$

(d)) $f(x) = (\frac{1}{2})^x, \quad x = 1, 2, 3, 4, \ldots$.

5. Consider a throw of two fair dice. Let X denote the sum of the numbers on the two dice.

   (a) Find the probability mass function of X.

   (b) Find: (i) $P(X = 7)$, (ii) $P(X > 8)$, (iii) $P(3 < X < 11)$.

6. The sample space of an experiment is {a, b, c, d, e, f }, and each outcome is equally likely. A random variable X, is defined as follows:

outcome	a	b	c	d	e	f
X	0	0	1.5	1.5	2	3

   (a) Determine the probability distribution of X.
   (b) Find: (i) $P(X = 1.5)$, (ii) $P(0.5 < X < 2.7)$, (iii) $P(X > 3)$, (iv) $P(0 \leq X < 2)$.

7. Determine the value of c so that each of the following can serve as a probability mass function of a discrete random variable.

   (a) $f(x) = c(x^2 + 4), \qquad x = 0, 1, 2, 3.$

   (b) $f(x) = c\binom{2}{x}\binom{3}{3-x}, \qquad x = 0, 1, 2.$

8. A discrete random variable X has the probability mass function given by

$$f(x) = \begin{cases} a\left(\frac{1}{3}\right)^{x-1}, & x = 1, 2, 3, \ldots \\ 0, & \text{elsewhere.} \end{cases}$$

   (a) Find the value of a. (b) Find $P(X = 3)$.

9. Show that the following functions are probability density functions for some value of c and determine c.

(a) $$f(x) = \begin{cases} c\,e^{-4x}, & x \geq 0 \\ 0, & \text{elsewhere.} \end{cases}$$

(b) $$f(x) = \begin{cases} c\,x^{2}, & -1 \leq x \leq 10 \\ 0, & \text{elsewhere.} \end{cases}$$

(c) $$f(x) = \begin{cases} c\,(1+2x), & 0 < x < 2 \\ 0, & \text{elsewhere.} \end{cases}$$

(d) $$f(x) = \begin{cases} \frac{1}{2}e^{-cx}, & x \geq 0 \\ 0, & \text{elsewhere.} \end{cases}$$

10. Suppose that in a certain region, the daily rainfall (in inches) is a continuous random variable X with p.d.f. f (x) given by

$$f(x) = \begin{cases} \frac{3}{4}(2x - x^{2}), & 0 < x < 2 \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability that on a given day in this region the rainfall is

(a) not more than 1 inch, (b) greater than 1.5 inches,
(c) equal to 1 inch, (d) less than 1 inch.

11. Let X be a continuous random variable with p.d.f.

$$
f(x) = \begin{cases} -4x, & -0.5 < x < 0 \\ 4x, & 0 < x < 0.5 \\ 0, & \text{elsewhere.} \end{cases}
$$

(a) Sketch the graph of f (x).
(b) Find:
(i) $P(X \leq -0.3)$, (ii) $P(X \leq 0.3)$, (iii) $P(-0.2 \leq X \leq 0.2)$.

12. The pressure (measured in kg/cm2) at a certain valve is a random variable X whose p.d.f. is

$$
f(x) = \begin{cases} \frac{6}{27}\left(3x - x^2\right), & 0 < x < 3 \\ 0, & \text{elsewhere.} \end{cases}
$$

Find the probability that the pressure at this valve is
(a) less than 2 kg/cm2, (b) greater than 2 kg/cm2, (c) between 1.5 and 2.5 kg/cm2.

13. Let X denote the length in minutes of a long-distance telephone conversation. Assume that the p.d.f. of X is given by

$$
f(x) = \begin{cases} \frac{1}{10} e^{-\frac{x}{10}}, & x > 0 \\ 0, & \text{elsewhere.} \end{cases}
$$

(a) Verify that f is a p.d.f. of a continuous random variable.
(b) Find the probability that a randomly selected call will last:
(i) at most 7 minutes, (ii) at least 7 minutes, (iii) exactly 7 minutes.

14. A continuous random variable X has the p.d.f.

$$f(x) = \begin{cases} \frac{2}{27}(1+x), & 2 < x < 5 \\ 0, & \text{elsewhere.} \end{cases}$$

Find: (a) $P(X < 4)$, (b) $P(3 < X < 4)$.

15. The proportion of people who respond to a certain mail-order solicitation is a continuous random variable X with p.d.f.

$$f(x) = \begin{cases} \frac{2}{27}(1+x), & 2 < x < 5 \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find $P(0 < X < 0.8)$.
(b) Find the probability that more than 0.25 but fewer than 0.5 of the people contacted will respond to this type of solicitation.

16. A continuous random variable X that can assume values between x = 1 and x = 3 has a p.d.f. given by

$$f(x) = \frac{1}{2}$$

(a) Show that the area under the curve is equal to 1.

(b) Find: (i) $P(2 < X < 2.5)$, (ii) $P(X \leq 1.6)$.

17. Which of the following functions are probability density functions?

(a) $$f(x) = \begin{cases} x, & -0.5 < x < 0.5 \\ 0, & \text{elsewhere.} \end{cases}$$

(b) $g(x) = \begin{cases} \dfrac{1}{2}, & -1 < x < 1 \\ 0, & \text{elsewhere.} \end{cases}$

(c) $f(x) = \begin{cases} \dfrac{1}{x}, & 0 < x < 1 \\ 0, & \text{elsewhere.} \end{cases}$

(d) $h(x) = \begin{cases} \dfrac{1}{3}, & 0 < x < 1 \\ \dfrac{2}{3}, & 2 < x < 3 \end{cases}$

18. The random variable X has the p.d.f.

$$f(x) = \begin{cases} \dfrac{1}{30}, & 0 < x < 30 \\ 0, & \text{elsewhere.} \end{cases}$$

Find:

(a) $P(X > 25 \mid X > 15)$, (b) $P(X < 20 \mid X > 15)$,

(c) $P(X > 15 \mid X < 22)$, (d) $P(X < 13 \mid X < 18)$.

19. Let X be a discrete random variable whose only possible values are 1, 2 and 5. Find the cumulative distribution function of X if the probability mass function of X is defined by the following values:

$f(1) = \dfrac{1}{4}, f(2) = \dfrac{1}{2}, f(5) = \dfrac{1}{4}.$

20. Let X be a discrete random variable whose cumulative distribution function is

$$F(x) = \begin{cases} 0, & x < -3 \\ \frac{1}{6}, & -3 \le x < 6 \\ \frac{1}{2} & 6 \le x < 10 \\ 1, & x \ge 10. \end{cases}$$

(a) Find:

(i) $P(X \le 4)$, (ii) $P(-5 < X \le 4)$, (iii) $P(X = 4)$.

(b) Find the probability mass function of X.

21. Let X be a discrete random variable with cumulative distribution function

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.1, & 1 \le x < 3 \\ 0.4 & 3 \le x < 5 \\ 0.9 & 5 \le x < 5.5 \\ 1, & x \ge 5.5. \end{cases}$$

(a) Find :

(i) $P(X \le 3)$, (ii) $P(X \le 4)$, (iii) $P(1.5 < X \le 5.2)$.

(b) Find the probability mass function of X.

22. Let X be a discrete random variable whose only possible values are −5, −1, 0, and 7.

Find the cumulative distribution function of X if the probability mass function of X is defined by the values :

$f(-5) = 0.3$, $f(-1) = 0.1$, $f(0) = 0.2$, and $f(7) = 0.4$.

23. Let

$$F(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

be the cumulative distribution function of the random variable X. Find:

(a) $P(X \leq -1)$, (b) $P(X \leq 0.5)$, (c) $P(X > 0.4)$,
(d) $P(0.2 < X \leq 0.5)$, (e) $P(X > 0.4 \mid X > 0.2)$.

24. The weekly profit (in thousands) from a certain concession is a random variable X whose distribution function is given by:

$$F(x) = \begin{cases} 0, & x < 0 \\ 3x\text{-}3x^2 + x^3, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

(a) Find the probability of a weekly profit of less than 2 000.00.
(b) Find the probability of a weekly profit of at least 500.00.

25. Suppose the cumulative distribution function of the random variable X is

$$F(x) = \begin{cases} 0, & x < -2 \\ 0.25x+0.5, & -2 \le x < 2 \\ 1, & x \ge 2 \end{cases}$$

(a) Find the p.d.f. of X.

(b) Calculate:
(i) $P(X < 1.8)$, (ii) $P(X > -1.5)$, (iii) $P(X < -2)$, (iv) $P(X > 1 \mid X > 0.5)$.

26. A continuous random variable X has the p.d.f.

$$\mathbf{f(x)} = \begin{cases} e^{-cx}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where c is a constant.

(a) Find the value of c.
(b) Find the cumulative distribution function of X and sketch its graph.
(c) Find (i) $P(X > 8 \mid X > 3)$, (ii) P(X >1 X < 4).

27. The distribution function of X is given by:

$$\mathbf{F(x)} = \begin{cases} 0, & x < 0 \\ \frac{x}{8}, & 0 \leq x < 2 \\ \frac{x^2}{16}, & 2 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

(a) Find the p.d.f. of X.
(b) Find (i) $P(1 \leq X \leq 3)$, (ii) $P(X < 3)$, (iii) $P(X > 1 \mid X > 2)$.

28. The shelf life, in days, for bottles of a certain prescribed medicine is a random variable with p.d.f.

$$\mathbf{f(x)} = \begin{cases} \frac{20000}{(x+100)^3}, & x \geq 0 \\ 0, & elsewhere \end{cases}$$

Find the probability that a bottle of this medicine will have a shelf life of
(a) at least 200 days, (b) between 80 and 120 days.

29. A random variable X has distribution function given by

$$F(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{2}(x+1)^2, & -1 \le x < 0 \\ 1 - \frac{1}{2}(1-x)^2, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

Find the p.d.f. of X.

30. A random variable X has the p.d.f. given by

$$f(x) = \begin{cases} \frac{1}{2}x, & 0 \le x \le 2 \\ 0, & elsewhere \end{cases}$$

(a) Find the cumulative distribution function of X and sketch its graph.
(b) Compute (i) $P(X \le 2)$, (ii) $P(X \le 2 \mid 1 \le X < 3)$.

31. A discrete random variable X has probability mass function of the form

$$p(x) = \begin{cases} c(8-x), & for\ x = 0,1,2,3,4,5. \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the constant c. (b) Find P (X > 2).
(c) Find the expected value E(X) for the random variable X.

32. A random variable X has a cumulative distribution function

$$F(x) = \begin{cases} \frac{1}{2}x, & 0 \le x < 1 \\ x - \frac{1}{2}, & 1 \le x < \frac{3}{2} \end{cases}$$

(a) Graph F (x). (b) Graph f(x). (c) Find P ($X \le 0.5$).
(d) Find P ($X \ge 0.5$). (e) Find P ($X \le 1.25$). (f) Find P (X = 1.25).

33. Let X be a random variable with probability density function

$$p(x) = \begin{cases} \frac{1}{15}x, & for\ x = 1,2,3,4,5. \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the expected value of X. (b) Find the variance of X.
(c) Find the expected value of 2X (d) Find the variance of 2X + 3.
(e) Find the expected value of $3X - 5X^2 + 1$.

34. 5. The measured radius of a circle, R, has probability density function

$$f(x) = \begin{cases} 6r(1-r), & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

(a) Find the expected value of the radius.
(b) Find the variance of r.

Chapter Three

3.1 Special Probability Distribution

1- Binomial Distribution

Consider a fixed number n of mutually independent Bernoulli trails. Suppose these trials have same probability of success, say p. A random variable X is called a binomial random variable if it represents the total number of successes in n independent Bernoulli trials.
Now we determine the probability mass function of a binomial random variable. Recall that the probability mass function of X is defined as

$p(x) = P(X = x)$.

Thus, to find the probability mass function of X we have to find the probability of x successes in n independent trails.

If we have x successes in n trails, then the probability of each n-tuple with x successes and n - x failures is

$$p^x (1 - p)^{n-x}.$$

However, there are $\binom{n}{x}$ tuples with x successes and n- x failures in n trials.
Hence

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}.$$

Therefore, the probability density function of X is

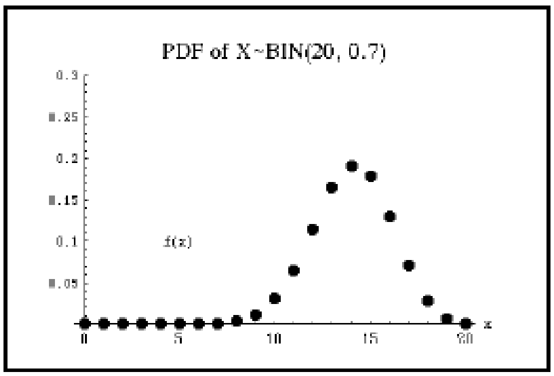
$$P(x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

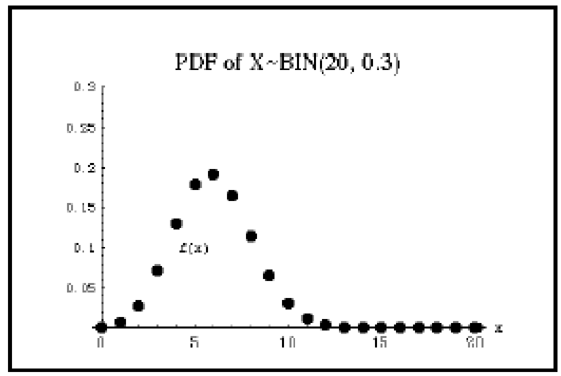
Definition 5.2. The random variable X is called the binomial random variable with parameters p and n if its probability mass function is of the form

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

where $0 < p < 1$ is the probability of success.

We will denote a binomial random variable with parameters p and n as $X \sim BIN(n, p)$.





Example 3.1. Is the real valued function f(x) given by:

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

where n and p are parameters, a probability mass function?

Answer: To answer this question, we have to check that p(x) is nonnegative And $\sum_{x=0}^{n} p(x)$ is 1 . It is easy to see that p(x) ≥ 0.

We show that sum is one.

$$\sum_{x=0}^{n} p(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$
$$= (p+1-p)^n = 1$$

Hence p(x) is really a probability mass function.

Example 3.2.

On a five-question multiple-choice test there are five possible answers, of which one is correct. If a student guesses randomly and independently, what is the probability that she is correct only on two questions?

Answer:

Here the probability of success is p = 1/5, and thus

1 −p = 4/5 .

There different ways she can be correct on two questions.

Therefore, the probability that she is correct on two questions is

$$P(\text{correct on two questions}) = \binom{5}{2} p^2 (1-p)^3$$

$$= 10 \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^3$$

$$= \frac{640}{5^5} = 0.2048.$$

Example 3.3.

What is the probability of rolling two sixes and three nonsixes in 5 independent casts of a fair dice?

Answer:
Let the random variable X denote the number of sixes in 5 in- dependent casts of a fair die. Then X is a binomial random variable with probability of success p and n= 5. The probability of getting a six is p = 1/6.
Hence

$$p(x=2) = p(2) = \binom{5}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3$$

$$= 0.160751.$$

Example 3.4.

What is the probability of rolling at most two "three" in 5 independent casts of a fair dice?

Answer:

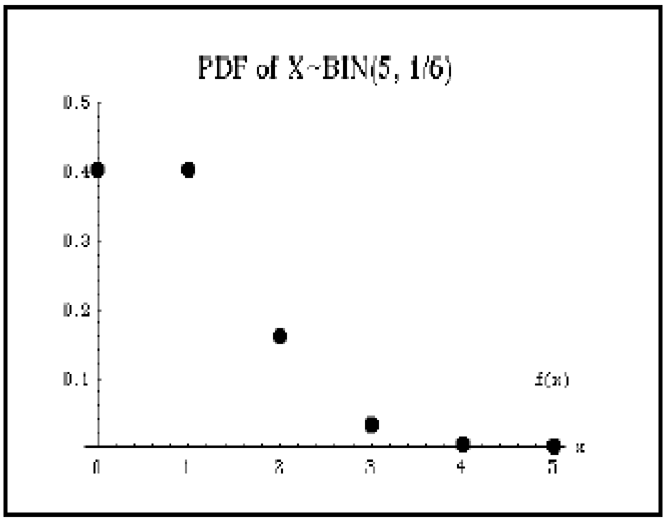
Let the random variable X denote number of three in 5 independent casts of a fair dice. Then X is a binomial random variable with probability of success p and n = 5. The probability of getting a three is p = 1/6 . Hence, the probability of rolling at most two three is

P (X ≤ 2) = F (2) = P(0) +P(1) + P(2)

$$= \binom{5}{0}\left(\frac{1}{6}\right)^{0}\left(\frac{5}{6}\right)^{5} + \binom{5}{1}\left(\frac{1}{6}\right)^{1}\left(\frac{5}{6}\right)^{4} + \binom{5}{2}\left(\frac{1}{6}\right)^{2}\left(\frac{5}{6}\right)^{3}$$

$$= \sum_{k=0}^{2}\binom{5}{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{5-k}$$

$$= \frac{1}{2}(0.9421 + 0.9734) = 0.9577 \qquad \text{(from binomial table)}$$



Theorem 3.1.1.

If X is a binomial random variable with parameters p and n, then the mean, variance and moment generating functions are respectively given by

$$\mu_X = n\,p$$

$$\sigma_X^2 = n\,p\,(1 - p)$$

$$M_X(t) = [(1-p)+pe^t]^n .$$

Proof: First, we determine the moment generating function M(t) of X and then we generate mean and variance from M(t).

$$\begin{aligned} M(t) &= E\left(e^{tX}\right) \\ &= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} \left(p\,e^t\right)^x (1-p)^{n-x} \\ &= \left(p\,e^t + 1 - p\right)^n . \end{aligned}$$

$$E(x) = \sum x . P(x)$$

$$= \sum_{x=0}^{n} x . \binom{n}{x} p^x q^{n-x}, \quad where \quad q = 1 - p$$

$$= p \sum_{x=0}^{n} x . \frac{n!}{x!(n-x)!} p^{x-1} q^{n-x}$$

$$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

$$where, \quad y = x - 1, \quad m = (n-1).$$

$$= np \sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^y q^{m-y} = np$$

$$E(x^2) = \sum x^2.P(x)$$

$$= \sum_{x=0}^{n} x^2.\binom{n}{x} p^x q^{n-x}, \text{ where } q = 1 - p$$

$$= \sum_{x=0}^{n} x^2.\frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$= np\sum_{x=1}^{n} x.\frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

$$= np\sum_{x=1}^{n} (x-1+1).\frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

$$= np + np\sum_{x=1}^{n} (x-1).\frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

$$= np + np^2(n-1)\sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} q^{n-x}$$

$$\text{where, } y = x - 2, \; m = (n-2).$$

$$= np + np^2(n-1)\sum_{y=0}^{m} \frac{m!}{y!(m-y)!} p^{y} q^{m-y}$$

$$= np + n(n-1)p^2$$

$$V(x) = E(x^2) - (E(x))^2$$

$$= np + n(n-1)p^2 - n^2p^2 = np - np^2$$

$$= np(1-p) = npq$$

Example 3.5:

Afair coin is tossed 12 times:

Find the probability of getting five heads and seven tails.

Solution:

$$P(x) = \binom{n}{x} p^{x} \; q^{n-x}$$

$$n = 12, \;\; x = 5$$

$$P(5) = \binom{12}{5} \left(\frac{1}{2}\right)^{5} \left(\frac{1}{2}\right)^{7}$$

2. Poisson Distribution

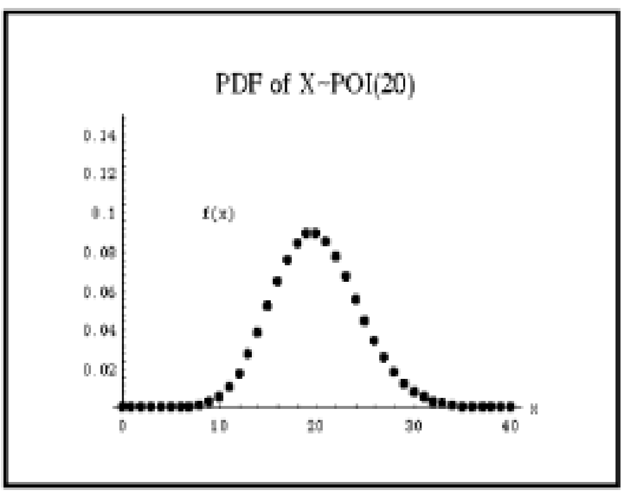
In this section, we define an important discrete distribution which is widely used for modeling many real life situations. First, we define this distribution and then we present some of its important properties.

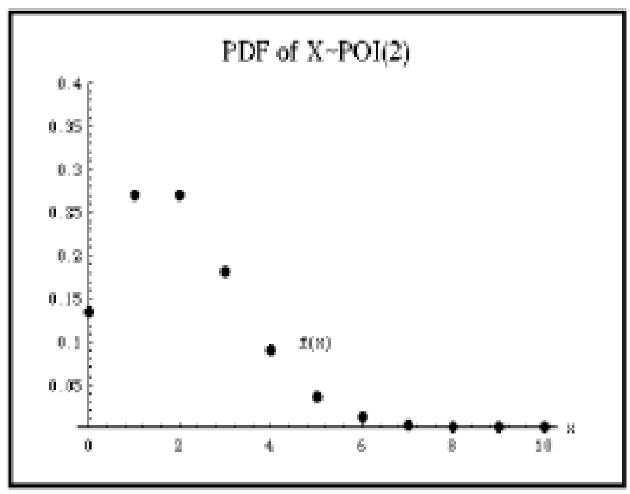
Definition 3.2.

A random variable X is said to have a Poisson distribution if its probability density function is given by

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

where $0 < \lambda < \infty$ is a parameter. We denote such a random variable by $X \sim POI(\lambda)$.





The probability density function f is called the Poisson distribution after Simeon D. Poisson (1781-1840).

Example 3.6.

Is the real valued function defined by

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

where $0 < \lambda < \infty$ is a parameter, a probability mass function?

Answer: It is easy to check $p(x) \geq 0$. We show that $\sum_{x=0}^{\infty} p(x)$ is equal to one.

$$\sum_{x=0}^{\infty} p(x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1$$

Theorem 3.7.

If X ~ POI(λ), then

$$E(X) = \lambda$$

$$V(X) = \lambda$$

$$M(t) = e^{\lambda(e^t - 1)}.$$

Proof:

$$E(x) = \sum x.P(x)$$

$$= \sum_{x=0}^{\infty} x.\frac{e^{-\lambda}\lambda^x}{x!},$$

$$= \lambda\sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} = \lambda\sum_{y=0}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{y!}, where \ y = x - 1$$

$$= \lambda$$

$$E(x^2) = \sum x^2 . P(x)$$

$$= \sum_{x=0}^{\infty} x^2 . \frac{e^{-\lambda}\lambda^x}{x!},$$

$$= \lambda \sum_{x=1}^{\infty} (x-1+1) \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$$

$$= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} + \lambda \sum_{x=1}^{\infty} (x-1) \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!},$$

$$= \lambda + \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!}$$

$$= \lambda + \lambda^2 \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^{y}}{(y)!}, \quad where \;\; y = x-2$$

$$= \lambda + \lambda^2$$

$$V(x) = E(x^2) - (E(x))^2$$

$$V(x) = \lambda + \lambda^2 - \lambda^2$$
$$= \lambda$$

We find the moment generating function of X.

$$\begin{aligned}
M(t) &= E(e^{tx}) \\
&= \sum_{x=0}^{\infty} e^{tx}\, \frac{e^{-\lambda}\,\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} e^{tx}\, \frac{\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t\lambda)^x}{x!} \\
&= e^{-\lambda}\, e^{\lambda e^t} \\
&= e^{\lambda\,(e^t-1)}.
\end{aligned}$$

Thus,

$$M'(t) = \lambda\, e^t\, e^{\lambda\,(e^t-1)},$$

and

$$E(X) = M'(0) = \lambda.$$

Similarly,

$$M''(t) = \lambda\, e^t\, e^{\lambda\,(e^t-1)} + (\lambda\, e^t)^2\; e^{\lambda\,(e^t-1)}.$$

Hence

$$M''(0) = E(X^2) = \lambda^2 + \lambda.$$

Therefore

$$V(X) = E(X^2) - (\,E(X)\,)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

This completes the proof.

Example 3.8.

A random variable X has a Poisson distribution with a mean of 3. What is the probability that X is bounded by 1 and 3, that is, P $(1 \le X \le 3)$?

Answer:

$$E(x) = 3 = \lambda$$

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Hence

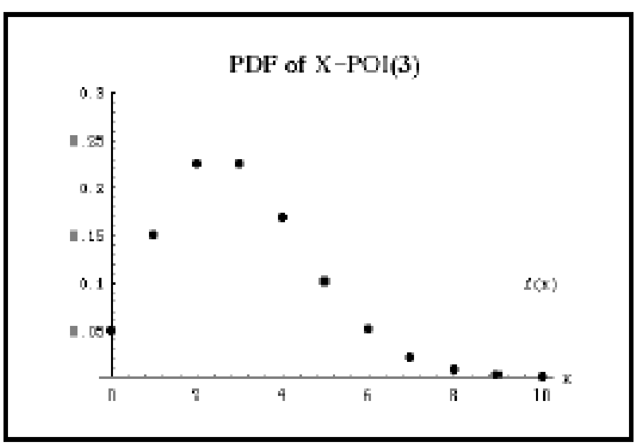
$$p(x) = \frac{3^x e^{-3}}{x!}, \quad x = 0, 1, 2, \ldots$$

Therefore

$$P\,(1 \le X \le 3) = p(1) + p(2) + p(3)$$

$$= 3\,e^{-3} + \frac{9}{2}\,e^{-3} + \frac{27}{6}\,e^{-3}$$

$$= 12\,e^{-3}.$$



Example 3.9.

The number of traffic accidents per week in a small city has a Poisson distribution with mean equal to 3. What is the probability of exactly 2 accidents occur in 2 weeks?

Answer:

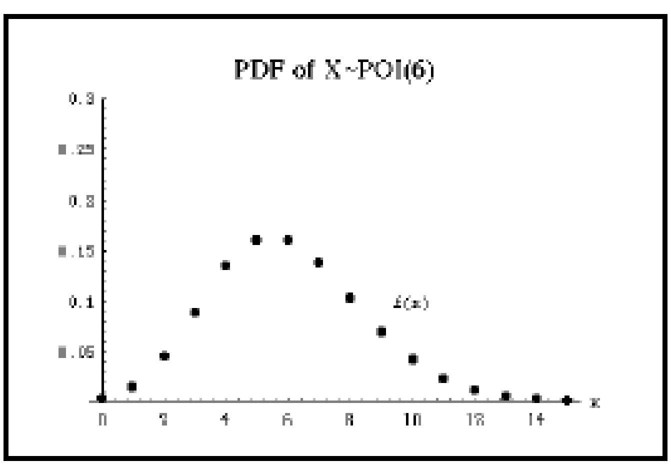
The mean traffic accident is 3. Thus, the mean accidents in two weeks are $\lambda = (3)\,(2) = 6$.

Since

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

We get

$$p(2) = \frac{6^2 e^{-6}}{2!} = 18\, e^{-6}$$



Example 3.10.

Let X have a Poisson distribution with parameter $\lambda = 1$. What is the probability that $X \geq 2$ given that $X \leq 4$?

Answer:

$$P(X \geq 2 \mid X \leq 4) = \frac{P(2 \leq X \leq 4)}{P(X \leq 4)}.$$

$$P(2 \leq X \leq 4) = \sum_{x=2}^{4} \frac{\lambda^x e^{-\lambda}}{x!}$$

$$= \frac{1}{e} \sum_{x=2}^{4} \frac{1}{x!}$$

$$= \frac{17}{24\,e}.$$

And

$$P(X \leq 4) = \frac{1}{e} \sum_{x=0}^{4} \frac{1}{x!}$$

$$= \frac{65}{24\,e}.$$

Therefore, we have

$$P(X \geq 2 \mid X \leq 4) = 17/65.$$

Example 3.11.

If the moment generating function of a random variable X is $M(t) = e^{4.6\,(e^t-1)}$, then what are the mean and variance of X? What is the probability that X is between 3 and 6, that is P $(3 < X < 6)$?

Answer: Since the moment generating function of X is given by

$$M(t) = e^{4.6\,(e^t-1)}$$

we conclude that $X \sim POI(\lambda)$ with $\lambda = 4.6$. Thus, by Theorem 5.8, we get

$$E(X) = 4.6 = V(X).$$

$$\begin{aligned} P\,(3 < X < 6) &= p(4) + p(5) \\ &= F\,(5) - F\,(3) \\ &= 0.686 - 0.326 \\ &= 0.36. \end{aligned}$$

3.2 SOME SPECIAL CONTINUOUS DISTRIBUTIONS

In this chapter, we study some well-known continuous probability density functions. We want to study them because they arise in many applications. We begin with the simplest probability density function.

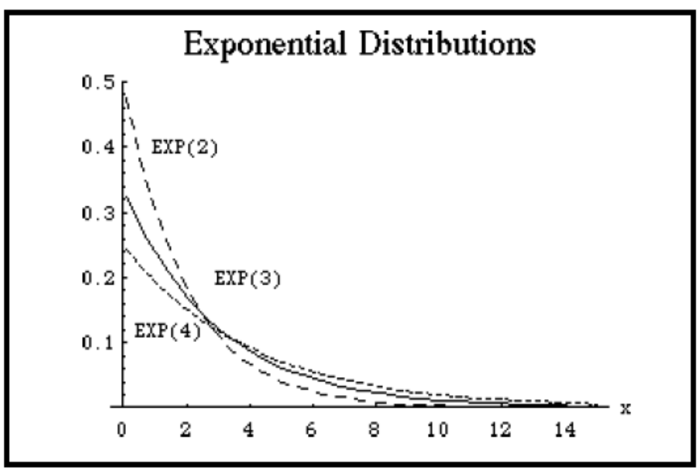
1- An Exponential Distribution

Definition 3.2.1.

A continuous random variable is said to be an exponential random variable with parameter θ if its probability density function is of the form

$$f(x) = \begin{cases} \frac{1}{\theta}\, e^{-\frac{x}{\theta}} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$. If a random variable X has an exponential density function with parameter θ, then we denote it by writing $X \sim EXP(\theta)$.



An exponential distribution is a special case of the gamma distribution. If the parameter $\alpha = 1$, then the gamma distribution reduces to the exponential distribution. Hence most of the information about an exponential distribution can be obtained from the gamma distribution.

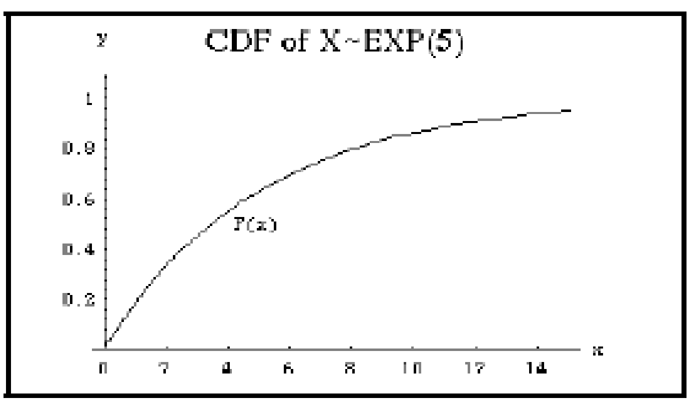
Example 3.12.

What is the cumulative density function of a random variable which has an exponential distribution with variance 25?

Answer:

Since an exponential distribution is a special case of the gamma distribution with $\alpha = 1$, from Theorem 6.3, we get $V(X) = \theta^2$. But this is given to be 25. Thus, $\theta^2 = 25$ or $\theta = 5$. Hence, the probability density function of X is

$$\begin{aligned} F(x) &= \int_0^x f(t)\,dt \\ &= \int_0^x \frac{1}{5} e^{-\frac{t}{5}}\,dt \\ &= \frac{1}{5}\left[-5\,e^{-\frac{t}{5}}\right]_0^x \\ &= 1 - e^{-\frac{x}{5}}. \end{aligned}$$



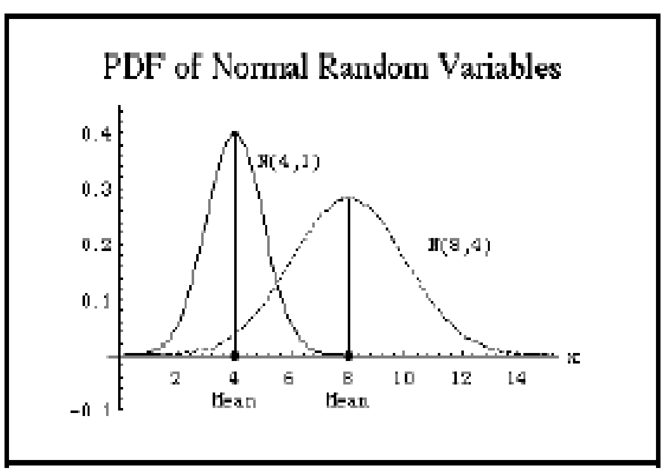
2. Normal Distribution

Among continuous probability distributions, the normal distribution is very well known since it arises in many applications. Normal distribution was discovered by a French mathematician Abraham DeMoivre (1667-1754). DeMoivre wrote two important books. One is called the *Annuities Upon Lives*, the first book on actuarial sciences and the second book is called the *Doctrine of Chances*, one of the early books on the probability theory. Pierre- Simon Laplace (1749-1827) applied normal distribution to astronomy. Carl Friedrich Gauss (1777-1855) used normal distribution in his studies of problems in physics and astronomy. Adolphe Quetelet (1796-1874) demonstrated that man's physical traits (such as height, chest expansion, weight etc.) as well as social traits follow normal distribution. The main importance of normal distribution lies on the central limit theorem which says that the sample mean has a normal distribution if the sample size is large.
Definition 6.7. A random variable X is said to have a normal distribution if its probability density function is given by

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

where $-\infty < \mu < \infty$ and $0 < \sigma^2 < \infty$ are arbitrary parameters. If X has a normal distribution with parameters μ and σ^2, then we write

$X \sim N(\mu, \sigma^2)$.



Example 3.13.

Is the real valued function defined by

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

a probability density function of some random variable X?

Answer:

To answer this question, we must check that f is nonnegative and it integrates to 1. The nonnegative part is trivial since the exponential function is always positive. Hence using property of the gamma function, we show that f integrates to 1 on R .

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

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

$$= \frac{2}{\sigma\sqrt{2\pi}} \int_{0}^{\infty} e^{-z} \frac{\sigma}{\sqrt{2z}}\,dz, \qquad \text{where} \quad z = \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

$$= \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \frac{1}{\sqrt{z}} e^{-z}\,dz$$

$$= \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = \frac{1}{\sqrt{\pi}} \sqrt{\pi} = 1.$$

The following theorem tells us that the parameter μ is the mean and the parameter σ2 is the variance of the normal distribution.

Theorem 3.2.1. If X ~ N(μ, σ2), then

$$E(X) = \mu$$
$$Var(X) = \sigma^2$$
$$M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}.$$

$$M_X(t) = \int_{-\infty}^{\infty} \frac{e^{Xt}}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} dX$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu-\sigma^2 t}{\sigma}\right)^2 + \mu t + \sigma^2 \frac{t^2}{2}} dX$$

$$= e^{\mu t + \sigma^2 \frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

$$= e^{\mu t + \sigma^2 \frac{t^2}{2}} \quad , \quad y = \frac{X-\mu-\sigma^2 t}{\sigma}.$$

$$M_X(t) = e^{\mu t + \sigma^2 \frac{t^2}{2}}$$

$$M_X(t) = e^{\mu t + \sigma^2 \frac{t^2}{2}}$$

$$M'_X(t) = \left(\mu + \sigma^2 t\right) e^{\mu t + \sigma^2 \frac{t^2}{2}},$$

$$M'_X(0) = \mu = E(x).$$

$$M''_X(t) = \left(\mu + \sigma^2 t\right)^2 e^{\mu t + \sigma^2 \frac{t^2}{2}} + \sigma^2 e^{\mu t + \sigma^2 \frac{t^2}{2}},$$

$$M''_X(0) = \mu^2 + \sigma^2,$$

$$E(x^2) = M''_X(0) = \mu^2 + \sigma^2.$$

$$V(x) = E(x^2) - (E(x))^2$$

$$= \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

Example 3.14.

If X is any random variable with mean μ and variance $\sigma^2 > 0$, then what are the mean and variance of the random variable $Y = \dfrac{X - \mu}{\sigma}$?

Answer:

The mean of the random variable Y is

$$
\begin{aligned}
E(Y) &= E\left(\frac{X-\mu}{\sigma}\right) \\
&= \frac{1}{\sigma} E(X-\mu) \\
&= \frac{1}{\sigma} (E(X)-\mu) \\
&= \frac{1}{\sigma} (\mu-\mu) \\
&= 0.
\end{aligned}
$$

The variance of Y is given by

$$
\begin{aligned}
Var(Y) &= Var\left(\frac{X-\mu}{\sigma}\right) \\
&= \frac{1}{\sigma^2} Var(X-\mu) \\
&= \frac{1}{\sigma} Var(X) \\
&= \frac{1}{\sigma^2} \sigma^2 \\
&= 1.
\end{aligned}
$$

Hence, if we define a new random variable by taking a random variable and subtracting its mean from it and then dividing the resulting by its standard deviation, then this new random variable will have zero mean and unit variance.

Definition 3.2.2.

A normal random variable is said to be standard normal, if its mean is zero and variance is one. We denote a standard normal random variable X by X ~ N(0, 1).

The probability density function of standard normal distribution is the following:

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

Example 3.15.

If $X \sim N(0, 1)$, what is the probability of the random variable X less than or equal to -1.72?

Answer:

$$P(X \leq -1.72) = 1 - P(X \leq 1.72)$$

$$= 1 - 0.9573 \qquad \text{(from table)}$$

$$= 0.0427.$$

Example 3.16.

If $Z \sim N(0, 1)$, what is the value of the constant c such that $P(|Z| \leq c) = 0.95$?

Answer:

$$0.95 = P(|Z| \leq c)$$

$$= P(-c \leq Z \leq c)$$

$$= P(Z \leq c) - P(Z \leq -c)$$

$$= 2\,P(Z \leq c) - 1.$$

Hence
$P(Z \leq c) = 0.975$,
and from this using the table we get
$c = 1.96$.

The following theorem is very important and allows us to find probabilities by using the standard normal table.

Theorem 3.2.2.

If X ~ N(μ, σ2), then the random variable

$$Z = \frac{x-\mu}{\sigma} \sim N(0, 1).$$

Proof:

We will show that Z is standard normal by finding the probability density function of Z. We compute the probability density of Z by cumulative distribution function method.

$$\begin{aligned}
F(z) &= P(Z \le z) \\
&= P\left(\frac{X-\mu}{\sigma} \le z\right) \\
&= P(X \le \sigma z + \mu) \\
&= \int_{-\infty}^{\sigma z+\mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&= \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} \sigma e^{-\frac{1}{2}w^2} dw, \qquad \text{where } w = \frac{x-\mu}{\sigma}.
\end{aligned}$$

Hence

$$f(z) = F'(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}.$$

The following example illustrates how to use standard normal table to find probability for normal random variables.

Example 3.17.

If X ~ N(3, 16), then what is P (4 ≤ X ≤ 8)?

Answer:

$$
\begin{aligned}
P(4 \le X \le 8) &= P\left(\frac{4-3}{4} \le \frac{X-3}{4} \le \frac{8-3}{4}\right) \\
&= P\left(\frac{1}{4} \le Z \le \frac{5}{4}\right) \\
&= \mathbf{P(Z \le 1.25) - P(Z \le 0.25)} \\
&= 0.8944 - 0.5987 \\
&= 0.2957.
\end{aligned}
$$

Example 3.18.

If X ~ N(25, 36), then what is the value of the constant c such that

P (|X − 25| ≤ c) = 0.9544?

Answer:

$$
\begin{aligned}
0.9544 &= P(|X-25| \le c) \\
&= P(-c \le X - 25 \le c) \\
&= P\left(-\frac{c}{6} \le \frac{X-25}{6} \le \frac{c}{6}\right) \\
&= P\left(-\frac{c}{6} \le Z \le \frac{c}{6}\right) \\
&= P\left(Z \le \frac{c}{6}\right) - P\left(Z \le -\frac{c}{6}\right) \\
&= 2\,P\left(Z \le \frac{c}{6}\right) - 1.
\end{aligned}
$$

Hence

$$P\left(Z \leq \frac{c}{6}\right) = 0.9772$$

and from this, using the normal table, we get

$$c/6 = 2 \quad \text{or} \quad c = 12.$$

Example 3.19.

a) What is the probability that Z for a standard normal probability distribution is between -0.76 and 0.76 ?
b) What is the probability that Z for a standard normal probability distribution is smaller than – 0.76 or larger than 0.76 ?

Solution:

The calculation of probability is as follows:

a) $P(-0.76 < z < 0.76) = P(z < 0.76) - P(z < -0.76)$

$= 0.7764 - 0.2236$

$= 0.5528$

b) $P(z < -0.76) \cup (z > 0.76) = P(z < -0.76) + (z > 0.76) =$

$= 0.2236 + (1 - 0.7664) = 0.4472$

Example 3.20

A city installs 2000 electric lamps for street lighting. These lamps have a mean burning life of 1000 hours with a standard deviation of 200 hours. The normal distribution is a close approximation to this case.

a) What is the probability that a lamp will fail in the first 700 burning hours?

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{700 - 1000}{200} = -1.50$$

From Table A1 for $z_1 = -1.50 = (-1.5) + (-0.00)$,

$$\Pr[X < 700] = \Pr[Z < -1.50]$$
$$= \Phi(-1.50)$$
$$= 0.0668$$

Then Pr [burning life < 700 hours] = 0.0668 or 0.067.

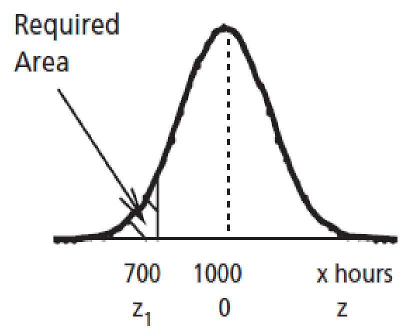


Figure 7.7: Probabilities for Example 7.2(a)

b) What is the probability that a lamp will fail between 900 and 1300 burning hours?

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{900 - 1000}{200} =$$
$$= -0.50 = (-0.5) + (-0.00)$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{1300 - 1000}{200} =$$
$$= +1.50 = (+1.5) + (0.00)$$

From Table A1, $\Phi(z_1) = \Phi(-0.50) = 0.3085$

and $\Phi(z_2) = \Phi(1.50) = 0.9332$

Then Pr [900 hours < burning life < 1300 hours]

$$= \Phi(z_2) - \Phi(z_1)$$
$$= 0.9332 - 0.3085$$
$$= 0.6247 \text{ or } 0.625.$$

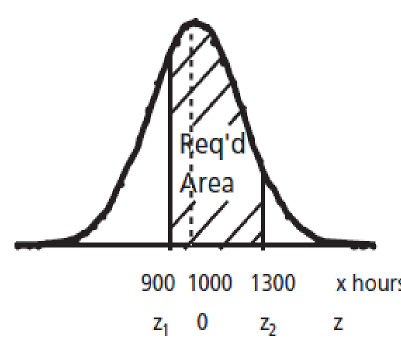


Figure 7.8: Probabilities for Example 7.2(b)

c) How many lamps are expected to fail between 900 and 1300 burning hours?

This is a continuation of part (b). The expected number of failures is given by the total number of lamps multiplied by the probability of failure in that interval. Then the expected number of failures = (2000) (0.6247) = 1249.4 or 1250 lamps. Because the burning life of each lamp is a random variable, the actual number of failures between 900 and 1300 burning hours would be only approximately 1250.

d) What is the probability that a lamp will burn for exactly 900 hours?
Since the burning life is a continuous random variable, the probability of a life of exactly 900 burning hours (not 900.1 hours or 900.01 hours or 900.001 hours, etc.) is zero. Another way of looking at it is that there are an infinite number of possible lifetimes between 899 and 901 hours, so the probability of any one of them is one divided by infinity, so zero. We saw this before in Example 6.2.

e) What is the probability that a lamp will burn between 899 hours and 901 hours before it fails?

Since this is an interval rather than a single exact value, the probability of failure in this interval is not infinitesimal (although in this instance the probability is small).

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{899 - 1000}{200} = -0.505$$

$$z_2 = \frac{901 - 1000}{200} = -0.495$$

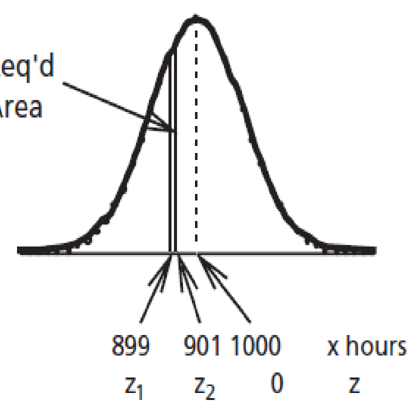


Figure 7.9: Probabilities for Example 7.2(e)

We could apply linear interpolation between the values given in Table A1. However, considering that in practice the parameters are not known exactly and the real distribution may not be exactly a normal distribution, the extra precision is not worthwhile.

Pr [899 hours < burning life < 901 hours]

$$\approx \Phi(-0.49) - \Phi(-0.50)$$
$$= \Phi(-0.4 - 0.09) - \Phi(-0.5 - 0.00)$$
$$= 0.3121 - 0.3085$$
$$= 0.0036 \text{ or } 0.4\%$$

(0.3% would also be a reasonable approximation).

f) After how many burning hours would we expect 10% of the lamps to be left?

This corresponds to the time at which

Pr [burning life > x_1 hours] = 0.10,

so Pr [burning life < x_1 hours] = 1 – 0.10 = 0.90.

Thus, Pr [$Z < z_1$] = 0.90

or $\Phi(z_1) = 0.90$

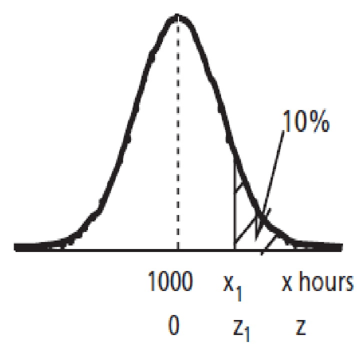


Figure 7.10: Probabilities for Example 7.2(f)

From Table A1,

$$\Phi(1.2 + 0.08) = 0.8997$$

and $$\Phi(1.2 + 0.09) = 0.9015$$

Once again, we could apply linear interpolation but the accuracy of the calculation probably does not justify it.

Since (0.90 – 0.8997) << (0.9015 – 0.90), let us take $z_1 = 1.28$. Then we have

$$z_1 = \frac{x_1 - \mu}{\sigma} = 1.28$$

$$\frac{x_1 - 1000}{200} = 1.28$$

$$x_1 = (200)(1.28) + 1000 = 1256$$

Then after 1256 hours of burning, we would expect 10% of the lamps to be left. And again, because the burning time is a random variable, performing the experiment would give a result which would be close to 1256 hours but probably not exactly that, even if the normal distribution with the given values of the mean and standard deviation applied exactly.

g) After how many burning hours would we expect 90% of the lamps to be left?

We won't draw another diagram, but imagine looking at Figure 7.10 from the back.

$\Pr[Z < z_2] = 0.10$ or $\phi(z_2) = 0.10$. From Table A1 we find

$$\phi(-1.2 - 0.08) = 0.1003$$

$$\phi(-1.2 - 0.09) = 0.0985$$

so $z_2 \approx -1.28$. (Do you see any resemblance to the answer to part (f)? Look again at equation 7.9.)

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{x_2 - 1000}{200} = -1.28$$

$$x_2 - 1000 = -256$$

$$x_2 = 744$$

After 744 hours we would expect 90% of the lamps to be left.

Exercise 3

1. What is the probability of getting exactly 3 heads in 5 flips of a fair coin?

2. On six successive flips of a fair coin, what is the probability of observing 3 heads and 3 tails?

3. If a fair coin is tossed 4 times, what is the probability of getting at least two heads?

4. Suppose X has a Poisson distribution with a standard deviation of 4. What is the conditional probability that X is exactly 1 given that $X \geq 1$?

5. Let X have a Poisson distribution with parameter $\lambda = 2$. What is the probability that $X \geq 5$ given that $X \leq 8$?

6. Find the mean and variance of an exponential distribution?

7. What is the probability that a normal random variable with mean 6 and standard deviation 3 will fall between 5.7 and 7.5 ?

8. If in a certain normal distribution of X, the probability is 0.5 that X is less than 500 and 0.0227 that X is greater than 650. What is the standard deviation of X?

9. 11. If $X \sim N(5, 4)$, then what is the probability that $8 < Y < 13$ where $Y = 2X + 1$?