MATHEMATICS

FOR

LFK, HLP AND HGL COMBINATIONS

SENIOR SIX

STUDENT'S BOOK

Experimental version

Kigali, 2023

FOREWORD

Dear Student,

Rwanda Basic Education Board (REB) is honoured to present Mathematics book for Senior six (S6) students in the following combinations: Literature in English, French, Kinyarwanda and Kiswahili (LFK); History, Literature in English and Psychology (HLP); History, Geography and Literature in English (HGL).

This book will serve as a guide to competence-based teaching and learning to ensure consistency and coherence in the learning of the Mathematics.

The Rwanda educational philosophy is to ensure that you achieve full potential at every level of education which will prepare you to be well integrated in society and exploit employment opportunities. The government of Rwanda emphasizes the importance of aligning teaching and learning materials with the syllabus to facilitate your learning process. Many factors influence what you learn, how well you learn and the competences you acquire. Those factors include the relevance of the specific content, the quality of teachers' pedagogical approaches, the assessment strategies and the instructional materials available. In this book, we paid special attention to the activities that facilitate the learning process in which you can develop your ideas and make new discoveries during concrete activities carried out individually or with peers.

In competence-based curriculum, learning is considered as a process of active building and developing knowledge and meanings by the learner where concepts are mainly introduced by an activity, situation or scenario that helps the learner to construct knowledge, develop skills and acquire positive attitudes and values.

For efficiency use of this textbook, your role is to:

- Work on given activities which lead to the development of skills;
- Share relevant information with other learners through presentations, discussions, group work and other active learning techniques such as role play, case studies, investigation and research in the library, on internet or outside;
- Participate and take responsibility for your own learning;
- Draw conclusions based on the findings from the learning activities.

To facilitate you in doing activities, the content of this book is self explanatory so that you can easily use it yourself, acquire and assess your competences. The book is made of units as presented in the syllabus. Each unit has the following structure: the key unit competence is given and it is followed by the introductory activity before the development of mathematical concepts that are connected to real world problems or to other sciences.

The development of each concept has the following points:

- It starts by a learning activity: it is a hand on well set activity to be done by students in order to generate the concept to be learnt;
- Main elements of the content to be emphasized;
- Worked examples; and
- Application activities: those are activities to be done by the user to consolidate competences or to assess the achievement of objectives.

Even though the book has some worked examples, you will succeed on the application activities depending on your ways of reading, questioning, thinking and grappling ideas of calculus not by searching for similar-looking worked out examples.

Furthermore, to succeed in Mathematics, you are asked to keep trying; sometimes you will find concepts that need to be worked at before you completely understand. The only way to really grasp such a concept is to think about it and work-related problems found in other reference books.

I wish to sincerely express my appreciation to the people who contributed towards the development of this book, particularly, REB staff, UR-CE Lecturers and teachers for their technical support. A word of gratitude goes to Head Teachers and TTCs principals who availed their staff for various activities.

Any comment or contribution would be welcome to the improvement of this text book for the next edition.

Dr. MBARUSHIMANA Nelson

Director General, REB

ACKNOWLEDGEMENT

I wish to express my appreciation to the people who played a major role in the development of this Mathematics book for Senior six (S6) students in the following combinations: Literature in English, French, Kinyarwanda and Kiswahili (LFK); History, Literature in English and Psychology (HLP); History, Geography and Literature in English (HGL).

It would not have been successful without active participation of different education stakeholders.

I owe gratitude to different universities and schools in Rwanda that allowed their staff to work with REB in the in-house textbooks production initiative.

Finally, my word of gratitude goes to the Rwanda Education Board staffs who were involved in the whole process of in-house textbook Elaboration.

Joan MURUNGI

Head of CTLR Department

Table of Contents

UNIT 1: BIVARIATE STATISTICS

Key unit Competence: Extend understanding, analysis and interpretation of bivariate data to correlation coefficients and regression lines.

1.0 Introductory activity

In Kabeza village, after her 9 observations about farming, UMULISA saw that in every house observed, where there are X cows there are also Y domestic ducks. She got the following results in form of ordered pairs $(\boldsymbol{x}, \boldsymbol{y})$:

$(1,4),(2,8),(3,4),(4,12),(5,10)$
$(6,14),(7,16),(8,6),(9,18)$

a) Represent this information graphically in $(x, y)-coordinates$.

b) Find the equation of a straight line joining many points of the graph and guess the name of this line.

c) According to your observation from (a), explain in your own words if there is any relationship between the number of Cows (X) and the number of domestic ducks (Y).

1.1 Bivariate data and scatter diagram

Activity 1.1

Consider the situation in which the mass y (g) of a chemical is related to the time, x minutes, for which the chemical reaction has been taking place. Data were given in the following table:

Time, x min	5	7	12	16	20
Mass, y g	4	12	18	21	24

a) Plot the above information in (x, y) coordinates.

b) Explain in your own words the relationship between the variation of x *and* y.

Content summary

In statistics, **bivariate** or **double series** includes technique of analysing data in two variables, the focus on the relationship between a dependent variable y and an independent variable x.

For example, the relationship or dependency between age and weight, weight and height, years of education and salary, amount of daily exercise and cholesterol level, indicate the bivariate data. As with data for a single variable, we can describe bivariate data both graphically and numerically. In both cases, we will be primarily concerned with determining whether there is a *linear* relationship between the two variables under consideration or not.

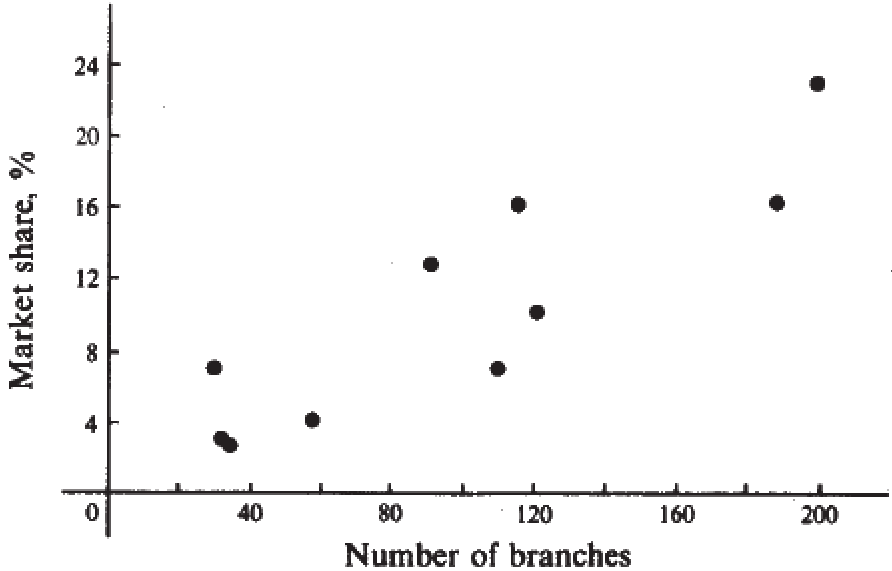
It should be kept in mind that a statistical relationship between two variables does not necessarily imply a *causal* relationship between them. For example, a strong relationship between weight and height does not imply that either variable causes the other.

Scatter plots or Scatter diagram

Consider the following data which relate *x*, the respective number of branches that 10 different banks have in a given common market, with *y*, the corresponding market share of total deposits held by the banks:

x	198	186	116	89	120	109	28	58	34	31
y	22.7	16.6	15.9	12.5	10.2	6.8	6.8	4.0	2.7	2.8

If each point (x, y) of the data is plotted in an x, y *coordinate* plane, **the *scatter plot* *or* *Scatter diagram*** is obtained.



Application Activity 1.1

One measure of personal fitness is the time taken for an individual's pulse rate to return to normal after strenuous exercise, the greater the fitness, the shorter the time. Following a short program of strenuous exercise Norman recorded his pulse rates P at time t minutes after he had stopped exercising. Norman's results are given in the table below.

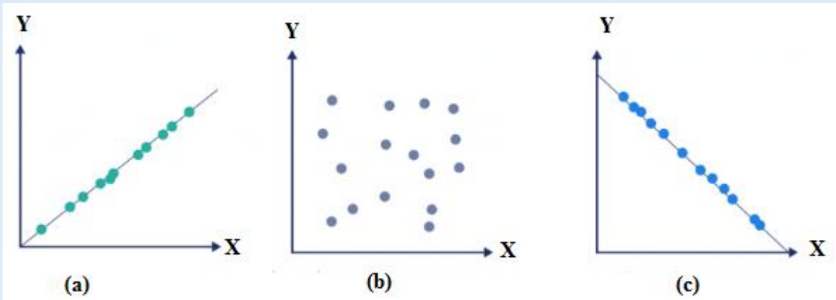
t	0.5	1.0	1.5	2.0	3.0	4.0	5.0
P	125	113	102	94	81	83	71

Draw a scatter diagram to represent this information in (x, y) *coordinates*.

1.2 Correlation and types of correlation

Activity 1.2

Observe the diagram and the graphs plotted by joining some (x,y) points of the function y = f(x):



a) At which diagram do we have many points on a straight line?

b) If we consider the linear function y = f(x), which line has a positive slope, Which one has a negative slope?

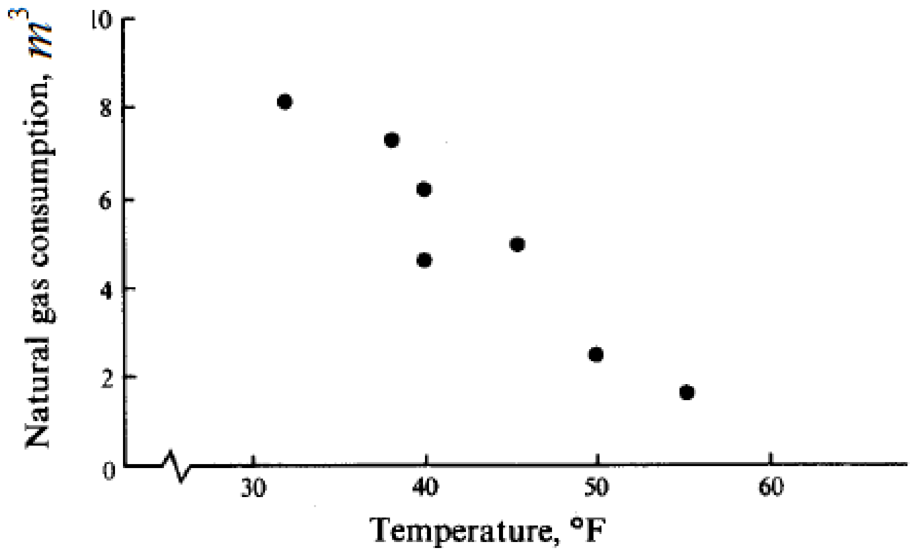
c) Consider each case; Can you say that there is a relationship between the variation of x and the variation of y?

Content summary

The scatter plot or scatter diagram (in the figure above) indicates that, roughly speaking, the market share increases as the number of branches increases. We say that *x* and *y* have a ***positive correlation*.**

On the other hand, consider the data below, which relate average daily temperature *x*, in degrees Fahrenheit, and daily natural gas consumption *y*, in cubic metre.

$x, {}^0F$	50	45	40	38	32	40	55
y, m^3	2.5	5.0	6.2	7.4	8.3	4.7	1.8

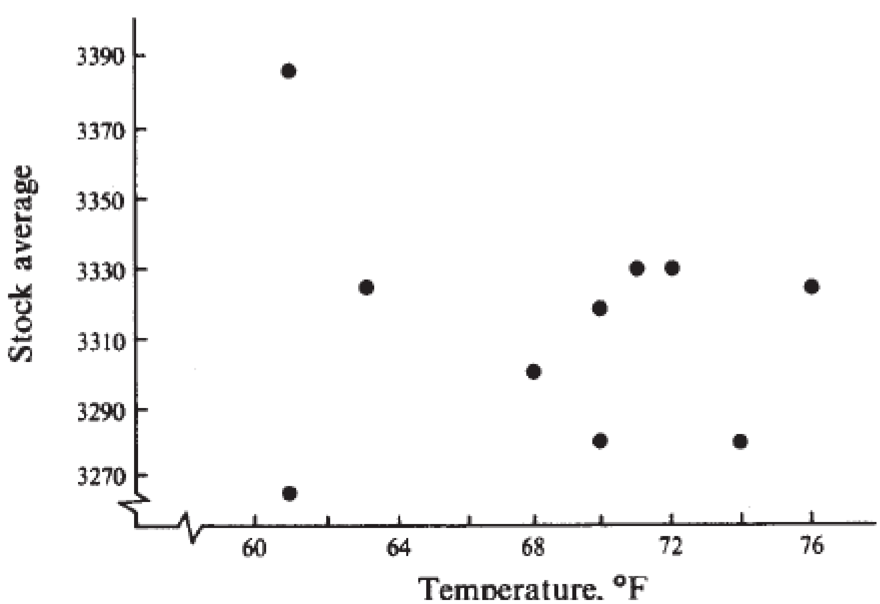


We see that *y* tends to decrease as *x increases*. Here, *x* and *y* have a ***negative correlation***

Finally, consider the data items (*x*, *y*) below, which relate daily temperature *x* over a 10-day period to the Dow Jones stock average *y*.

(63, 3385); (72, 3330); (76, 3325); (70, 3320); (71, 3330); (65, 3325); (70, 3280); (74, 3280)

(68, 3300); (61, 3265).



There is no apparent relationship between *x* and *y* (*no correlation* or *Weak correlation*).

Note that the correlation is a mutual relationship between two or more things.

Examples of correlation in real life include:

• As students' study time increases, the tests average increases too.

• As the number of trees cut down increases, soil erosion increases too.

• The more you exercise your muscles, the stronger they get.

• As a child grows, so does the clothing size.

• The more one smokes, the fewer years he will have to live.

Types of correlation

Correlation can be **negative or positive** depending on the situation:

In a situation where one variable positively affects another variable, we say positive correlation has occurred.

Examples:

1) The more times people have unprotected sex with different partners, the more the rates of HIV is in a society.

2) As students study time increases, the tests average increases too.

On the other hand, when one variable affects another variable negatively, we say negative correlation has occurred.

Examples:

The more alcohol one consumes, the less the judgment one has

Correlation is a scatter diagram which can be determined whether it is positive or negative by following the trend of the points and the gradient of the line of the best fit.

Application Activity 1.2

One measure of personal fitness is the time taken for an individual's pulse rate to return to normal after strenuous exercise, the greater the fitness, the shorter the time. Following a short program of strenuous exercise Norman recorded his pulse rates P at time t minutes after he had stopped exercising. Norman's results are given in the table below.

t	0.5	1.0	1.5	2.0	3.0	4.0	5.0
P	125	113	102	94	81	83	71

Explain the relationship between Norman's pulse P and time t.

1.3 Covariance

Activity 1.3

Complete the following table

i	x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
1	3	6			
2	5	9			
3	7	12			
4	3	10			
5	2	7			
6	6	8			
	$\sum_{i=1}^{6} x_i = \ldots$	$\sum_{i=1}^{6} y_i = \ldots$			$\sum_{i=1}^{6} (x_i - \bar{x})(y_i - \bar{y}) = \ldots$
	$\bar{x} = \ldots..$	$\bar{y} = \ldots..$			

What can you get from the following expressions?

1) $\sum_{i=1}^{k} (x_i - \bar{x})(x_i - \bar{x})$

2) $\sum_{i=1}^{k} (x_i - \bar{x})(y_i - \bar{y})$

Content summary

In case of two variables, say x and y, there is another important result called **covariance of *x* and *y***, denoted $\operatorname{cov}(x, y)$.

The **covariance of variables *x* and *y*** is a measure of how these two variables change together. If the greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the smaller values, i.e. the variables tend to show similar behaviour, the covariance is positive. In the opposite case, when the greater values of one variable mainly correspond to the smaller values of the other, i.e. the variables tend to show opposite behaviour, the covariance is negative. If covariance is zero the variables are said to be **uncorrelated,** it means that there is no linear relationship between them.

Therefore, the sign of covariance shows the tendency in the linear relationship between the variables. The magnitude of covariance is not easy to interpret.

Covariance of variables *x* and *y*, where the summation of frequencies $\sum_{i=1}^{k} f_i = n$ are equal for both variables, is defined to be

$$\operatorname{cov}(x,y) = \frac{1}{n}\sum_{i=1}^{k} f_i\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$$

Developing this formula, we have

$$\begin{aligned}
\operatorname{cov}(x,y) &= \frac{1}{n}\sum_{i=1}^{k} f_i\left(x_i y_i - x_i\,\bar{y} - \bar{x}\,y_i + \bar{x}\,\bar{y}\right) \\
&= \frac{1}{n}\sum_{i=1}^{k} f_i x_i y_i - \frac{1}{n}\sum_{i=1}^{k} f_i x_i\;\bar{y} - \frac{1}{n}\sum_{i=1}^{k} f_i \bar{x}\,y_i + \frac{1}{n}\sum_{i=1}^{k} f_i \bar{x}\;\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{k} f_i x_i y_i - \frac{1}{n}\bar{y}\sum_{i=1}^{k} f_i x_i - \frac{1}{n}\bar{x}\sum_{i=1}^{k} f_i\,y_i + \bar{x}\;\bar{y}\frac{1}{n}\sum_{i=1}^{k} f_i \qquad \left[\frac{1}{n}\sum_{i=1}^{k} f_i = \frac{1}{n}\times n = 1\right] \\
&= \frac{1}{n}\sum_{i=1}^{k} f_i x_i y_i - \bar{x}\;\bar{y} - \bar{x}\;\bar{y} + \bar{x}\;\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{k} f_i x_i y_i - \bar{x}\;\bar{y}
\end{aligned}$$

Thus, the covariance is also given by

$$\operatorname{cov}(x,y) = \frac{1}{n}\sum_{i=1}^{k} f_i x_i y_i - \bar{x}\;\bar{y}$$

Examples

1) Find the covariance of x and y in following data sets

x	3	5	6	8	9	11
y	2	3	4	6	5	8

Solution

We have

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
3	2	-4	-2.6	10.4
5	3	-2	-1.6	3.2
6	4	-1	-0.6	0.6
8	6	1	1.4	1.4
9	5	2	0.4	0.8
11	8	4	3.4	13.6
$\sum_{i=1}^{6} x_i = 42$	$\sum_{i=1}^{6} y_i = 28$			$\sum_{i=1}^{6}(x_i - \bar{x})(y_i - \bar{y}) = 30$
$\bar{x} = 7$	$\bar{y} = 4.6$			

Thus,

$$\begin{aligned} \operatorname{cov}(x, y) &= \frac{1}{6}\sum_{i=1}^{6}(x_i - \bar{x})(y_i - \bar{y}) \\ &= \frac{1}{6}(30) \\ &= 5 \end{aligned}$$

2) Find the covariance of the following distribution

y \ x	0	2	4
1	2	1	3
2	1	4	2
3	2	5	0

Solution

Convert the double entry into a simple table and compute the arithmetic means

x_i	y_i	f_i	$x_i f_i$	$y_i f_i$	$x_i y_i f_i$
0	1	2	0	2	0
0	2	1	0	2	0
0	3	2	0	6	0
2	1	1	2	1	2
2	2	4	8	8	16
2	3	5	10	15	30
4	1	3	12	3	12
4	2	2	8	4	16
4	3	0	0	0	0
		$\sum_{i=1}^{9} f_i = 20$	$\sum_{i=1}^{9} x_i f_i = 40$	$\sum_{i=1}^{9} y_i f_i = 41$	$\sum_{i=1}^{9} x_i y_i f_i = 76$

$$\bar{x} = \frac{40}{20} = 2\,,\ \bar{y} = \frac{41}{20} = 2.05$$

$$\operatorname{cov}(x, y) = \frac{76}{20} - 2 \times 2.05 = -0.3$$

Alternative method

$$\operatorname{cov}(x, y) = \frac{1}{n}\sum_{i=1}^{k} f_i x_i y_i - \bar{x}\,\bar{y}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i f_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{k} y_i f_i$$

y \ x	0	2	4	Total
1	2	1	3	6
2	1	4	2	7
3	2	5	0	7
Total	5	10	5	20

$$\bar{x} = \frac{1}{20}(0\times5+2\times10+4\times5)$$
$$= \frac{40}{20} = 2$$

$$\bar{y} = \frac{1}{20}(1\times6+2\times7+3\times7)$$
$$= \frac{41}{20} = 2.05$$

$$\operatorname{cov}(x,y) = \frac{1}{20}(0\times1\times2+0\times2\times1+0\times3\times2+2\times1\times1+2\times2\times4+2\times3\times5+4\times1\times3+4\times2\times2+4\times3\times0)$$
$$-2\times2.05$$
$$= \frac{1}{20}(0+0+0+2+16+30+12+16+0)-4.1$$
$$= \frac{76}{20}-4.1 \; = -0.33$$

Application activity 1.3

1. The scores of 12 students in their mathematics and physics classes are

Mathematics	2	3	4	4	5	6	6	7	7	8	10	10
Physics	1	3	2	4	4	4	6	4	6	7	9	10

Find the covariance of the distribution

2. The values of two variables x and y are distributed according to the following table

y \ x	100	50	25
14	1	1	0
18	2	3	0
22	0	1	2

Calculate the covariance

1.4 Coefficient of correlation and properties related to coefficient of correlation

Activity 1.4

Consider the following table

x	y
3	6
5	9
7	12
3	10
2	7
6	8

1. Find σ_x, σ_y
2. Find $\operatorname{cov}(x, y)$
3. Calculate the ratio $\dfrac{Cov(x, y)}{\sigma_x . \sigma_y}$

Content summary

The **Pearson's coefficient of correlation** (or **Product moment coefficient of correlation** or simply **coefficient of correlation**), denoted by *r*, is a measure of the strength of linear relationship between two variables.

The coefficient of correlation between two variables *x* and *y* is given by

$$r = \frac{\operatorname{cov}(x, y)}{\sigma_x \sigma_y}$$

Where, $\operatorname{cov}(x, y)$ is covariance of *x* and *y*

σ_x is the standard deviation for x;

σ_y is the standard deviation for y.

Properties of the coefficient of correlation

a) The coefficient of correlation does not change the measurement scale. That is, if the height is expressed in meters or feet, the coefficient of correlation does not change.

b) The sign of the coefficient of correlation is the same as the covariance.

c) The square of the coefficient of correlation is equal to the product of the gradient of the regression line of y on x, and the gradient of the regression line of x on y.

In fact, $r = \frac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y}$. Squaring both sides gives

$$r^2 = \left[\frac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y}\right]^2$$

$$= \frac{\operatorname{cov}^2(x,y)}{\sigma_x^2 \sigma_y^2}$$

$$= \frac{\operatorname{cov}(x,y)}{\sigma_x^2} \times \frac{\operatorname{cov}(x,y)}{\sigma_y^2}$$

$= ac$, where $y = ax + b$ is the equation of the regression line of y on x, and $x = cy + d$ is the equation of the regression line of x on y

d) If the coefficient of correlation is known, it can be used to find the gradients or slopes of two regression lines.

We know that the gradient of the regression line y on x is $\frac{\operatorname{cov}(x,y)}{\sigma_x^2}$.

From this we have,
$$\frac{\operatorname{cov}(x,y)}{\sigma_x^2} = \frac{\operatorname{cov}(x,y)}{\sigma_x \sigma_x} \times \frac{\sigma_y}{\sigma_y}$$

$$= \frac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y} \times \frac{\sigma_y}{\sigma_x}$$

$$= r\frac{\sigma_y}{\sigma_x}$$

We know that the gradient of the regression line of x on y is $\frac{\operatorname{cov}(x,y)}{\sigma_y^2}$.

From this we have, $\frac{\operatorname{cov}(x,y)}{\sigma_y^2}=\frac{\operatorname{cov}(x,y)}{\sigma_y\sigma_y}\times\frac{\sigma_x}{\sigma_x}$

$$=\frac{\operatorname{cov}(x,y)}{\sigma_x\sigma_y}\times\frac{\sigma_x}{\sigma_y}$$

$$=r\frac{\sigma_x}{\sigma_y}$$

Thus, the gradient of the regression line of y on x is given by $r\frac{\sigma_y}{\sigma_x}$ and the gradient of the regression line of x on y is given by $r\frac{\sigma_x}{\sigma_y}$.

e) Cauchy Inequality: $\operatorname{cov}^2(x,y)\le\sigma_x^2\sigma_y^2$

In fact, $r=\frac{\operatorname{cov}(x,y)}{\sigma_x\sigma_y}\Leftrightarrow\operatorname{cov}(x,y)=r\sigma_x\sigma_y$.

Squaring both sides gives $\operatorname{cov}^2(x,y)=r^2\sigma_x^2\sigma_y^2$

Or $\operatorname{cov}^2(x,y)\le\sigma_x^2\sigma_y^2$

f) The coefficient of correlation takes value ranging between -1 and +1. That is $-1\le r\le 1$

In fact, from Cauchy Inequality we have,

$$\operatorname{cov}^2(x,y)\le\sigma_x^2\sigma_y^2$$

$$\Leftrightarrow\frac{\operatorname{cov}^2(x,y)}{\sigma_x^2\sigma_y^2}\le 1$$

$$\Leftrightarrow\left[\frac{\operatorname{cov}(x,y)}{\sigma_x\sigma_y}\right]^2\le 1$$

$$\Leftrightarrow r^2\le 1.$$

Taking square roots both side

$$\Leftrightarrow\sqrt{r^2}\le 1$$

$$\Leftrightarrow|r|\le 1 \text{ since } \sqrt{x^2}=|x|$$

$|r|\le 1$ is equivalent to $-1\le r\le 1$.

Thus, $-1 \le r \le 1$.

g) If the linear coefficient of correlation takes values closer to **−1**, the **correlation** is **strong and negative**, and will become stronger the closer ***r*** approaches -1.

h) If the linear coefficient of correlation takes values close to **1** the **correlation** is **strong and positive**, and will become stronger the closer ***r*** approaches 1

i) If the linear coefficient of correlation takes values close to **0**, the **correlation** is **weak**.

j) If ***r = 1***or ***r = −1***, there is **perfect correlation** and the line on the scatter plot is increasing or decreasing respectively.

k) If ***r = 0***, there is **no linear correlation**.

Example

A test is made over 200 families on number of children (x) and number of beds y per family. Results are collected in the table below

y \ x	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	2	7	5	2	0	0	0	0	0	0
2	2	2	10	8	15	1	0	0	0	0	0
3	1	3	5	6	8	6	1	0	0	0	0
4	0	2	8	2	6	12	10	8	0	0	0
5	0	1	0	2	5	6	10	5	7	3	3
6	0	0	0	2	4	5	5	2	3	3	2

a) What is the average number for children and beds per a family?

b) Find the covariance

c) Can we confirm that there is a high linear correlation between the number of children and number of beds per family?

Solution

a) Average number of children per family:

Contingency table:

y \ x	0	1	2	3	4	5	6	7	8	9	10	Total
1	0	2	7	5	2	0	0	0	0	0	0	16
2	2	2	10	8	15	1	0	0	0	0	0	38
3	1	3	5	6	8	6	1	0	0	0	0	30
4	0	2	8	2	6	12	10	8	0	0	0	48
5	0	1	0	2	5	6	10	5	7	3	3	42
6	0	0	0	2	4	5	5	2	3	3	2	26
Total	**3**	**10**	**30**	**25**	**40**	**30**	**26**	**15**	**10**	**6**	**5**	**200**

$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{k} f_i y_i$

Marginal series:

x_i	0	1	2	3	4	5	6	7	8	9	10	Total
f_i	3	10	30	25	40	30	26	15	10	6	5	$\sum f_i = 200$
$f_i x_i$	0	10	60	75	160	150	156	105	80	54	50	$\sum f_i x_i = 900$

y_i	1	2	3	4	5	6	Total
f_i	16	38	30	48	42	26	$\sum f_i = 200$
$f_i y_i$	16	76	90	192	210	156	$\sum f_i y_i = 740$

The means are

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{11} f_i x_i = \frac{900}{200} = 4.5$$

And $\bar{y} = \frac{1}{n}\sum_{i=1}^{6} f_i y_i = \frac{740}{200} = 3.7$

There are about 5 children per family and about 4 beds per family.

b) The covariance is calculated as follow:

$Cov(x,y) = \left(\frac{1}{n}\sum_{(i,j)=(1,1)}^{(p,q)=(11,6)} f_{i,j} x_i y_j\right) - \bar{x}\bar{y}$, where i assumes values from 1 to $p = 11$, and j assumes values from 1 to $q = 6$,or

$\operatorname{cov}(x,y) = \frac{1}{200}\sum_{i=1}^{66} f_i x_i y_i - \bar{x}\,\bar{y}$ Where $\bar{y} = 3.7$ and $\bar{x} = 4.5$

$$= \frac{1}{200}\begin{pmatrix} 0+2+14+15+8+0+4+40+48+120+10+0 \\ +9+30+54+96+90+18+0+8+64+24+96 \\ +240+240+224+0+5+0+30+100+150 \\ +300+175+280+135+150+0+36+96+150 \\ +180+84+144+162+120 \end{pmatrix} - 4.5\times 3.7$$

$$\begin{aligned} &= \frac{3751}{200} - 16.65 \\ &= 18.7555 - 16.65 \\ &= 2.105 \end{aligned}$$

c) The correlation coefficient is given by $\frac{\operatorname{cov}(x,y)}{\sigma_x \sigma_y}$

$$\sigma_y^2 = \frac{1}{200}\sum_{i=1}^{6} f_i y_i^2 - \left(\overline{y}\right)^2$$
$$= \frac{1}{200}\left(16 + 38\times 4 + 30\times 9 + 48\times 16 + 42\times 25 + 26\times 36\right) - \left(3.7\right)^2$$
$$= 15.96 - 13.69$$
$$= 2.27$$

$$\sigma_x^2 = \frac{1}{200}\sum_{i=1}^{11} f_i x_i^2 - \left(\overline{x}\right)^2 = \frac{1}{200}\begin{pmatrix} 0 + 10 + 4\times 30 + 9\times 25 + 40\times 16 + 30\times 25 + \\ 36\times 26 + 49\times 15 + 64\times 10 + 81\times 6 + 100\times 5 \end{pmatrix} - \left(4.5\right)^2$$
$$= 25.21 - 20.25 = 4.96$$

Therefore, the correlation coefficient is

$$r = \frac{2.105}{\sqrt{4.96}\sqrt{2.27}} \approx 0.63$$

There is a high linear correlation.

Application activity 1.4

1) The scores of 12 students in their mathematics and physics classes are

Mathematics	2	3	4	4	5	6	6	7	7	8	10	10
Physics	1	3	2	4	4	4	6	4	6	7	9	10

Find the correlation coefficient distribution and interpret it.

2) The values of the two variables X and Y are distributed according to the following table:

X \ Y	0	2	4
1	2	1	3
2	1	4	2
3	2	5	0

Calculate the correlation coefficient.

1.5 Regression lines

Activity 1.5

Given the data in the table below:

x	5	7	12	16	20
y	4	12	18	21	24

Determine:

a) Variance of x

b) Variance of y

c) Covariance *of* (x, y)

d) The value a given by $\dfrac{\operatorname{cov}(x,y)}{\operatorname{var} x} = \dfrac{S_{x,y}}{S_{x,x}}$

e) The value b given by $b = \overline{Y} - a\overline{X}$

f) Establish the equation of the line $y = ax + b$

g) In the Cartesian plane, represent the data (x, y) in the table above and the line $y = ax + b$ found in (f).

h) Discuss the position of the points of coordinate (x, y) with respect to the line $y = ax + b$

Content summary

We use the regression line of y on x to **predict** a value of y for any given value of x and vice versa, we use the regression line of x *on* y, to predict a value of x for a given value of y. The "best" line would make the best predictions: the observed y-values should stray as little as possible from the line. This straight line is the regression line from which we can adjust its algebraic expressions and it is written as $y = ax + b$, where a *is* the gradient and b is the y-intercept.

The regression line y on x is written as $y=\frac{\operatorname{cov}(x,y)}{\sigma_x^2}x+\left(\bar{y}-\frac{\operatorname{cov}(x,y)}{\sigma_x^2}\bar{x}\right)$

The above regression line can be re-written as

$$L_{y/x}\equiv y-\bar{y}=\frac{\operatorname{cov}(x,y)}{\sigma_x^2}\left(x-\bar{x}\right)$$

Note that the regression line x on y is $x=cy+d$, where c is the gradient of the line and

d is the x-intercept , it is given by $x=\frac{\operatorname{cov}(x,y)}{\sigma_y^2}y+\left(\bar{x}-\frac{\operatorname{cov}(x,y)}{\sigma_y^2}\bar{y}\right)$

This line is written as $L_{x/y}\equiv x-\bar{x}=\frac{\operatorname{cov}(x,y)}{\sigma_y^2}\left(y-\bar{y}\right)$

Shortcut method of finding regression line

To abbreviate the calculations, the two regression lines can be determined as follow:

a) The equation of the regression line of y on x is $L_{y/x}\equiv y=ax+b$ and the values of a and b are found by solving the simultaneous equations:

$$\begin{cases}\sum_{i=1}^{k} f_i y_i = a\sum_{i=1}^{k} f_i x_i + b\,n\\ \sum_{i=1}^{k} f_i x_i y_i = a\sum_{i=1}^{k} f_i x_i^2 + b\sum_{i=1}^{k} f_i x_i\end{cases}$$

These equations are called the **normal equations** for y on x.

b) The equation of the regression line of x on y is $L_{x/y}\equiv x=cy+d$ and the values of c and d are found by solving the simultaneous equations:

$$\begin{cases}\sum_{i=1}^{k} f_i x_i = c\sum_{i=1}^{k} f_i y_i + d\,n\\ \sum_{i=1}^{k} f_i x_i y_i = c\sum_{i=1}^{k} f_i y_i^2 + d\sum_{i=1}^{k} f_i y_i\end{cases}$$

These equations are called the **normal equations** for x on y.

Examples:

Let us consider the following bivariate data for this situation. A farmer sold hens in different periods and received interest in thousands of Rwandan francs. The following table shows the number x of hens sold and the corresponding interest y in thousands of Rwandan francs earned in different periods:

x (number of hens)	3	5	6	8	9	11
y (thousands Frw)	2	3	4	6	5	8

Find the equation of the regression line of y on x , and the equation of the regression line of x on y ,for the following data and estimate the value of y for $x=4, x=7, x=16$ *and the value of x for* $y=7, y=9, y=16$.

Solution

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
3	2	-4	-2.6	16	6.76	10.4
5	3	-2	-1.6	4	2.56	3.2
6	4	-1	-0.6	1	0.36	0.6
8	6	1	1.4	1	1.96	1.4
9	5	2	0.4	4	0.16	0.8
11	8	4	3.4	16	11.56	13.6
$\sum_{i=1}^{6} x_i = 42$	$\sum_{i=1}^{6} y_i = 28$			$\sum_{i=1}^{6}(x_i-\bar{x})^2 = 42$	$\sum_{i=1}^{6}(y_i-\bar{y})^2 = 23.36$	$\sum_{i=1}^{6}(x_i-\bar{x})(y_i-\bar{y}) = 30$

$\bar{x} = \frac{42}{6} = 7$, $\bar{y} = \frac{28}{6} = 4.7$

$$\text{cov}(x,y) = \frac{1}{n}\sum(x-\bar{x})(y-\bar{y}) = \frac{30}{6} = 5$$

$$\sigma_x^2 = \frac{42}{6} = 7 \text{ , } \sigma_y^2 = \frac{23.36}{6} = 3.89$$

$$L_{y/x} \equiv y - \bar{y} = \frac{\operatorname{cov}(x, y)}{\sigma_x^2}(x - \bar{x})$$

$$L_{y/x} \equiv y - 4.7 = \frac{5}{7}(x - 7)$$

The equation of the regression line of y on x is generally $L_{y/x} \equiv y - \bar{y} = \frac{Cov(x, y)}{\sigma_x^2}(x - \bar{x})$ where

$$Cov(x, y) = \frac{1}{n}\sum_{i=1}^{6}(x_i - \bar{x})(y_i - \bar{y}) = \frac{30}{6} = 5$$

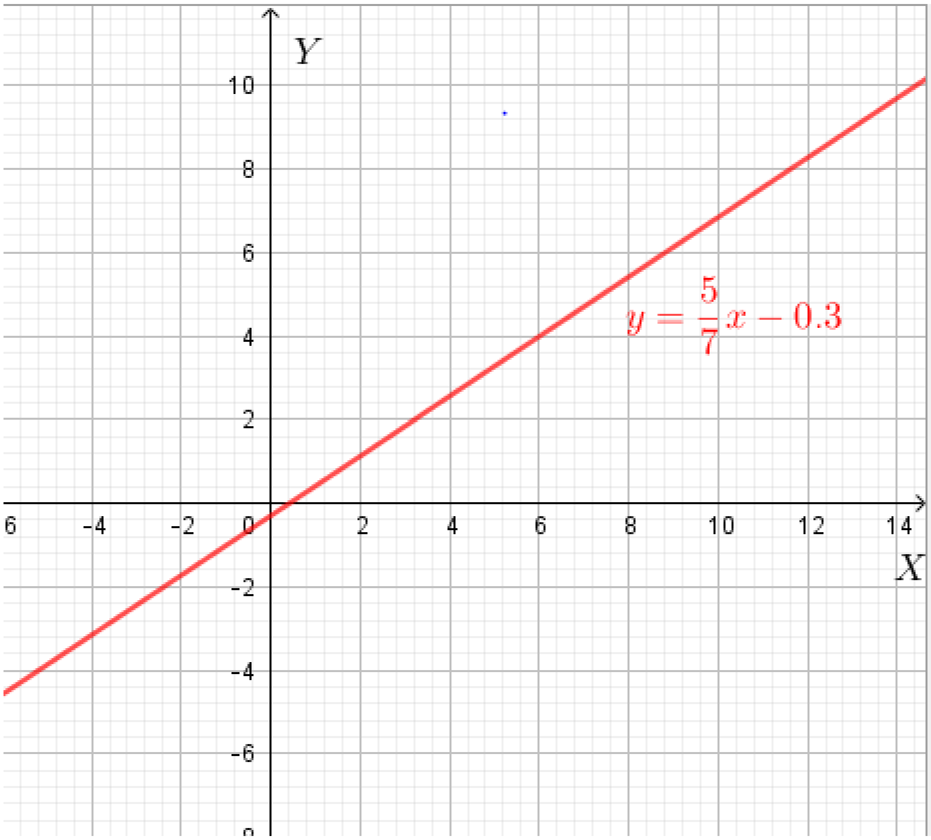
In this case it becomes:

$$L_{y/x} \equiv y = \frac{5}{7}x - 0.3\,.$$

Using this equation, we can predict the interest to be earned when the farmer sells 4 hens, 7 hens or 16 hens.

x	4	7	16
y	2.5	4.7	11.1

By joining the obtained points (number of hens, interest), we can get the following figure: Since the data appears to be linearly related, we can find a straight-line model that fits the data better than all other possible straight-line models.



And

$$L_{x/y} \equiv x - \bar{x} = \frac{\operatorname{cov}(x, y)}{\sigma_y^2}(y - \bar{y})$$

$$L_{x/y} \equiv x - 7 = \frac{5}{3.89}(y - 4.7)$$

Finally, the equation of the regression line of x on y is $L_{x/y} \equiv x = 1.3y + 1$

On the other hand, when we have the interest y, we can predict the number of hens sold using the equation of the regression line of x on y. This is $L_{x/y} \equiv x - \bar{x} = \frac{Cov(x, y)}{\sigma_y^2}(y - \bar{y})$.

Hence, $L_{x/y} \equiv x = 1.3y + 1$.

Using this equation, we can predict the number of hens to be sold if the farmer received the interest of 7000Frw, 9000Frw or 16000Frw

y	7	9	16
x	10.1	12.7	21.8

Note: As x represents the number of hens, it should be an integer.

Therefore, the interest of 7000Frw, 9000Frw and 16000Frw correspond to 10 hens, 13 hens and 22 hens respectively.

Alternative method

x	y	x^2	y^2	xy
3	2	9	4	6
5	3	25	9	15
6	4	36	16	24
8	6	64	36	48
9	5	81	25	45
11	8	121	64	88
$\sum_{i=1}^{6} x_i = 42$	$\sum_{i=1}^{6} y_i = 28$	$\sum_{i=1}^{6} x_i^2 = 336$	$\sum_{i=1}^{6} y_i^2 = 154$	$\sum_{i=1}^{6} x_i y_i = 226$

$$L_{y/x} \equiv y = ax + b$$

$$\begin{cases} \sum_{i=1}^{k} f_i y_i = a \sum_{i=1}^{k} f_i x_i + b\,n \\ \sum_{i=1}^{k} f_i x_i y_i = a \sum_{i=1}^{k} f_i x_i^2 + b \sum_{i=1}^{k} f_i x_i \end{cases}$$

$$\begin{cases} 28 = 42a + 6b \\ 226 = 336a + 42b \end{cases} \Leftrightarrow \begin{cases} a = \frac{5}{7} \\ b = -0.3 \end{cases}$$

Thus, the line of y on x is $L_{y/x} \equiv y = \frac{5}{7}x - 0.3$

$x = 4 \Rightarrow y = 2.5$

$x = 7 \Rightarrow y = 4.7$

$x = 16 \Rightarrow y = 11.1$

$$L_{x/y} \equiv x = cy + d$$

$$\begin{cases}\sum_{i=1}^{k} f_i x_i = c\sum_{i=1}^{k} f_i y_i + d\,n \\ \sum_{i=1}^{k} f_i x_i y_i = c\sum_{i=1}^{k} f_i y_i^2 + d\sum_{i=1}^{k} f_i y_i\end{cases}$$

$$\begin{cases}42 = 28c + 6d \\ 226 = 154c + 28d\end{cases} \Leftrightarrow \begin{cases}c = 1.3 \\ d = 1\end{cases}$$

Thus, the line of x on y is $L_{x/y} \equiv x = 1.3y + 1$

$y = 7 \Rightarrow x = 10.1$
$y = 9 \Rightarrow x = 12.7$
$y = 16 \Rightarrow x = 21.8$

Application activity 1.5

1. Consider the following table

x	y
60	3.1
61	3.6
62	3.8
63	4
65	4.1

   a) Find the regression line of y on x

   b) Calculate the approximate y value for the variable $x = 64$

2. The values of two variables x and y are distributed according to the following table

y \ x	100	50	25
14	1	1	0
18	2	3	0
22	0	1	2

Find the regression lines

1.6 Interpretation of statistical data

Activity 1.6

Explain in your own words how statistics, especially bivariate statistics, can be used in our daily life.

Content summary

Bivariate statistics can help in prediction of a value for one variable if we know the value of the other.

Examples

1) One measure of personal fitness is the time taken for an individual's pulse rate to return to normal after strenuous exercise, the greater the fitness, the shorter the time. Following a short program of strenuous exercise Norman recorded his pulse rates P at time t minutes after he had stopped exercising. Norman's results are given in the table below.

t	0.5	1.0	1.5	2.0	3.0	4.0	5.0
P	125	113	102	94	81	83	71

Estimate Norman's pulse rate 2.5 minutes after stopping the exercise program.

Solution

t	P	t^2	P^2	tP
0.5	125	0.25	15625	62.5
1	113	1	12769	113
1.5	102	2.25	10404	153
2	94	4	8836	188
3	81	9	6561	243
4	83	16	6889	332
5	71	25	5041	355
$\sum_{i=1}^{7} t_i = 17$	$\sum_{i=1}^{7} P_i = 669$	$\sum_{i=1}^{7} t_i^2 = 57.5$	$\sum_{i=1}^{7} P_i^2 = 66125$	$\sum_{i=1}^{7} t_i P_i = 1446.5$

We need the line $P = at + b$

Use the formula

$$\begin{cases} \sum_{i=1}^{7} P_i = a\sum_{i=1}^{7} t_i + bn \\ \sum_{i=1}^{7} t_i P_i = a\sum_{i=1}^{7} t_i^2 + b\sum_{i=1}^{7} t_i \end{cases}$$

We have: $\begin{cases} 669 = 17a + 7b \\ 1446.5 = 57.5a + 17b \end{cases}$

Solving we have $\begin{cases} a = -11 \\ b = 122.3 \end{cases}$

Then $P = -11t + 122.3$

So, the Norman's pulse rate 2.5 minutes after stopping the exercise program is estimated to be $P = -11(2.5) + 122.3$ or 94.8.

2) Collect data on the mass and the height of 10 people. If X represents the mass and Y represents the height.

a) Organise the data in the table

b) Plot the scatter diagram

c) Calculate the mean of x, the mean of y, the variance of x and the variance of y

d) Calculate the $\text{cov}(x, y)$, the correlation coefficient (r) then interpret the relationship between the mass and the height of these people

e) Establish the equation of the regression line y in function of x

Application activity 1.6

An old film is treated with a chemical in order to improve the contrast. Preliminary tests on nine samples drawn from a segment of the film produced the following results.

Sample	A	B	C	D	E	F	G	H	I
x	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
y	49	60	66	62	72	64	89	90	90

The quantity x is a measure of the amount of chemical applied, and y is the constant index, which takes values between 0 (no contrasts) and 100 (maximum contrast).

i) Plot a scatter diagram to illustrate the data.

ii) It is subsequently discovered that one of the samples of film was damaged and produced an incorrect result. State which sample you think this was**.**

In all subsequent calculations, this incorrect sample was ignored. The remaining data can be summarized as follows: $\sum x = 23.5$, $\sum y = 584$, $\sum x^2 = 83.75$, $\sum y^2 = 44622$, $\sum xy = 1883$, $n = 8$.

iii) Calculate the product moment correlation coefficient,

iv) State with a reason whether it is sensible to conclude from your answer to part (iii) that x and y are linearly related.

v) The line of regression of y on x has equation $y = a + bx$. Calculate the value of a and b, each correct to three significant figures.

vi) Use your regression line to estimate what the contrast index corresponding to the damaged piece of film would have been if the piece has been undamaged.

vii) State with a reason, whether it would be sensible to use your regression equation to estimate the contrast index when the quantity of chemical applied to the film is zero.

1.7 End unit assessment

1) The scores awarded by two judges at an ice skating competition are shown in the table.

Competitor	A	B	C	D	E	F	G	H	I	J
Judge A	5	6.5	8	9	4	2.5	7	5	6	3
Judge b	6	7	8.5	9	5	4	7.5	5	7	4.5

a) Construct a scatter diagram for this data with Judge A's scores on the horizontal axis and Judge B's scores on the vertical axis.

b) Copy and complete the following comments about the scatter diagram: There appears to be ____correlation between Judge A's scores and Judge B's scores. This means that as Judge A's scores increase, Judge B's scores ______

2) The following results were obtained from line-ups in Mathematics and Physics examinations:

	Mathematics (x)	*Physics* (y)
Mean	47.5	39.5
Standard deviation	16.8	10.8

$r = 0.95$

Find both equations of the regression lines. Also estimate the value of y for $x = 30$.

3) For a set of 20 pairs of observations of the variables x and y, it is known that $\sum_{i=1}^{k} f_i x_i = 250$, $\sum_{i=1}^{k} f_i y_i = 140$, and that the regression line of y on x passes through $(15,10)$. Find the equation of that regression line and use it to estimate y when $x = 10$.

4) In a partially destroyed laboratory record of an analysis of correlation data, the following results only are legible: Variance of x is 9

Equations of regression lines: $8x - 10y + 66 = 0$ and $40x - 18y - 214 = 0$

What were

a) the mean values of x and y

b) the standard deviation of y, and

the coefficient of correlation between x and y.

UNIT 2: COUNTING TECHNIQUES

Key Unit competence: Use counting techniques to determine the number of possible outcomes of events occurring in real life

2.0 Introductory activity

Suppose a college has 3 different history courses, 4 different literature courses, and 2 different sociology courses.

(a) Find the number m of ways a student can choose one of each kind of courses.

($m = 3(4)(2) = 24$).

(b) Find the number n of ways a student can choose just one of the courses.

($n = 3 + 4 + 2 = 9$).

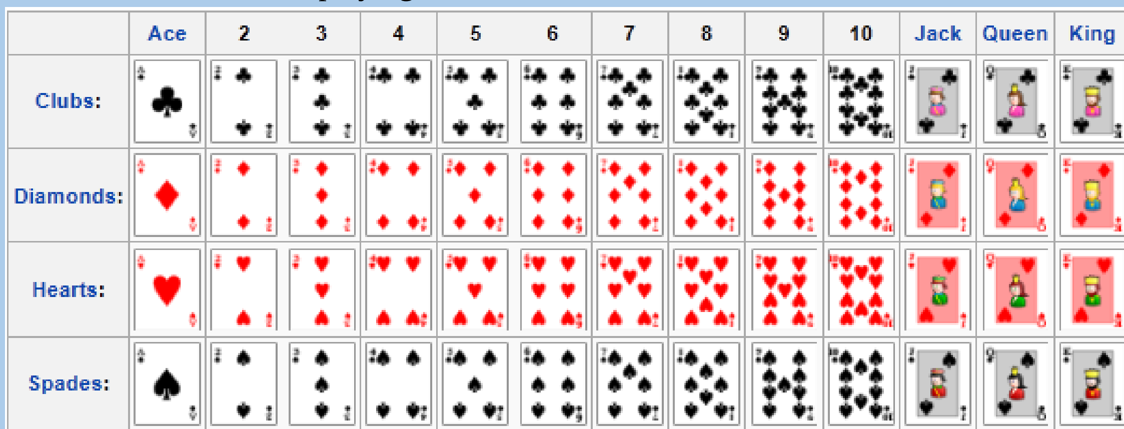
There are some techniques for determining, without direct enumeration, the number of possible outcomes of a particular event or the number of elements in a set.

Tell some example of counting techniques. Which technique can you use to find the required number in a) or in b)?

2.1 Meaning of event, sample space and related terms

Activity 2.1

Consider the deck of 52 playing cards



1. Suppose that you are choosing one card

a) How many possibilities do you have for the cards to be chosen?

b) How many possibilities do you have for the kings to be chosen?

c) How many possibilities do you have for the aces of hearts to be chosen?

2. If “selecting a queen is an example of event, give other examples of events.

Content summary

Consider the context of a game of 52 playing cards. In a deck of 52 playing cards, cards are divided into four suits of 13 cards each. If a player selects a card at random (by simple random sampling), then each card has the same chance of being selected.

In this case, there are 52 different ways of selecting one card.

When a coin is tossed, it may show Head (H-face with logos) or Tail (T-face with another symbol).

We cannot say beforehand whether it will show head up or tail up. That depends on chance.

Random experiments and Events

A **random experiment** is an experiment whose outcome cannot be predicted or determined in advance.

Example of experiments:

-Tossing a coin,

-Throwing a dice

- Selecting a card from a pack of playing cards, etc.

In all these cases there are a number of possible results (outcomes) which can occur but there is an uncertainty as to which one of them will actually occur.

Each performance in a random experiment is called a **trial**.

The result of a trial in a random experiment is called an **outcom**e, an elementary event, or a sample point.

The totality of all possible outcome (or sample points) of a random experiment constitutes the **sample space** which is denoted by Ω . Sample space may be discrete or continuous.

Discrete sample space:

- Firstly, the number of possible outcomes is **finite**.

- Secondly, the number of possible outcomes is **countable infinite**, which means that there is an infinite number of possible outcomes, but the outcomes can be put in a one-to-one correspondence with the positive integers.

Example

If a die is rolled and the number that show up is noted, then $\Omega = \{1, 2, 3, ..., 6\}$.

If a die is rolled until a "6" is obtained, and the number of rolls made before getting first "6" is counted, then we have that $\Omega = \{0, 1, 2, 3, ...\}$.

Continuous sample space: If the sample space contains one or more intervals, the sample space is then **uncountable infinite**.

Example

A die is rolled until a "6" is obtained and the time needed to get this first "6" is recorded. In this case, we have that $\Omega = \{t \in \mathbb{R} : t > 0\} = (0, \infty)$.

An event is a set of outcomes of an experiment; it is a subset of the sample space.

The null set ϕ is thus an event known as the **impossible event**. The sample space Ω corresponds to the **sure event**.

In particular, every elementary outcome is an event, and so is the sample space itself.

Remarks

- An elementary outcome is sometimes called a **simple event**, whereas a **compound** event is made up of at least two elementary outcomes.
- To be precise we should distinguish between the elementary outcome *w*, which is an element of Ω and the elementary event $\{w\} \subset \Omega$.
- The events are denoted by capital letters such as *A*, *B*, *C* and so on.

Example

Consider the experiment that consists in rolling a die and recording the number that shows up. Let A be the event "the even number is shown" and B be the event "the odd number less than 5 is shown". Define the events A and B.

Solution

We have the sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$.

$\boldsymbol{A} = \{2, 4, 6\}$ and $\boldsymbol{B} = \{1, 3\}$

Definitions

- Two or more events which have an equal probability (same chance) of occurrence are said to be **equally likely**, i.e. if on taking into account all the conditions, there should be no reason to except any one of the events in preference over the others. Equally likely events may be simple or compound events.
- Two events, *A* and *B* are said to be **incompatible** (or **mutually exclusive**) if their intersection is empty. We then write that $A \cap B = \varnothing$.
- Two events, *A* and *B* are said to be **exhaustive** if they satisfy the condition $A \cup B = \Omega$.
- An event is said to be impossible if it cannot occur.

Example

Consider the experiment that consists in rolling a die and recording the number that shows up.

We have that $\Omega = \{1, 2, 3, 4, 5, 6\}$.

We define the events

$A=\{1,2,4\}$, $B=\{2,4,6\}$, $C=\{3,5\}$, $D=\{1,2,3,4\}$ and $E=\{3,4,5,6\}$

We have

$A\cup B=\{1,2,4,6\}$,

$A\cap B=\{2,4\}$,

$A\cap C=\varnothing$ and

$D\cup E=\Omega$.

Therefore, *A* and *C* are incompatible events.

D and *E* are exhaustive events.

Moreover, we may write that $A'=\{3,5,6\}$, where the symbol A' denotes the **complement** of the event *A*.

This suggests the following definition:

If *E* is an event, then E' is the event which occurs when *E* does not occur. Events *E* and E' are said to be **complementary events.**

Example of event and sample spaces

-Tossing a coin: there are two possible outcomes, you gain Head up or Tail up. Then, $\Omega=\{H,T\}$ - throwing a die and noting the number of its uppermost face. There are 6 possible outcomes: one number from 1 to 6 can be up. Then, $\Omega=\{1,2,3,4,5,6\}$.

- Two coins are thrown simultaneously. $\Omega=\{HH,HT,TH,TT\}$

- Tree coins are thrown simultaneously. $\Omega=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}$

- Two dice are thrown simultaneously. The sample space consists of 36 points: .....

$\Omega=\{(1,1),(1,2),....\}$. Please complete other points!

Note: The determination of sample space for some events such as the one for dice thrown simultaneously requires the use of complex reasoning but it can be facilitated by different counting techniques.

Application activity 2.1

Two dice are thrown simultaneously and the sum of points is noted, determine the sample space. How many elements does it have?

2.2 Simple counting techniques

2.2.1 Use of Venn diagram

Activity 2.2.1

In a class of 30 students, 10 got A on the first test, 9 got A on a second test, and 15 did not get an A on either test. Find: the number of students who got:

a) an A on both tests;

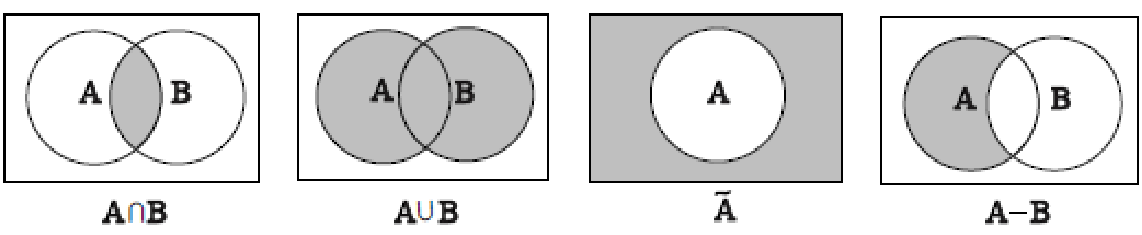
b) an A on the first test but not the second;

c) an A on the second test but not the first.

Content summary

A **Venn diagram** refers to representing mathematical or logical sets pictorially as circles or closed curves within an enclosing rectangle (the universal set), common elements of the sets represented by intersections of the circles.

In many cases, events can be described in terms of other events through the use of the standard constructions of set theory. We will briefly review definitions of these constructions. The reader is referred to the following figure for Venn diagrams which illustrate these constructions.



Let A and B be two sets. Then the union of A and B is the set $A \cup B = \{x / x \in A \, or \, x \in B\}$.

The intersection of a and b is the set $A \cap B = \{x / x \in A \, and \, x \in B\}$

The difference of A and B is the set $A-B=\{x/x\in A \, and \; x\notin B\}$.

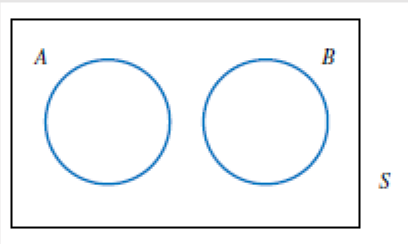
The set A is a subset of B, written $A\subset B$, if every element of A is also an element of B.

If the number of elements in a set is zero or a positive integer, we say that the set is **finite**. Otherwise, the set is said to be **infinite**. The area of mathematics that deals with the study of finite sets is called **finite mathematics**.

The number of elements in a set A is noted $n(A)$

The number of elements of the empty set is 0. $n(\varnothing)=0$

Sum rule Principle: Consider the set A and B such that $A\cap B=\phi$, A and B are **disjoint sets**.



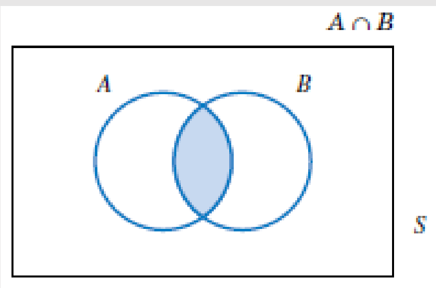
$n(A\cup B)=n(A)+n(B)$

If the two sets represent two events, the two events are said to be **mutually exclusive**. This means that they cannot occur at the same time, they do not have outcomes in common.

In this case, $n(A\cup B)=n(A)+n(B)$.

Consider the two set A and B such that $A\cap B\neq\phi$.

If A and B represent events, the two events are **not mutually exclusive**. This means that they have some outcomes in common.



In this case, $n(A\cup B)=n(A)+n(B)-n(A\cap B)$

Example: A survey of a group of children indicated there were 25 with brown eyes and 15 with black hair.

(a) If 10 children had both brown eyes and black hair, how many children interviewed had either brown eyes or black hair?

(b) If 23 children had neither brown eyes nor black hair, how many children in all were interviewed?

Solution: Let *A* denote the set of children with brown eyes and *B* the set of children with black hair. Then the data given tell us

$n(A) = 25$ and $n(b) = 15$

a) Since 10 children had both brown eyes and black hair, we know that $n(A \cap B) = 10$

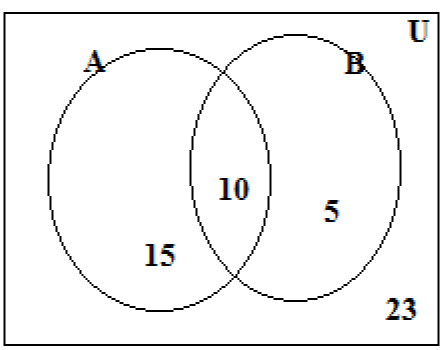
The number of children with either brown eyes or black hair is $n(A \cup B)$.

But $n(A \cup B)$ cannot be $n(A) + n(B)$ because those elements with both characteristics would then be counted twice. To obtain $n(A \cup B)$ we need to subtract those with both characteristics from $n(A) + n(B)$ to avoid counting them twice. So,

$n(A \cup B) = n(A) + n(B) - n(A \cap B)$=25 + 15 – 10 = 30.

There are 30 children with either brown eyes or black hair.

b) The sum of the number of children either in *A* or in *B* (30) and the number of children neither in *A* nor in *B* (23) is the total interviewed. The total number of children interviewed is 30 + 23 = 53.

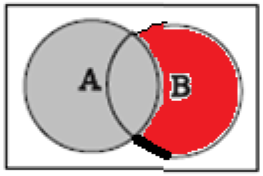


Venn diagram can be used to represent this situation.

Product Rule Principle: Let *A* × *B* be the Cartesian product of sets *A* and *B*.

Then $n(A \times B) = n(A) \times n(B)$

The complement of an event A is the set of outcomes in the sample space Ω that are not included in the outcomes of event A. The complement of A is denoted A'



Finally, the complement of A is the set $\bar{A} = \{x / x \in \Omega \text{ and } x \notin A\}$

Example:

1) Determine which events are mutually exclusive and which are not when a single die is rolled.

a) Getting an odd number and getting an even number.

b) Getting a 3 and getting an odd number.

c) Getting an odd number and getting a number less than4

d) Getting a number greater than 4 and getting a number less than 4.

Solution:

a) Events are mutually exclusive.

b) Events are not mutually exclusive.

c) Events are not mutually exclusive.

d) Events are mutually exclusive

Application activity 2.2.1

In a survey of 100 students, it was found that 77 students were studying Mathematics; 47 students were studying Physics; 44 students were studying Chemistry; 43 students were studying both Mathematics and Physics; 37 students were studying both Mathematics and Chemistry; 12 students were studying both Physics and Chemistry ; 12 students were studying all three subjects.

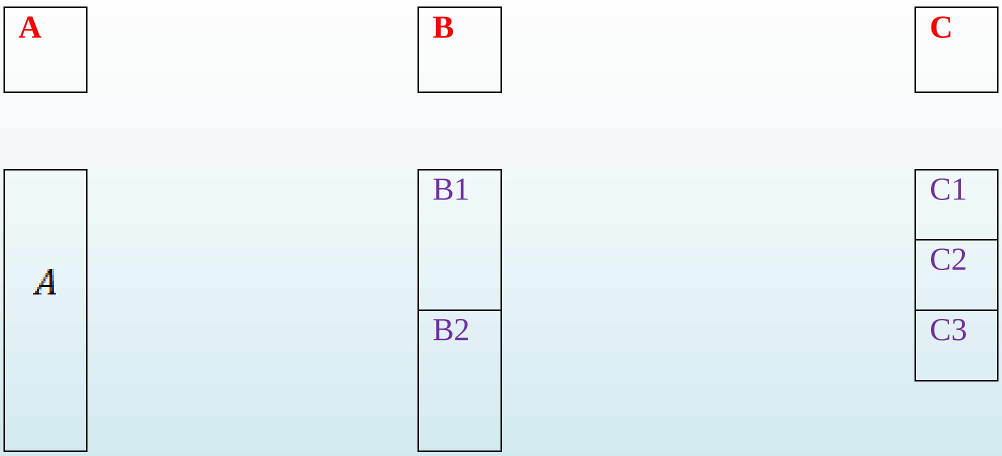
a) Find the number of students from among the 100 who were not studying any one of the three sciences.

b) Find the number of students from among the 100 who were studying both Physics and Mathematics but not Chemistry.

2.2.2 Use of tree diagrams

Activity 2.2.2

1) There are 2 roads joining A and B and 3 roads joining B and C. Write down different roads from A to C via B. How many are they?



2) Use a tree diagram to find the gender for 3 children in a family.

3) You are going to a restaurant for dinner. You can either get a crunchy or a soup. For meat, you can choose either beef, chicken, or fish.

a) Create a tree diagram and list all possible request;
b) Find the total number of all possible selections.

A **tree diagram** is simply a way of representing a sequence of events. Tree diagrams are particularly useful in probability since they record all possible outcomes in a clear and uncomplicated manner.

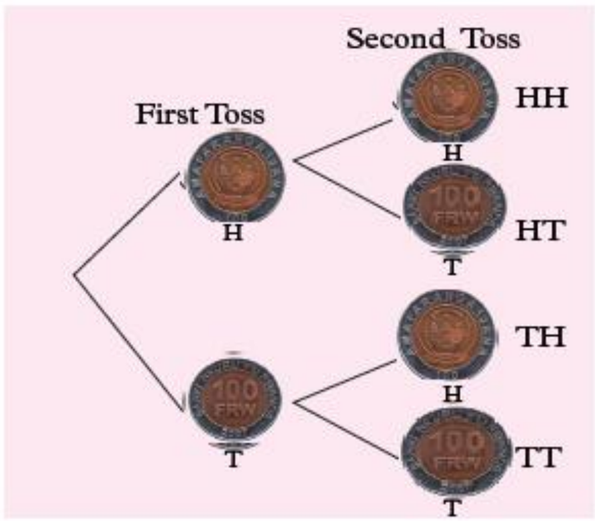
It has branches and sub-branches which help us to see the sequence of events and all the possible outcomes at each stage.

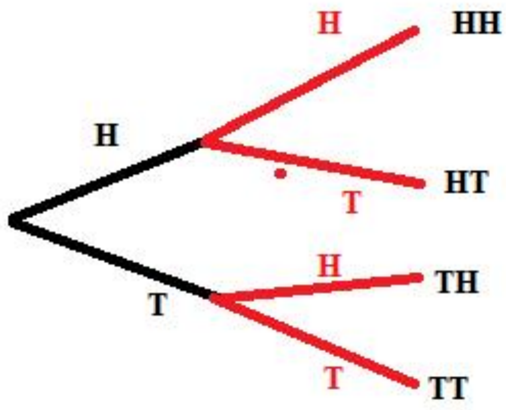
Example:

1. Using a tree diagram, determine all the possible outcomes that can be obtained when a coin is tossed twice.

***Solution**:*

In the first toss, we get either a head (H) or a tail (T). On getting a H in the first toss, we can get a H or T in the second toss. Likewise, after getting a T in the first toss, we can get a H or T in the second toss.

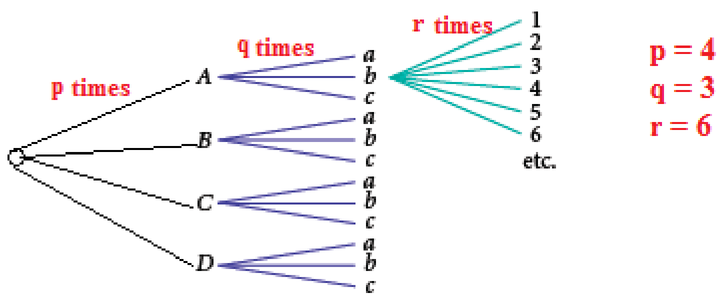




$S = \{HH, HT, TH, TT\}$

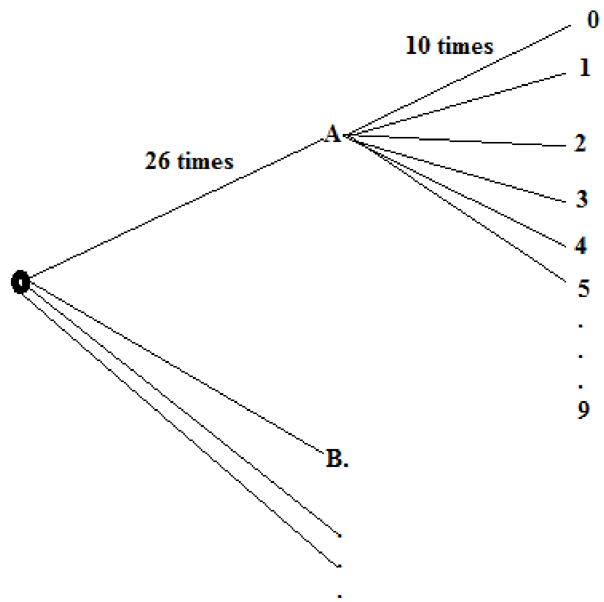
We see that the number of possible outcomes is 4. There 2 selections of the first choice and 2 selections of the second. This is 2 x 2 = 4.

Multiplication principle: If a task consists of a sequence of choices in which there are *p* selections for the first choice, *q* selections for the second choice, *r* selections for the third choice, and so on, then the task of making these selections can be done in $n = p \times q \times r$.



Example 1: How many two-symbol code words can be formed if the first symbol is a letter (uppercase) and the second symbol is a digit?

Solution: It sometimes helps to begin by listing some of the possibilities. The code consists of a letter (uppercase) followed by a digit, so some possibilities are A0, A1, A2, ….B0, B1, B2, B3,….. X9.



The task consists of making two selections.

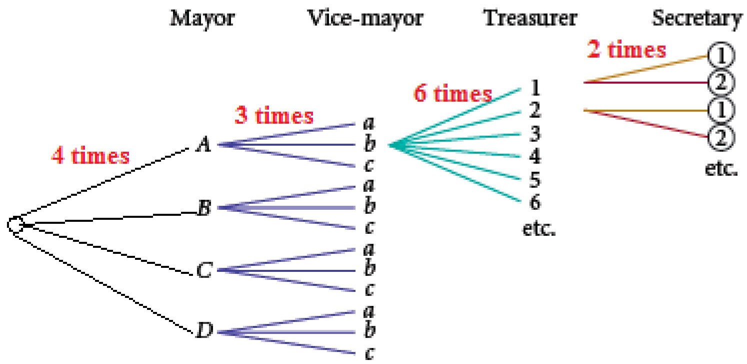
The first selection requires choosing an uppercase letter (26 choices) and the second task requires choosing a digit (10 choices).

By the Multiplication Principle, there are:

$n = 26 \times 10 = 260$ different code words of the type described.

Example 2: In a city election there are four candidates for mayor, three candidates for vice-mayor, six candidates for treasurer, and two for secretary. In how many ways can these four offices be filled?

Solution: The task of filling an office can be divided into four choices:

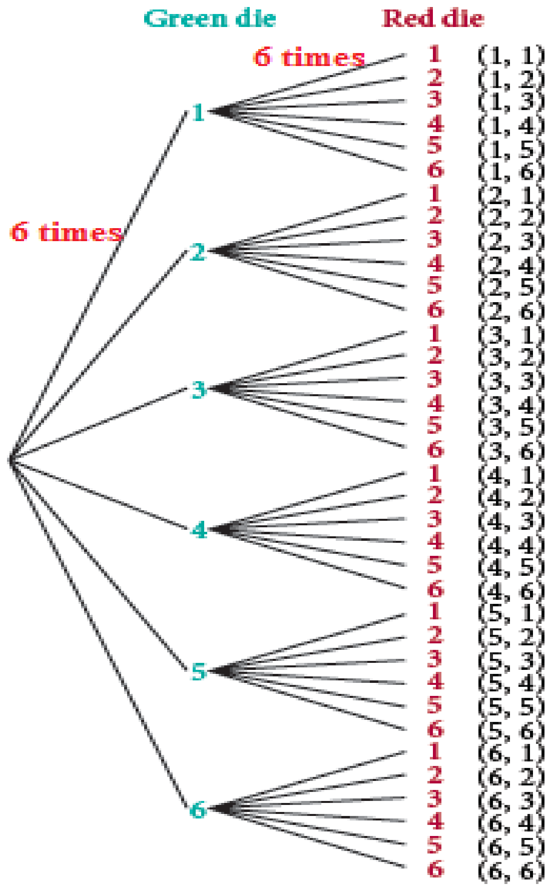


Corresponding to each of the four possible mayors, there are three vice-mayors. These two offices can be filled in 4x 3=12 different ways. Also, corresponding to each of these 12 possibilities, we have six different choices for treasurer, giving 21 x 6 =72 different possibilities. Finally, to each of these 72 possibilities there correspond two choices for secretary. In all, these offices can be filled in 4 x 3 x6 x2 =144 different ways.

Example 3:

Consider an experiment in which two dice are rolled, one green and the other red. The set of outcomes consists of all the different ways that the dice may come to rest. Find the number of all possible outcomes. It is called the cardinal of the sample space for this experiment.

Solution: Since there are 6 possible ways for the green die to come up and 6 ways for the red die to come up, the number of outcomes of rolling the two dice is 6 x 6 = 36.



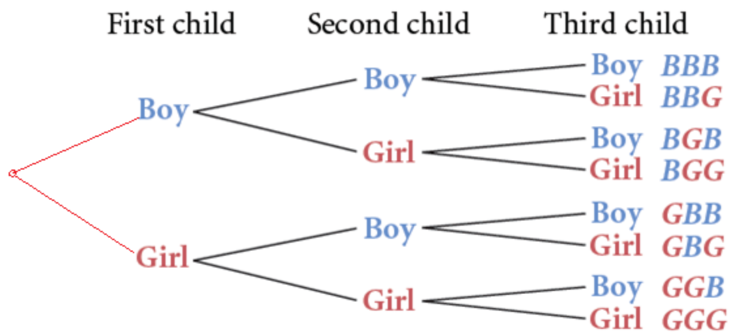
We use a tree diagram to obtain a list of the 36 outcomes.

Example 4: Find the cardinal of a sample space for the experiment of selecting one family from the set of all possible three-child families.

Solution:

One way of describing a sample space for this experiment is to list the number of girls in the family. The only possibilities are members of the set $\{0,1,2,3\}$.

That is, a three-child family can have 0, 1, 2, or 3 girls. This sample space has four outcomes. Another way of describing a sample space for this experiment is to denote *B* as “boy” and *G* as “girl.” Then the sample space can be given as {*BBB*, *BBG*, *BGB*, *BGG*, *GBB*, *GBG*, *GGB*, *GGG*} where *BBB* means first child is a boy, second child is a boy, third child is a boy, and so on. This sample space has 2 x 2 x 2 = 8 outcomes.



Application activity 2.1.2

1. Angelus has asked his girlfriend to make all the decisions for their date on her birthday. She will pick a restaurant and an activity for the date. Andy will choose a gift for her. The local restaurants include Mexican, Chinese, Seafood, and Italian. The activities she can choose from are swimming, bowling, and movies. Andy will buy her either candy or flowers.

   a) Draw a tree diagram to illustrate the choices

   b) How many outcomes are there for these three decisions?

2. A designer has 3 fabric colours he/she may use for a dress: red, green, and blue. Four different patterns are available for the dress. If each dress design requires one color

2.2.3 Use of a table

Activity 2.2.3

James has to go to a party, and he has three different shirts and two different jeans to choose from. The shirts are red, black and white in color while the jeans are blue and green in color and James can wear any jeans with his shirt.

Complete the following table to show all possible outfits can James choose from?

Jeans / **Shirts**	**Blue jeans**	**Green jeans**
Red shirt		
Black shirt		
White shirt		

A **table** is simply a way of representing a sequence of events. It is a rectangular array in which the fist column has elements of the first set while the first low has elements of the second set to be associated with the first.

Product Rule Principle: Let $A \times B$ be the Cartesian product of sets A and B.

Then $n(A\times B)=n(A)\times n(B)$

Example:

1) Find the cardinal for the sample space of "rolling two dice".

Solution:

As each die cam land in 6 different ways, and two dice are rolled, the sample space can be presented as a rectangular array.

Die 2 / Die 1	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Thus the total number of all possible outcomes while rolling two dice is 6 x 6 = 36.

There is a technique of counting without necessarily listing the total number of all possible outcomes.

This is known as **Basic product principle of counting:**

If a sequence of n events in which the first one has n_1 possibilities, the second with n_2 possibilities the third with n_3 possibilities, and so forth until n_k, the total number of possibilities of the sequence will be: Total number $= n_1 \times .n_2 \times .n_3 ... \times \mathrm{n_k}$.

Examples:

1) A car license plate is to contain three letters of the alphabet, the first of which must be R, S, T or U, followed by three decimal digits. How many different license plates are possible?

Solution

The first letter can be chosen in 4 different ways, the second and third letters in 26 different ways each, and each of the three digits can be chosen in ten ways.

By using basic product principle of counting, we get that there are $4 \times 26 \times 26 \times 10 \times 10 \times 10 = 2\,704\,000$ plates possible.

2) a) How many numbers of four different digits can be formed?

b) How many of these are even?

Solution

a) There are nine ways to choose the first digit since 0 cannot be the first digit, and nine, eight and seven ways to choose the next three digits since no digit may be repeated.

   Therefore there are $9\times9\times8\times7=4536$ numbers possible

b) The last digit must be a $0, 2, 4, 6$ or 8. There are five ways of choosing it. Then the first digit can be chosen in eight different ways since it cannot be a zero or the number chosen for the last digit. The other two digits can be chosen in eight and seven ways respectively.

   Therefore the number of even numbers is $8\times8\times7\times5=2240$.

Application activity 2.2.3

Two standard dice are rolled and their face values multiplied. How many ways are there such that the product is a prime or ends in 6?

2.3 Permutations of n unlike objects in a row

Activity 2.2

I have a suitcase with a number lock, labelled with 3 first digits from 0 to 2. The lock can be opened if 3 specific digits are set in a particular sequence. Unfortunately, I forgot that specific sequence. In how many ways can I try to unlock my suitcase?

Content summary

If I forgot the specific sequence to unlock this suitcase, I will immediately start listing all possible sequences of 3 digits taken from digits $\{0,1,2\}$ and reversed order of these digits is also included such as $\{012, 210, 021, 120, \cdots\}$.

To get the number of ways in which some or all digits (in this case), at a time are arranged, we note that the first digit to be written down can be chosen in 3 ways; the second digit can be chosen in 2 ways because there are 2 remaining digits to be written down and then the third digit can be chosen in 1 way because it is only one digit remain to be written down. Thus, the three operations can be performed in $3\times2\times1=6$ ways.

Example:

Give all different ways three students: Cauchy, Emmanuel and Alexis can be sit on the same bench. Two ways were given in the table, complete others.

Cauchy	Emmanuel	Alexis
Cauchy	Alexis	Emmanuel
..	..	..

This suggests the following fact:

The number of different permutations of *n* different objects (unlike objects) in a row is

$$n\times(n-1)\times(n-2)\times(n-3)\times...\times2\times1$$

This corresponds to the number of ways of arranging n unlike objects in a line. It is a **permutation or** ordered arrangement of **n *different* objects chosen from *n* objects**.

A useful short hand of writing this operation is $n!$ (read ***n*** **factorial)**. Then, $n!=n\times(n-1)\times(n-2)\times(n-3)\times...\times2\times1$

Thus, $1!=1$, $2!=2\times1=2$, $3!=3\times2\times1=6$, $4!=4\times3\times2\times1=24$, $5!=5\times4\times3\times2\times1=120$ and so on.

Note that $0!=1$

Example:

1) $\frac{6!}{2!4!}=\frac{6\times5\times4\times3\times2\times1}{2\times1\times4\times3\times2\times1}=15$

2) $\frac{7!}{4!2!}=\frac{7\times6\times5\times4\times3\times2\times1}{4\times3\times2\times1\times2\times1}=105$

3. Five children have to be seated on a bench. In how many ways they can be seated? How many arrangements are they, if the youngest child has to sit at the left end of the bench?

Solution

Since there are five children, the first child can be chosen in 5 ways, the next child in 4 ways, the next in 3 ways, the next in 2 ways and the last in 1 way. Then, the number of ways is $5! = 5\times4\times3\times2\times1 = 120$.

Now, if the youngest child has to sit at the left end of the bench, this place can be filled in only 1 way. The next child can then be chosen in 4 ways, the next in 3 ways and so on. Thus, the number of total arrangement is $1\times4! = 1\times4\times3\times2\times1 = 24$.

4. What is the number of ways of sitting n_1 men and n_2 women on a bench if men sitt together and women sit together?

Solution: Note that the number of ways of arranging n objects in a line is $n! = 1\times2\times3\times4\times5\times...\times n$.

Given that the total number $n = n_1 + n_2$ of people made by n_1 men and n_2 women, the total number of ways of sitting (arranging) n_1 men and n_2 women on a bench such that men sit together and women sit together considering that one group can sit at the left side or the right side of the other is $2\times n_1!\times n_2!$.

5. Three different mathematics books and five other books are to be arranged on a bookshelf. Find :

a) the number of possible arrangements of the book.

b) the number of possible arrangements if the three mathematics books must be kept together?

Solution:

We have 8 books altogether.

a) Since we have 8 books altogether, the first book can be chosen in 8 ways, the next in 7 ways, the next in 6 ways and so on. Thus, the total arrangement is $8! = 40320$

b) Since the 3 mathematics books have to be together, consider these bound together as one book. There are now 6 books to be arranged and these can be performed in $6! = 720$.

Note that we have taken the three mathematics book as one book; these three books can be arranged in $3! = 6$ ways. Thus, the total number of arrangements is $720 \times 6 = 4320$.

Application activity 2.3

1. Simplify

   a. $\frac{5!}{2!}$ b. $\frac{10!}{6!7!}$

2. Four different English books, five different Biology books and ten other books are to be arranged on a bookshelf. Find

   a. The number of possible arrangements of the books.

   b. The number of possible arrangements if the three Biology books must be kept together?

2.4 Permutations of indistinguishable objects: Permutations with repetitions of n objects chosen from *n* objects

Activity 2.4

1) Make a list of all arrangements formed by 4 numbers: 1,2,3,4. How many arrangements are they possible?

2) Consider the arrangements of four letters in the word “BOOM”.

- Write down all possible different arrangements.
- How many arrangements are they possible of four letters in the word “BOOM”?

Content summary

A **permutation** is an ordered arrangement of *r* objects chosen from *n* objects.

Consider the arrangements of six letters in the word “**AVATAR**” (a title used for the movie).

We see that there are three **A**'s (or 3 alike letters).

- Let the three **A**'s in the word be distinguished as $\boldsymbol{A_1}$, $\boldsymbol{A_2}$ and $\boldsymbol{A_3}$ respectively. Then all the six letters are different, so the number of permutations of them (called labeled permutations) is $n! = 6!$.

- However, consider each of the real permutations without distinguishing the three **A**'s, for example **W=RATAVA**.

- The following are all of the 6 (=3!) labeled permutations among the 6! ones, which come from permuting the three labeled **A**'s in **W=RATAVA**:

$\mathbf{RA_1TA_2VA_3, RA_1TA_3VA_2, RA_2TA_1VA_3, RA_2TA_3VA_1, RA_3TA_1VA_2, RA_3TA_2VA_1}$

- All these six labeled permutations should be considered as an identical real permutation, which is **W=RATAVA**.

- Since each real permutation has six of such labeled permutations coming from the three **A**'s, we conclude that the desired number of real permutations is just $\frac{6!}{3!} = \frac{6\times5\times4\times3!}{3!} = 6\times5\times4 = 120$

This suggests the following fact:

The number of different permutations of n ***indistinguishable*** *objects with* n_1 *alike,* n_2 *alike, ..., is* $\frac{n!}{n_1!\times n_2!\times\ldots}$.

This corresponds to the number of arranging all n objects in line of which n_1 of one type are alike, n_2 of the second type are alike and so on.

Note: Alike means that the objects in a group are indistinguishable from one another.

Example:

1. How many distinguishable six digit numbers can be formed from the digits 5, 4, 8, 5, 5, 4?

Solution

There are 6 letters in total with three 5’s and two 4’s. Then the required numbers are $\frac{6!}{3!2!}=\frac{720}{12}=60$

2. How many arrangements can be made from the letters of the word **TERRITORY**?

Solution

There are 9 letters in total with three **R**’s and two **T**’s.

Thus, we have $\frac{9!}{3!\times 2!}=\frac{9\times 8\times 7\times 6\times 5\times 4}{2\times 1}=\frac{60,480}{2}=30,240\, arrangements$.

3. In how many different ways can 4 identical red balls, 3 identical green balls and a yellow ball be arranged in a row?

Solution

There are 8 balls in total with 4 red, 3 green and one yellow.

Thus, we have $\frac{8!}{4!\times 3!}=\frac{8\times 7\times 6\times 5}{3\times 2\times 1}=280\, ways$.

Application activity 2.4

1. How many different arrangements can be made from the letters of the word
   a) **ENGLISH** b) MATHEMATICS
2. How many arrangements can be made from the letters of English alphabet?
3. In how many ways can 4 red, 3 yellow and 2 green discs be arranged in a row, if discs of the same colour are indistinguishable?

2.5 Circular permutation of all n objects

Activity 2.5

Take 5 different note books

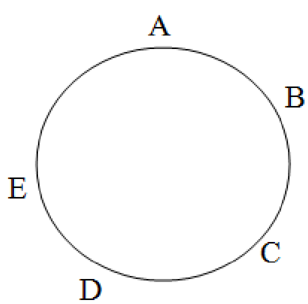
- Put them on a circular table
- Fix one note book, for example A;
- Try to interchange other 4 note books as possible
- How many different ways obtained?

Remember that there is one note book that will not change its place.

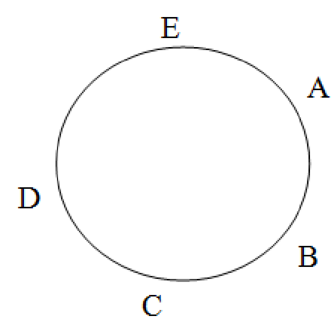
Content summary

We have seen that if we wish to arrange n different things in a row, we have $n!$ possible arrangements. Suppose that we wish to arrange n things around a circular table. The number of possible arrangements will no longer be $n!$ because there is now no distinction between certain arrangements that were distinct when written in a row.

For example ABCDE is different arrangement from EABCD, but



is not a different arrangement from



With circular arrangement of this type, it is the relative positions of the items being arranged which is important. One item can therefore be fixed and the remaining items arranged around it.

The number of arrangements of n unlike things in a circle will therefore be $(n-1)!$. *In those cases where clockwise and anticlockwise arrangements are not considered to be different, this reduces to* $\frac{1}{2}(n-1)!$.

Example:

1. Four men Peter, Rogers, Smith and Thomas are to be seated at a circular table. In how many ways can this be done?

Solution

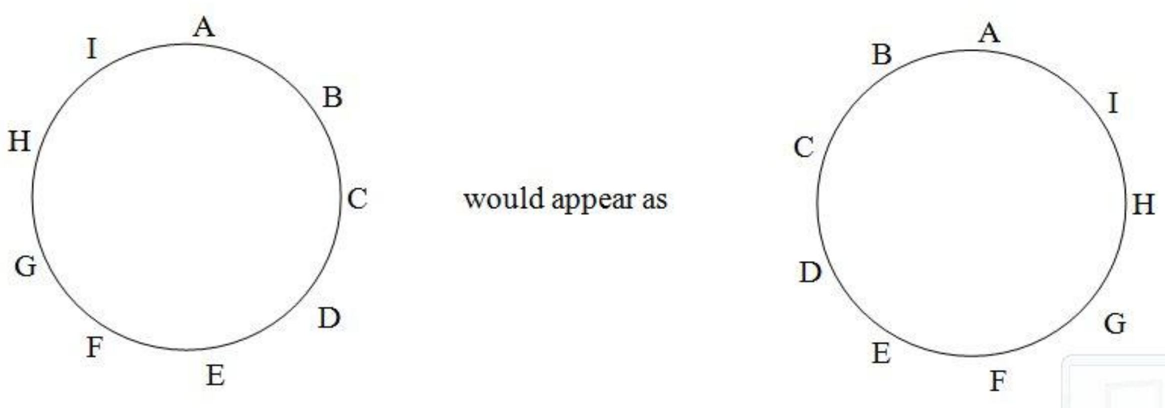
Suppose Peter is seated at some particular place. The seats on his left can be filled in 3 ways, the next seat can then be filled in 2 ways and the remaining seat in 1 way.

Thus, total number of arrangements is $3! = 6$.

2. Nine beads, all of different colors are to be arranged on a circular wire. Two arrangements are not considered to be different if they appear the same when the ring is turned over. How many different possible arrangements are there?

Solution

When the ring is turned over, the arrangements



When viewed from one side, these arrangements are only different in that one is a clockwise arrangement and the other is anticlockwise. If one bead is fixed, there are $(9-1)! = 8!$ ways of arranging the remaining beads relative to the fixed one.

But, half of these arrangements will appear the same as the other half when the ring is turned over, because for every clockwise arrangement there is a similar anticlockwise arrangement.

Hence, number of arrangements is $\frac{1}{2}8! = 20160$.

Application activity 2.5

1. Five men Eric, Peter, John, Smith and Thomas are to be seated at a circular table. In how many ways can this be done?
2. Eleven different books are placed on a circular table. In how many ways can this be done?

2.6 Basic sum principle of counting for mutually exclusive situations

Activity 2.6

1) Suppose that you go to a restaurant and you are allowed a soup or juice. Will you pick one, the other or both?

2) How many different four digits numbers, end in a 3 or a 4, can be formed from the figures 3,4,5,6 if each figure is used only once in each number.

Content summary

Two experiments 1 and 2 are mutually exclusive, if when experiment 1 occurs, experiment 2 cannot occur. Likewise, if experiment 2 occurs, experiment 1 cannot occur.

Basic sum principle of counting

In such cases, the number of permutations of either experiment 1 or experiment 2 occuring can be obtained by adding the number of permutations of experiment 1 to the number of permutations of experiment 2.

This suggests the following result:

“*If* the first e*xperiment has* **m** *possible outcomes and if* the second *experiment has* **n** *possible outcomes, then an experiment which might be* "the first *experiment or* the second *experiment*", *called* **experiment 1 or 2**, *has* $(m+n)$ *possible outcomes.*”

Example:

1. In tossing an object which might be a coin (with two sides H and T) or a die (with six sides 1 through 6), how many possible outcomes will appear?

Solution

- The experiment may be tossing a coin (experiment 1) or tossing a die (experiment 2), or just experiment 1 or 2.
- So the number of outcomes is $2+6=8$ according to the above basic sum principle of counting.

The number of permutations in which a certain experiment 1 occurs will clearly be mutually exclusive with those permutations in which that experiment does not occur. Thus,

$$\begin{aligned}&\textit{Number of permutations in which experiment 1 does not occur}\\&= \textit{total number of permutation - number of permutations in which experiment 1 occurs}\end{aligned}$$

2. In how many ways can five people Abijuru, Bwenge, Cyusa, Keza, and Manzi, be arranged around a circular table in each of following cases:

a) Abijuru must sit next Bwenge?

b) Abijuru must not sit next Bwenge?

Solution

There are five people.

a) Since Abijuru and Bwenge must sit next to each other, consider these two bound together as one person. There are now, 4 people to seat. Fix on of these, and then the remaining 3 people can be seated in $3\times2\times1=6$ ways relative to the one who was fixed.

In each of these arrangements Abijuru and Bwenge are seated together in a particular way. Abijuru and Bwenge could now change the seats giving another 6 ways of arranging the five people. Total number of arrangements is $6\times 2=12$.

b) If Abijuru must not sit next Bwenge, then this situation is a mutually exclusive with the situation in a).

Total number of arrangements of 5 people at a circular table is $(5-1)!=4!=24$.

Thus, if Abijuru must not sit next Bwenge, the number of arrangements is $24-12=12$.

Generalized sum principle of counting

"*If Experiments* 1 *through* k *have respectively* n_1 *through* n_k *outcomes*,
then the experiment 1 *or* 2 *or* … *or* k *has* $\mathrm{n}_1+n_2+...+\mathrm{n}_k$ *outcomes.*

Example:
How many even numbers containing one or more digits can be formed from the digits 2, 3, 4, 5, 6 if no digit may be repeated?

Solution

Since the required numbers are even, last digit must 2 or 4 or 6. Note that there are 5 digits.

So we can form one digit, two digits, three digits, four digits or five digits as follow

One digit: 2 or 4 or 6. That is 3 numbers

Two digits: 3 ways to choose the last and 4 ways to choose the first. That is $3\times 4=12$ numbers.

Three digits: 3 ways to choose the last, 4 ways to choose the first and 3 ways to choose the second. That is $3\times 4\times 3=36$

Four digits: 3 ways to choose the last, 4 ways to choose the first, 3 ways to choose the second and 2 ways to choose the fourth. That is $3\times 4\times 3\times 2=72$

Five digits: 3 ways to choose the last, 4 ways to choose the first, 3 ways to choose the second, 2 ways to choose the fourth and 1 way to choose the fifth. That is $3\times 4\times 3\times 2\times 1=72$

Adding we have $3+12+36+72+72=195$ even numbers in total.

Application activity 2.6

A palindrome number is a number that is the same when written forwards or backwards (for example 11, 121,999 are palindrome numbers). How many palindrome numbers less than 1000 are there?

2.7 Distinguishable arrangements (Permutations of *r* unlike objects chosen from *n* distinct objects without repetition)

Activity 2.7

Make a selection of any three letters from the word "*PRODUCT*" and fill them in 3 empty spaces

Use a box like this for empty spaces		

Write down all different possible permutations of 3 letters selected from the letters of the word "*PRODUCT*". How many are they?

Content summary

It is the same as the permutations of r **unlike** objects selected from n distinct objects. Consider the number of ways of placing 3 of the letters A, B, C, D, E, F, G in 3 empty spaces.

The first space can be filled in 7 ways, the second in 6 ways and the third in 5 ways. Therefore there are $7\times6\times5$ ways of arranging 3 letters taken from 7 letters. This is the number of permutations of 3 objects taken from 7 and it is written ${}^7P_3$.

So ${}^7P_3 = 7\times6\times5 = 210$.

Note that the order in which the letters are arranged is important: ABC is a different permutation from ACB.

Now, $7\times6\times5$ could be written $\dfrac{7\times6\times5\times4\times3\times2\times1}{4\times3\times2\times1}$

i.e $^{7}P_{3}=\dfrac{7\times6\times5\times4\times3\times2\times1}{4\times3\times2\times1}=\dfrac{7!}{4!}=\dfrac{7!}{(7-3)!}$

A **permutation** is an ordered arrangement of *r* objects chosen from *n* objects. If there is no repetition, this suggests the following fact:

The number of arrangements of *n* objects using $\boldsymbol{r \le n}$ of them, in which

1. the *n* objects are distinct,

2. once an object is used it cannot be repeated, and

3. order is important.

The number of different permutations (ways) of r unlike objects selected from n different objects is $^{n}P_{r}=\dfrac{n!}{(n-r)!}$ *or we can use the denotation* $P_{r}^{n}=\dfrac{n!}{(n-r)!}$ *or* $P(n,r)=\dfrac{n!}{(n-r)!}$

This is the number of Permutations of *r* objects chosen from *n* distinct objects without repetition.

Note that if $r=n$, we have $^{n}P_{n}=\dfrac{n!}{(n-n)!}=\dfrac{n!}{0!}=n!$ which is the ways of arranging *n* unlike objects.

Example:

1) *How many permutations of 3 letters chosen from eight letters of the word* ***RELATION*** *are there?*

Solution

We see that all those eight letters are distinguishable (unlike). So the required arrangements are given by

$^{8}P_{3}=\dfrac{8!}{(8-3)!}=\dfrac{8!}{5!}=336$.

2) How many permutations are there of 2 letters chosen from letters A, B, C, D, E?

Solution

There are 5 letters which are distinguishable (unlike). So the required arrangements are given by

$$^5P_2 = \frac{5!}{(5-2)!} = \frac{5!}{3!} = 20 .$$

4) *How many different ways can a chairperson and an assistant chairperson be selected for a research project if there are seven scientists available?*

Solution

It is $_7P_2 = \frac{7!}{(7-2)!} = 42$

5) All we know about Shannon, Patrick, and Ryan is that they have different birthdays.
If we listed all the possible ways this could occur, how many would there be?
Assume that there are 365 days in a year.

Solution: This is an example of a permutation in which 3 birthdays are selected from a possible 365 days, and no birthday may repeat itself.

The number of ways that this can occur is

$$\frac{365!}{(365-3)!} = 365 \times 364 \times 363 = 48{,}228{,}180$$

Application activity 2.7

1. How many permutations are there of 4 letters chosen from letters of the word ENGLISH?
2. How many permutations are there of 5 letters chosen from letters A, B, C, D, E, F, and G.
3. How many permutations are there of 10 letters chosen from English alphabet?

2.8 Indistinguishable arrangements: Permutations of *r* objects selected from mixture of *n* alike and unlike objects

Activity 2.8

How many permutations are there of 2 letters chosen from letters of the word BLOOM?

Content summary

When determining the number of all possible **permutations of r objects selected from a mixture of n alike and unlike objects**, start by determining all possible mutually exclusive events from the given experiment that may occur and then apply basic sum principle.

Example

How many different arrangements are there of 3 letters chosen from the word COMBINATION?

Solution

There are 11 letters including two O's, two I's and two N's. to find the total number of different arrangements we consider the possible arrangements as four mutually exclusive situations.

a) Arrangements in which all 3 letters are different: there are ${}^{8}P_{3} = 336$

b) Arrangements containing two O's and one other letter: the other letter can be one of seven letters (C, M, B, I, N, A or T) and can appear in any of the three positions (before the two O'S, between the two O's, or after the two O's). i.e $3\times 7 = 21$ arrangements containing two O's and one other letter.

c) Arrangements containing two I's and one other letter: by the same reasoning in b) there will be $3\times 7 = 21$ arrangements containing two I's and one other letter.

d) Arrangements containing two N's and one other letter: by the same reasoning in b) there will be $3\times 7 = 21$ arrangements containing two N's and one other letter.

Thus the total number of arrangements of 3 letters chosen from the word COMBINATION will be $336+21+21+21 = 399$

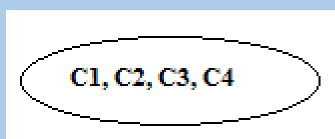
Application activity 2.8

How many permutations are there of 2 letters chosen from letters of the word PACIFIC?

2.9 Ordered Samples or product rule of counting

Activity 2.9

There are 4 red counters in a basket.



If you take 3 of them one after another where ordering matters,

a) In how many ways can you do it given that when you take the counter your colleague replaces it?

b) In how many ways can you do it given that there is no replacement?

Many problems are concerned with choosing an element from a set S, say ,with n elements. When we choose one element after another, say, r times, we call the choice **an ordered sample of size** r. We consider two cases:

(1) Sampling with replacement

Here we have a set with n elements (e.g.: A={1,2,3,⋯, n} and we want to draw ***r*** samples from the set such that ordering matters and repetition is allowed. For example, if A= {1,2,3} and ***r =2*** , there are 9 different possibilities:

1. (1,1);
2. (1,2);
3. (1,3);
4. (2,1);
5. (2,2);

6. (2,3);
7. (3,1);
8. (3,2);
9. (3,3).

In general, we can argue that there are r positions in the chosen list: (Position 1, Position 2, ..., Position r). There are n options for each position.

Thus, when ordering matters and repetition is allowed, the total number of ways to choose r objects from a set with n elements is $\underbrace{n\times n\times n\times ...\times n}_{r\ factors}=n^r$

Note that this is a special case of the multiplication principle where there are r "experiments" and each experiment has n possible outcomes.

Example: The International Airline Transportation Association (IATA) assigns three-letter codes to represent airport locations. For example, the airport code for Ft Lauderdale, Florida is FLF. Notice that repetition is allowed in forming this code. How many airport codes are possible given that there exist 26 letters?

Solution:

We are choosing 3 letters from 26 letters and arranging them in order. In the ordered arrangement a letter may be repeated. This is an example of a permutation with repetition in which 3 objects are chosen from 26 distinct objects.

The task of counting the number of such arrangements consists of making three selections. Each selection requires choosing a letter of the alphabet (26 choices).

By the Multiplication Principle, there are (26 x 26 x 26) possible airport codes.

This is: $26\times 26\times 26=26^3=17,576$.

(2) Sampling without replacement

Consider the same setting as above, but now repetition is not allowed. For example, if A={1,2,3} and r = 2, there are 6 different possibilities:

1. (1,2);
2. (1,3);
3. (2,1);
4. (2,3);

5. (3,1);
6. (3,2).

In general, we can argue that there are r positions in the chosen list: (Position 1, Position 2, ..., Position r). There are n options for the first position, (n−1) options for the second position (since one element has already been allocated to the first position and cannot be chosen here), (n−2) options for the third position, (n−k+1) options for the r^{th} position.

Thus, when ordering matters and repetition is not allowed, the total number of ways to choose r objects from a set with n elements is $n\times(n-1)\times(n-2)\times\cdots\times(n-r+1)$

Any of the chosen lists in the above setting (choose k elements, ordered and no repetition) is called a k-permutation of the elements in set A.

Example 1:

a) How many numbers of four different digits can be formed?

b) How many of these are even?

Solution

a) There are nine ways to choose the first digit since 0 cannot be the first digit, and nine, eight and seven ways to choose the next three digits since no digit may be repeated.

Therefore there are $9\times9\times8\times7=4536$ numbers possible

b) The last digit must be a $0, 2, 4, 6$ or 8. There are five ways of choosing it. Then the first digit can be chosen in eight different ways since it cannot be a zero or the number chosen for the last digit. The other two digits can be chosen in eight and seven ways respectively.

Therefore the number of even numbers is $8\times8\times7\times5=2240$.

Example 2:

A car license plate is to contain three letters of the alphabet, the first of which must be R, S, T or U, followed by three decimal digits. How many different license plates are possible?

Solution

The first letter can be chosen in 4 different ways, the second and third letters in 26 different ways each, and each of the three digits can be chosen in ten ways.

By using basic product principle of counting, we get that there are $4\times 26\times 26\times 10\times 10\times 10 = 2\,704\,000$ plates possible.

Example 3:

Three cards are chosen one after the other from a 52-card deck. Find the number m of ways this can be done:

(a) with replacement;

(b) without replacement.

Solution

(a) Each card can be chosen in 52 ways. Thus $m = (52)(52)(52) = 140608$

(b) Here there is no replacement. Thus, the first card can be chosen in 52 ways, the second in 51 ways, and the third in 50 ways.

Therefore: $m = (52,3) = 52(51)(50) = 132600$

Application activity 2.9

1. The school soccer team is ordering new knee pads for their uniforms. The knee pads come in 4 different colours, 6 sizes, and 2 styles. How many different outcomes of knee pads are available?
2. How many even numbers containing 2 digits can be formed from the digits 2, 3, 4 if no digit may be repeated?
3. There are 20 teams in the local football competition. In how many ways can the first four places in the premiership table be filled?

2.10 Combinations or combinatorics

Activity 2.10

Take 8 different Mathematics books and form different groups each containing 2 mathematics books. How many groups obtained?

Content summary

From permutation of r unlike objects selected from n different objects, we have seen that the order in which those objects are placed is important. But when considering the number of combinations of r unlike objects selected from n different objects, the order in which they are placed is not important.

For example, the one combination **ABC** gives rise to 3! permutations: **ABC, ACB, BCA, BAC, CAB, CBA**.

Consider the number of permutations of 3 letters selected from the 7 letters **A, B, C, D, E, F, G**.

That is

$$^{7}P_{3} = \frac{7!}{(7-3)!} = \frac{7!}{4!}.$$

If we need the combinations of 3 letters selected from those 7 letters, we will take this number of permutations divided by $3!$ because each permutation gives rise to 3! permutations.

That is, the number of combinations of 3 letters selected from those 7 letters is

$$\frac{^{7}P_{3}}{3!} = \frac{\frac{7!}{4!}}{3!} = \frac{7!}{4!3!} = \frac{7!}{(7-3)!3!}.$$

This number is denoted by $^{7}C_{3}$.

Thus, the number of combinations of 3 letters selected from those 7 unlike letters is $^{7}C_{3} = \frac{7!}{(7-3)!3!} = \frac{7!}{4!3!} = 35$.

A **combination** is an arrangement, without regard to order, of *r* objects selected from *n* distinct objects without repetition, where $\boldsymbol{r \le n}$. The notation $\boldsymbol{C(n,r)}$ represents the number of combinations of *n* distinct objects using *r* of them.

This suggests the following fact:

The number of arrangements of *n* objects using $\boldsymbol{r \le n}$ of them, in which

1. the *n* objects are distinct,

2. once an object is used, it cannot be repeated, and

3. order is not important,

is given by the formula $\boldsymbol{C(n,r) = \frac{n!}{(n-r)!r!}}$.

The number of different groups of r items that could be formed from a set of n distinct objects with the order of selections being ignored is

$$^{n}C_{r} = \frac{n!}{(n-r)!r!}.$$

We can write $^{n}C_{r} = \frac{^{n}P_{r}}{r!}$

$^{n}C_{r}$ is sometimes denoted by C_{r}^{n} or $_{n}C_{r}$ or $\begin{pmatrix} n \\ r \end{pmatrix}$ or $C(n,r)$.

Note that the objects selected to be in a group are regarded as indistinguishable (unlike).

Example:

1. From a group of 5 men and 7 women, how many different committees consisting of 2 men and 3 women can be formed?

Solution

- Experiment 1: select 2 men from 5.

- Number of possible outcomes of experiment 1 is ${}^{5}C_2 = \frac{5!}{2!3!} = \frac{5\times4\times3!}{2\times3!} = 10$

- Experiment 2: select 3 women from 7.

- Number of possible outcomes of experiment 2 is ${}^{7}C_3 = \frac{7!}{3!4!} = \frac{7\times6\times5\times4!}{6\times4!} = 35$

- Experiment of forming a committee: experiment 1 & 2.

- Number of possible outcomes of experiment 1 and 2 is ${}^{5}C_2 \times {}^{7}C_3 = 10\times35 = 350$ by the basic product principle of counting

  That is, the desired number of possible outcomes of the experiment of forming a committee is 350.

2. *A committee of three men and one woman is obtained from five men and three women. In how many ways can the members be chosen?*

Solution

Three men can be selected from five men, i.e ${}^{5}C_3 = \frac{5!}{(5-3)!3!} = \frac{5!}{2!3!}$ ways

One woman can be selected from three women, i.e ${}^{3}C_1 = \frac{3!}{(3-1)!1!} = \frac{3!}{2!1!}$ ways

By the basic product principle of counting, there are ${}^{5}C_3 \times {}^{3}C_1 = \frac{5!}{2!3!} \times \frac{3!}{2!1!} = \frac{5!}{2!2!} = 30$ ways of selecting the committee.

The following two identities are true:

a) ${}^{n}C_r = {}^{n}C_{n-r}$

b) Pascal's identity: ${}^{n+1}C_r = {}^{n}C_r + {}^{n}C_{r-1}$

3. How many different committees of 3 people can be formed from a pool of 7 people.

The 7 people are distinct. More important, though, is the observation that the order of being selected for a committee is not significant. The problem asks for the number of combinations of 7 objects taken 3 at a time.

$$\boldsymbol{C}(7,3)=\frac{7!}{4!3!}=35$$

4. In how many ways can a committee consisting of 2 faculty members and 3 students be formed if 6 faculty members and 10 students are eligible to serve on the committee?

Solution:

The problem can be separated into two parts: the number of ways that the faculty members can be chosen $\boldsymbol{C}(6,2)$, and the number of ways that the student members can be chosen $\boldsymbol{C}(10,3)$, By the Multiplication Principle, the committee can be formed in $\boldsymbol{C}(6,2)$ x $\boldsymbol{C}(10,3)$=1800 ways.

4. A farmer buys 3 cows, 2 pigs, and 4 hens from a man who has 6 cows, 5 pigs, and 8 hens. Find the number *m* of choices that the farmer has.

Solution: The farmer can choose the cows in *C(*6, 3*)* ways, the pigs in *C(*5, 2*)* ways, and the hens in *C(*8, 4*)* ways. Thus the number *m* of choices follows:

$$\boldsymbol{C}(6,3)\times\boldsymbol{C}(5,2)\times\boldsymbol{C}(8,4)$$

5. A class contains 10 students with 6 men and 4 women. Find the number *n* of ways to:

Solution: The 2 men can be chosen from the 6 men in *C*(6, 2) ways, and the 2 women can be chosen from the 4 women in *C*(4, 2) ways. Thus, by the Product Rule: $\boldsymbol{n}=\boldsymbol{C}(6,2)\times\boldsymbol{C}(4,2)=90$

Application activity 2.10

1. A committee of four men and two women is obtained from 10 men and 12 women. In how many ways can the members be chosen?
2. A group containing 4 Mathematics books and 5 Physics books is formed from 9 Mathematics books and 10 Physics books. How many groups can be formed?

2.11 Combination with repetition

Activity 2.11

There are five colored balls in a pool. All balls are of different colors. If the order in which the balls can be selected does not matter and the selection of balls can be repeated, in how many ways can we choose four pool balls?

Let suppose there are p elements in a set A. We are asked to select q elements from this set, given that each element can be selected multiple times. This is known as a **combination with repetition**. For instance, we can make combinations of three elements of the set {p, q, r, s} in this way:

ppp, ppq, ppr, pps, pqq, pqr, pqs, prr, prs, pss, qqq, qqr, qqs, qrr, qrs, qss, rrr, rrs,rss, sss

You can see that most of the alphabets are repeated more than once.

Let us consider another example:

Three flavors of ice-cream are available in an ice-cream cafe. These flavors are chocolate, vanilla, and pineapple. A person can have only two scoops of ice cream. What will be the variations in this case?

Well, if the person can select two scoops at a time, then he can have one flavor two times. In this case, the examples of variations can be: chocolate, chocolate, vanilla chocolate, chocolate pineapple, etc.

The order does not matter, and flavors can be repeated. This is a combination **with Repetitions.**

In this case, if n = total number of elements in a set,

r = number of elements that can be selected from a set

The number of combinations when repetition is allowed is

$$C(n,r)=\frac{(r+n-1)!}{r!(n-1)!}$$

Example 1

There are five colored balls in a pool. All balls are of different colors. In how many ways can we choose four pool balls? The order in which the balls can be selected does not matter in this case. The selection of balls can be repeated.

Solution

The order in which the balls can be selected does not matter in this case. The selection of balls can be repeated.

Total number of balls in the pool= n = 5

The number of balls to be selected = r = 4

The number of arrangements in which the four pool balls can be chosen is

$$\boldsymbol{C}(\boldsymbol{n},\boldsymbol{r})=\frac{(\boldsymbol{r}+\boldsymbol{n}-1)!}{\boldsymbol{r}!(\boldsymbol{n}-1)!}=\frac{(4+5-1)!}{4!(5-1)!}=70$$

The pool balls can be selected in 70 ways.

Example 2

There are eight different ice-cream flavors in the ice-cream shop. In how many ways can we choose five flavors out of these eight flavors?

Solution

The order in which the flavors can be selected does not matter in this case. One ice-cream flavor can be selected multiple times.

Total number of ice-cream flavors = n = 8

The number of ice-cream flavors to be selected = r = 5

Use the following formula to get the number of arrangements in which the five ice-cream flavors can be chosen.

$$\boldsymbol{C}(\boldsymbol{n},\boldsymbol{r})=\frac{(\boldsymbol{r}+\boldsymbol{n}-1)!}{\boldsymbol{r}!(\boldsymbol{n}-1)!}$$

$$\boldsymbol{C}(8,5)=\frac{(5+8-1)!}{5!(8-1)!}=792$$

The ice cream flavors can be selected in 792 ways.

Example 3

Harry has six different colored shirts. In how many ways can he hang the four shirts in the cupboard?

Solution

The order in which the shirts can be selected does not matter in this case. The shirts can be repeated.

Total number of shirts = n = 6

The number of shirts to be selected = r = 4

The number of arrangements in which the four shirts can be chosen.

$$\boldsymbol{C}(\boldsymbol{n},\boldsymbol{r})=\frac{(\boldsymbol{r}+\boldsymbol{n}-1)!}{\boldsymbol{r}!(\boldsymbol{n}-1)!}$$

$$\boldsymbol{C}(6,4)=\frac{(4+6-1)!}{4!(6-1)!}=126$$

The shirts can be displayed in 126 ways.

Example 4

Alice has seven different chocolates. How many ways can five chocolates be selected?

Solution

The order in which the chocolates can be selected does not matter in this case. The flavors can be repeated.

Total number of chocolates = n = 7

The number of chocolates to be selected = r = 5

The number of arrangements in which the four shirts can be chosen:

$$\boldsymbol{C}(\boldsymbol{n},\boldsymbol{r})=\frac{(\boldsymbol{r}+\boldsymbol{n}-1)!}{\boldsymbol{r}!(\boldsymbol{n}-1)!}$$

$$\boldsymbol{C}(7,5)=\frac{(5+7-1)!}{5!(7-1)!}=462$$

The chocolates can be selected in 462 ways

Example 5

Sam has five colored pencils. In how many ways can he select three pencils?

Solution

The order in which the pencils can be selected does not matter in this case. The pencils can be repeated.

Total number of pencils = n = 5

The number of pencils to be selected = r = 3

The number of arrangements in which the five pencils can be chosen.

$$C(n,r)=\frac{(r+n-1)!}{r!(n-1)!}$$

$$C(5,3)=\frac{(3+5-1)!}{3!(5-1)!}=35$$

The pencils can be selected in 35 different ways.

Example 6

Mariah has ten different candies. How many ways can six candies be selected?

Solution

The order in which the candies can be selected does not matter in this case. The candies can be repeated.

Total number of candies = n = 10

The number of candies to be selected = r = 6

The number of arrangements in which the six candies can be chosen:

$$C(n,r)=\frac{(r+n-1)!}{r!(n-1)!}$$

$$C(10,6)=\frac{(6+10-1)!}{6!(10-1)!}=5005$$

The candies can be selected in 5005 different ways.

Application activity 2.11

1) A person is going to a candy shop where there are 8 types of flavours. If this person is only going to buy 3, define the number of all possible combinations.
2) A sportsman goes to the store to buy 4 pairs of shoes, if at the store there are a lot of shoes in 5 available colors, how many combination of colors can this man buy.

2.3 End of unit assessment

There are 3 men and 4 women who want to sit on a bench.

a) Suppose that 7 students sit on the bench in different ways. Then, explain some ways that can happen.

b) Arrange 7 objects on a row in different ways and calculate the number of ways of sitting 7 students on a bench.

c) Arrange 3 ordered objects (M_1, M_2, M_3) of the one colour representing 3 men in different possible ways. Then, calculate the number of different ways of sitting all 3 men on the bench.

d) Arrange 4 ordered objects (W_1, W_2, W_3, W_4) of another colour representing 4 women in different possible ways. Then, calculate the number of different ways of sitting all 4 women on the bench.

e) Calculate the number of ways of sitting 3 men and 4 women on the same bench, such that all men sit together on the right hand of all women.

f) Calculate the number of ways of sitting 3 men and 4 women on the same bench such that all men sit together on the right hand side or on the left hand side of all women.

UNIT 3: ELEMENTARY PROBABILITY

Key unit competence: Use counting techniques and concepts of probability to determine the probability of possible outcomes of events occurring under equally likely assumptions.

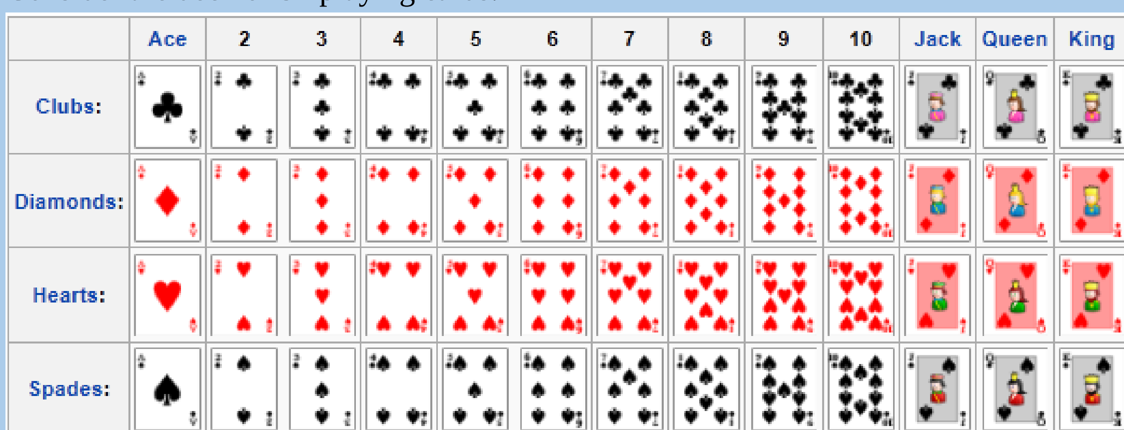
3.0 Introductory activity

A woman applying the family planning program considers the assumption that one boy or one girl can be born at each delivery. If she wishes to have 3 children including two girls and one boy, the family knows that this is a case among other cases which can happen for the 3 children they can get.
With your colleagues, discuss all those cases and deduce the chance that the woman has for having a girl at the first and the second delivery with a boy at the last delivery.

3.1 Concepts of probability

Activity 3.1

Consider the deck of 52 playing cards.



Suppose that you are choosing one card,

a) How many possibilities do you have for the cards to be chosen? What is the related chance?

b) How many possibilities do you have for the kings to be chosen? What is the related chance?

c) How many possibilities do you have for the aces of hearts to be chosen? What is the related chance?

Content summary

Probability is the chance that something will happen.

The concept of probability can be illustrated in the context of a game of 52 playing cards. In a deck of 52 playing cards, cards are divided into four suits of 13 cards each. If a player selects a card at random (by simple random sampling), then each card has the same chance or same probability of being selected. The chance of selecting one card is 1/52.

We cannot say beforehand whether it will show head up or tail up. That depends on chance. The same, a card drawn from a well shuffled pack of 52 cards can be red or black. That depends on chance. Such phenomena are called probabilistic. The theory of probability is concerned with this type of phenomena.

Probability is a concept which numerically measures the degree of uncertainty and therefore of certainty of occurrence of events.

In most sampling situations we are generally not concerned with sampling a specific individual but instead we concern ourselves with the probability of sampling certain types of individuals.

Example 1

If a dotted die is rolled and the number of dots that show up is noted, what is the probability of getting 5 dots?

If a die is rolled until a "6" is obtained, and the number of rolls made before getting first "6" is counted, then we have that the sample space $\Omega = \{1,2,3,4,5,6\}$. $n(\Omega) = 6$ There are 6 possible outcomes. If A is the event: getting 5, $A=\{5\}$, and $n(A)=1$. The chance of getting 5 is the probability of getting 5. That is $p(A)=\frac{n(A)}{n(\Omega)}=\frac{1}{6}$

Example 2

Consider the experiment that consists in rolling a die and recording the number that shows up. Let A be the event "the even number is shown" and B be the event "the odd number less than 5 is shown". Define the events A and B. and find their probability

Solution

We have the sample space $\Omega=\{1,2,3,4,5,6\}$.

$A=\{2,4,6\}$, $n(A)=3$. $B=\{1,3\}$ and $n(B)=2$

$p(A)=\frac{n(A)}{n(\Omega)}=\frac{3}{6}=0.5$ and $p(B)=\frac{n(B)}{n(\Omega)}=\frac{2}{6}=\frac{1}{3}$

Application activity 2.1

When three coins are thrown simultaneously, the sample space $\Omega=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}$. Find:

a) The probability of having 3 tails

b) the probability of having 2 heads.

c) The probability of having one head.

3.2 Determination of Probability of an event, rules and formulas

3.2.1: Probability of an event

Activity 3.2 .1

When a card is selected from an ordinary deck of 52 cards, one assumes that the deck has been shuffled, and each card has the same chance of being selected. Let A be the event of selecting a black card,

a) If n is the number of black cards in the pack, what is the value of n?

b) Calculate the value P (A) = $\frac{n}{number\ of\ all cards}$

c) If P (A) is the probability of electing a black card, deduce the definition of probability for any event E.

Content summary

The probability of an event $A \subset \Omega$, is a real number obtained by applying to A the function P defined by

$$P(A) = \frac{Number\ of\ outcomes\ in\ A}{Total\ number\ of\ outcomes\ in\ the\ sample\ space} = \frac{n(A)}{n(\Omega)}$$

This is the formula for the classical probability, it uses the **sample space** Ω **.**

Probability can be expressed as a fraction, decimal or percentage.

Example

1. For a card drawn from an ordinary deck, find the probability of getting a queen.

Solution: there are 52 cards in a deck and there are 4 queens, $P(\text{queen}) = \frac{n(\text{queen})}{n(\Omega)} = \frac{4}{52} = \frac{1}{13} = 0.076 = \frac{7.6}{100}$

2. A letter is chosen from the letters of the word "**MATHEMATICS**". What is the probability that the letter chosen is an "A"?

Solution

Since two of the eleven letters are A's, we have two favourable outcomes.

There are eleven letters, so we have 11 possible outcomes.

Thus, the probability of choosing a letter A is $\frac{2}{11}$.

Basic probability rules

1. **The probability cannot be negative or greater than 1**

Suppose that an experiment has only a finite number of equally likely outcomes. If A is an event, then $0 \le P(A) \le 1$.

2. **The probability of a certain event**

If the event A is certain to occur, $A = \Omega$, and $P(A) = 1$ and $P(\Omega) = 1$.

3. **Probability of impossible event**

The event that cannot occur is an impossible event $A = \varnothing$, and if $A = \varnothing$ then $P(A) = 0$.

Example

When a single die is rolled, find the probability of getting a 9.

Solution

$\Omega = \{1, 2, 3, 4, 5, 6\}$, it is impossible to get a 9. $A = \varnothing$

P (getting a 9) = 0.

4. **The sum of the probabilities of all the outcomes in the sample space is 1.**

Example

When you roll a fair die, there are six possible outcomes. Each outcome in the sample space has probability of $\frac{1}{6}$. Hence, the sum of the probabilities of all outcomes is given by

$$\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{6}{6}=1$$

5. Probability of complementary event

When *E* and E' are complementary events, $P(E)=1-P(E')$.

Consider two different events, *A* and *B*, which may occur when an experiment is performed.

- The event $A\cup B$ is the event which occurs if *A* or *B* or both *A* and *B* occur, i.e, at least one of *A* and *B* occurs.
- The event $A\cap B$ is the event which occurs if *A* and *B* occur.
- The event $A-B$ is the event which occurs when *A* occurs and *B* does not occur.
- The event A' is the event which occurs when *A* does not occur.

Example

1. When a die is rolled, the event E of getting odd number is that $E=\{1,3,5\}$ and $P(E)=\frac{3}{6}=\frac{1}{2}$

The event F of not getting an odd number is a complement of E. $F=E'=\{2,4,6\}$. As $P(\Omega)=1, P(F)=1-\frac{1}{2}=\frac{1}{2}$.

2. If the probability that a person lives in an industrialized country of the world is $\frac{1}{4}$, find the probability that a person does not live in an industrialized country.

Solution

P (not living in an industrialized country) = 1- P (living in an industrialized country)

$$P=1-\frac{1}{5}=\frac{4}{5}.$$

3) A survey involving 120 people about their preferred breakfast showed that;

55 eat eggs for breakfast, 40 drink juice for breakfast and 25 eat both eggs and drink juice for breakfast.

(a) Represent the information on a Venn diagram.

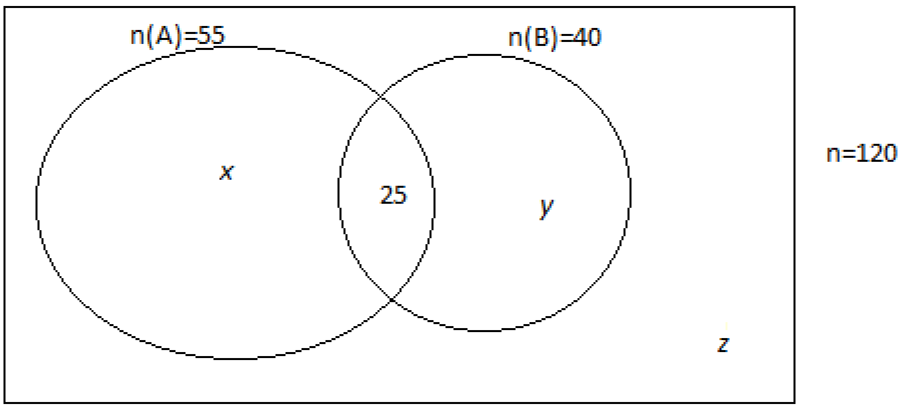
(b) Calculate the following probabilities.

(i) A person selected at random takes only one type for breakfast.

(ii) A person selected at random takes neither eggs nor juice for breakfast.

Solution:

a) Let A= those who eat eggs, B = those who take juice and z represents those who did not take anything.



Here, we can now solve for the number of people who take eggs only:

$x = 55 - 25 = 30$.

So 30 people took Eggs only.

Also, $y = 40 - 25 = 15$.

So, 15 people took Juice only.

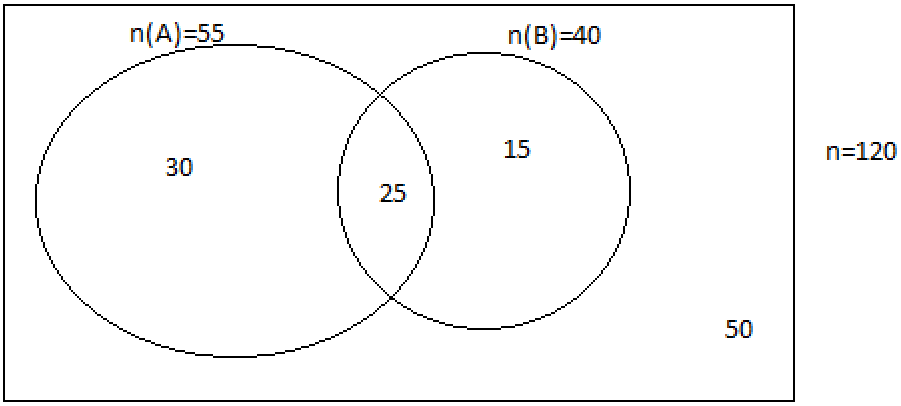
Hence $30 + 25 + 15 + z = 120$

$z = 120 - (30 + 25 + 15)$

$z = 120 - 70 = 50$.

The number of people who did not take anything for breakfast is 50.

Therefore,



b) i) Let E be the event " A person takes one type of breakfast", $P(E)=\frac{45}{120}$

ii) $A\cup B$ is the event "A person takes eggs or juice for breakfast"

then $(A\cup B)'$ is the event "A person takes neither eggs nor juice for breakfast"

for which the probability is $P\left[(A\cup B)'\right]=\frac{50}{120}$.

Properties of probabilities

Referring to rules mentioned above, probabilities assigned to events on a sample space Ω can be summarized in the following properties:

a) $P(E)\geq 0 \; for\, every\, E\subset\Omega$

b) $P(\Omega)=1$

c) $If\; E\subset F\subset\Omega, then\; P(E)\leq P(F)$

d) If A and B are disjoint subsets of Ω, then $If\; E\subset F\subset\Omega, P(\mathrm{A}\cup\mathrm{B})=P(\mathrm{A})+\mathrm{P(B)}$

e) $P(A')=1-\mathrm{P(A)}$ for every $\mathrm{A}\subset\Omega$..

Formula for empirical (classical) probability

Given a frequency distribution, the probability of an event being in a given class is

$$P(A)=\frac{frequency\; for\, the\, class}{Total\; frequency\, in\, the\, distribution}=\frac{f}{n}$$

This probability is called empirical probability and is based on observation.

Examples

1. A researcher asked 25 staff of an institution if they liked the way their breakfast is prepared. The responses were classified as "Yes", "No", and "Undecided". The results were categorized in a frequency distribution as follows:

Response	**Frequency**
Yes	15
No	8
Undecided	2
Total	**25**

a) What is the probability of selecting a person who disliked the way the breakfast is prepared?

b) What is the probability of selecting a person who liked the way the breakfast is prepared?

c) What is the probability of selecting a person who neither like nor disliked the way the breakfast is prepared?

Solution

a) $P(\text{non}) = \frac{f}{n} = \frac{8}{25}$

b) $P(\text{yes}) = \frac{f}{n} = \frac{15}{25}$

c) $P(\text{undecided}) = \frac{f}{n} = \frac{2}{25}$

Application activity 2.3.1

1. Hospital records indicated that maternity patients stayed in the hospital for the number of days shown in the distribution:

Number of days stayed	**Frequency**
3	15
4	32
5	56
6	19
7	5
Total	**127**

Find these probabilities:

a) A patient stayed exactly 5 days

b) A patient stayed less than 6 days

c) A patient stayed at most 4 days

d) A patient stayed at least 5 days

3.2.2: Probability of mutually exclusive (incompatible) and non- inclusive events

Activity 3.2.2

Given a deck of 52 playing cards . If a card is drawn from that pack, Find

(a) the probability P(A) that the card is a club

(b) the probability P(B) that the card is a diamond

(c) the probability P that the card is a club or a diamond.

(d) Compare the probability P to the P(A) and P(B).

Content summary

If two events A and B are such that $A \cup B = \Omega$ and $P(A \cup B) = 1$, these two events are said to be **exhaustive.**

When $A \cap B = \phi$, A and B are mutually exclusive or disjoint and $P(A \cup B) = P(A) + P(B)$

This is called **addition rule** for exclusive events.

When two events A and B are not mutually exclusive $\boldsymbol{A} \cap \boldsymbol{B} \neq 0$, the probability that A or B occurs is given by:

$$P(A \cup B) = \mathrm{P(A)} + \mathrm{P(B)} - \mathrm{P}(A \cap B)$$

This is called addition rule for non-exclusive events

Example

1. A die is thrown once. Let A be the event: "the number obtained is less than 5" and B be the event: "the number obtained is greater than 3". Find probability of $A \cup B$.

Solution

Here $A = \{1,2,3,4\}$ and $B = \{4,5,6\}$, then $A \cup B = \{1,2,3,4,5,6\}$ and then

$$P(A \cup B) = P(\Omega) = 1$$

Or

$$P(A) = \frac{4}{6},\ P(A) = \frac{3}{6},\ A \cap B = \{1\},\ \text{ then } P(A \cap B) = \frac{1}{6}$$

Therefore,

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \frac{4}{6} + \frac{3}{6} - \frac{1}{6} \\ &= 1 \end{aligned}$$

Generally, given a finite sample space, say $\Omega = \{a_1, a_2, a_3, ..., a_n\}$, we can find a finite probability by assigning to each point $a_i \in \Omega$ a real number P_i, called the probability of a_i, satisfying the following:

a) $P_i \geq 0$ for all integers i, $1 \leq i \leq n$;

b) $\sum_{i=1}^{n} P_i = 1$.

If E is an event, then the probability $P(E)$ is defined to be the sum of the probabilities of the sample points in E.

Example

1. A coin is weighted so that heads is three times as likely to appear as tails. Find $P(H)$ and $P(T)$.

Solution

Let $P(T) = p_1$, then $P(H) = 3p_1$.

But $P(H) + P(T) = 1$

Therefore $3p_1 + p_1 = 1 \Leftrightarrow 4p_1 = 1 \Rightarrow p_1 = \frac{1}{4}$

Thus, $P(H) = \frac{3}{4}$ and $P(T) = \frac{1}{4}$.

3. Events A and B are such that they are both mutually exclusive and exhaustive. Find the relation between these two events.

Solution

If A and B are mutually exclusive then

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

If A and B are mutually exclusive then

$P(A\cup B)=1$

Therefore, $P(A)+P(B)=1$ $P(B)=1-P(A)$

But, $P(A')=1-P(A)$

Therefore, $P(B)=P(A')$ i.e. $B=A'$

Similarly, $A=B'$

Thus, if events *A* and *B* are such that they are both mutually exclusive and exhaustive, then they are complementary.

4. A pen is drawn from a basket containing 10 pens of which 5 are red and 3 are black. If A is the event: "a pen is red" and B is the event: "a pen is black", find $P(A),P(B),P(A\cup B)$.

Solution

There are 5 red pens, then $P(A)=\frac{5}{10}=\frac{1}{2}$

There are 3 black pens, then $P(B)=\frac{3}{10}$

Since the pen cannot be red and black at the same time, then $A\cap B=\varnothing$ and two events are mutually exclusive so

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

$$=\frac{1}{2}+\frac{3}{10}=\frac{4}{5}$$

Note that if $A=A_1\cup A_2\cup A_3\cup ...\cup A_n$, where $A_1,A_2,A_3,...,A_n$ are incompatible events, then we may write that $P(A)=\sum_{i=1}^{n}P(A_i)$ for $n=2,3,...$

Therefore, **the addition rule for exclusive events can be extended to any number of exclusive events and be written as follows:**

$$P(A_1 or A_2 or\ A_3 ....A_n) = P(A_1) + P(A_2) + P(A_3) + ... + P(A_n)$$

Example

In race in which there are no dead hearts, the probability that John wins is 0.3, the probability that Paul wins is 0.2 and the probability that Marks wins is 0.4. Find the probability that

a) John or Mark wins

b) John or Paul or Mark wins

c) Someone else wins.

Solution

Since only one person wins, the events are mutually exclusive;

a) $P(John\ or\ Mark\ wins) = P(John\ wins) + P(Mark\ wins) = 0.3 + 0.4 = 0.7$

$P(John\ or\ Paul\ or\ Mark\ wins) = P(John\ wins) + P(Paul\ wins) + P(Mark\ wins)$
$= 0.3 + 0.2 + 0.4 = 0.9$

b) $P(someone\ else\ wins) = 1 - 0.9 = 0.1$

Application activities 3.2.2

An ordinary die is thrown. Find the probability that the number obtained is:

(a) Prime number

(b) Odd number

(c) Even or prime

(d) Less than 4 or multiple of 5

3.2.3: Probability of independent events and multiplication rule

Activity 2.2.3

A box contains 3 red pens, 4 green pens and 5 blue pens. One pen is taken from the box and then replaced. Another pen is taken from the box. Let *A* be the event "the first pen is red" and *B* be the event the second pen is blue."

Is the occurrence of event *B* affected by the occurrence of event *A*? Explain.

Content summary

Events A and B in a probability space *S* are said to be ***independent*** if the occurrence of one of them does not influence the occurrence of the other.

Probability of independent events

When two events are independent, the probability of both occurring is

$P(A\ and\ B) = P(A).\ P(B)$ or $P(A \cap B) = P(A).P(B)$

This is the **multiplication rule** of independent events or **and rule** for independent events.

Example

A fair die is thrown twice. Find the probability that two fives are thrown.

Solution

On one throw, $P(5) = \frac{1}{6}$

On two throws,

$$P(5_1\, and\; 5_2) = P(5_1 \times 5_2) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

$$P(two\;\; fives\; are\; thrown) = \frac{1}{36}$$

Example

A factory runs two machines. The first machine operates for 80% of the time while the second machine operates for 60% of the time and at least one machine operates for 92% of the time. Do these two machines operate independently?

Solution

Let the first machine be M_1 and the second machine be M_2,

then $P(M_1) = 80\% = 0.8$, $P(M_2) = 60\% = 0.6$ and $P(M_1 \cup M_2) = 92\% = 0.92$

Now,

$$P(M_1 \cup M_2) = P(M_1) + P(M_2) - P(M_1 \cap M_2)$$

$$\begin{aligned} P(M_1 \cap M_2) &= P(M_1) + P(M_2) - P(M_1 \cup M_2) \\ &= 0.8 + 0.6 - 0.92 \\ &= 0.48 \\ &= 0.8 \times 0.6 \\ &= P(M_1) \times P(M_2) \end{aligned}$$

Thus, the two machines operate independently.

The multiplication rule can be extended to any number of independent events:

$$P(A_1\; and\;\; A_2\; and\;\; A_3 ....A_n) = P(A_1) \times P(A_2) \times P(A_3) \times ....P(A_n)$$

Example

1) Events A, B and C are independent. If the $P(\text{A}) = \frac{1}{3}; P(\text{B}) = \frac{1}{4}\, and\;\; P(\text{C}) = \frac{2}{5}$. Find the probability of A, B, and C.

Solution

If events A, B and C are independent, then the P (A and B and C) =

$$P(A)\times P(B)\times P(C)=P(\frac{1}{3})\times P(\frac{1}{4})\times P(\frac{2}{5})=\frac{2}{60}=\frac{1}{30}.$$

Application activities 3.2.3

Work out the following questions

1. If A and B are mutually exclusive events, given the probability of A and B as $\frac{1}{5}$ and $\frac{1}{3}$ respectively, find the probability of at least any one event occurring at a time.
2. If X and Y are two events, the probability of the happening of X or Y is $\frac{7}{10}$ and the probability of X is $\frac{1}{3}$. If X and Y are mutually exclusive, find the probability of Y.

3. In a class of a certain school, there are 12 girls and 20 boys. If a teacher want to choose one student to answer the asked question
   a. What is the probability that the chosen student is a girl?
   b. What is the probability that the chosen student is a boy?
   c. If teacher doesn't care on the gender, what is the probability of choosing any student?

3.2.4 Dependent events

Activity 3.2.4

Suppose that you have a deck of cards; then draw a card from that deck, not replacing it, and then draw a second card, etc.

a) What is the sample space for each event?

b) Explain if there is any relationship (Independence or dependence) between those two events considering the sample space. Does the selection of the first card affect the selection of the second card?

Content summary

When the outcome or occurrence of the first event affects the outcome or occurrence of the second event in such a way that the probability is changed, the events are said to be ***dependent***.

Example

1) Suppose a card is drawn from a deck and not replaced, and then the second card is drawn. What is the probability of selecting an ace on the first card and a king on the second card?

Solution

The probability of selecting an ace on the first draw is $\frac{4}{52}$. But since that card is not replaced, the probability of selecting a king on the second card is $\frac{4}{51}$, since there are 51 cards remaining.

The outcomes of the first draw has affected the outcome of the second. By multiplication rule , the probability of both events occurring is : $\frac{4}{52}\times\frac{4}{51}=\frac{16}{2652}=\frac{4}{663}=0.006$.

In this case the $P(A\cap B)=P(A)\times P(B\mid A)$

Application activity 3.2.4

A die is tossed. Find the probability that the number obtained is a 4 given that the number is greater than 2.

3.3 Examples of Events in real life and determination of related probability

Activity 3.3

1. Two football teams in Rwanda " Team A" and "Team B" had to play 3 matches. Two boys Mary and Manasseh made a betting in the following ways in which the winner should be given 400,000Frw when his event succeeds.

Matayo said that team B will gain the first match only and Team A will gain the second and the third. Manasseh said that Team B will gain at least two matches.

a) Between Mary and Manasseh, discuss and determine the boy who has more chances of winning that money.

b) Is there any risk in betting? Referring to the results obtained in a) what are the points of advice you can give to the youth who spend their money in betting?

2. Carry out a research in the library or on internet to find other applications of probability in real life and present them in the classroom discussion.

Content summary

Many people don't care about the risks involved in some activities since they do not understand the concept of probability. On the other hand, people may fear activities that involve little risk to health or life because these activities have been sensationalized by the press and media.

We have to think big before taking decision regarding our engagement in the games of chance. Such games are for example: betting on card games, slot machine (ikiryabarezi), lotteries, and weather forecasting. In such games, predictions are based on probability and hypotheses are tested by using probability.

The following are example of applications of rules of probability to solve some problems we can meet in our life experience.

Examples

1) A box contains 3 blue marbles, 4 red marbles and 5 yellow marbles. If a person selects one marble at a random, find the probability that it is either a blue or yellow marble.

Solution

The total of marbles is 12. Since there are 3 blue and 5 yellow marbles,

$$P(\text{blue or yellow}) = \text{P(blue)} + \text{P(yellow)} = \frac{3}{12} + \frac{5}{12} = \frac{8}{12}$$

2) In a political rally, there are 200 republicans, 130 Democrats and 60 independents. If a person is selected at random, find the probability that he/she is either democrat or independent.

Solution

$$P(\text{Democrat or } independent) = \text{P(Democrat)} + \text{P(independent)} = \frac{130}{390} + \frac{60}{390} = \frac{19}{39}$$

3) A single card is drawn from a deck. Find the probability that it is a king or a club.

Solution

As the king of clubs is counted twice, one of the two probabilities must be subtracted (it a part of intersection)

$$P(\text{king or } c\text{lub}) = \text{P(king)} + \text{P(club)} - \text{P(king of clubs)} = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{4}{13}$$

4) In a hospital unit there are 8 nurses and 5 physicians; 7 nurses and 3 physicians are females. If a staff person is selected, find the probability that the subject is a nurse or a male.

Solution

The sample space is:

Staff	**Females**	**Males**	**Total**
Nurses	7	1	8
Physicians	3	2	5
Total	**10**	**3**	**13**

The probability is

$$P(\text{nurse or } male) = \text{P(nurse)} + \text{P(male)} - \text{P(male nurse)} = \frac{8}{13} + \frac{3}{13} - \frac{1}{13} = \frac{10}{13}.$$

5) A card is drawn from a deck of 52 playing cards. If A is an event of drawing an ace and B is an event of drawing a spade. Find $P(A), P(B), P(A \cap B), P(A \cup B)$

Solution

There are 4 aces, then $P(A) = \frac{4}{52} = \frac{1}{13}$

There are 13 spades, then $P(B) = \frac{13}{52} = \frac{1}{4}$

There is 1 ace of spades, then $\#(A \cap B) = 1$ and $P(A \cap B) = \frac{1}{52}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$$

Alternatively, there are 4 aces and 13 spades but also 1 ace of spades. Then

$\#(A \cup B) = 16$ and $P(A \cup B) = \frac{16}{52} = \frac{4}{13}$.

Application activity 3.3

1. In one state of America, the probability that a student owns a car is 0,65, and the probability that a student owns a computer is 0.82. If the probability that a student owns both is 0.55, what is the probability that a given student owns neither a car nor a computer?

2. At a particular school with 200 male students, 58 play football, 40 play basketball, and 8 play both. What is the probability that a randomly selected male student plays neither sport?

3.4 End unit assessment

Lottery:

An urn contains 20 lottery tickets numbered from 1 to 20.
To buy a ticket, each one is selected at random and replaced before the next selection. The organizer of the lottery decided to pay 1000Frw to the one who will select a number divisible by 4 and 3 at the same time. He will pay also 500Frw to the one who will select a number which is divisible by 5 and 2 at the same time.

1. Given that the 20 tickets numbered from 1 to 20 were bought at 200Frw per ticket, do the following:

   a) Play this lottery in your class and observe its outcomes. Is the game fair or not?
   b) The money received by the organizer of the lottery
   c) The probability for participants to win 1000Fr
   d) The probability for participants to win 500Frw
   e) The money to be made by the organizer of the lottery.

2. The parents of your friend Anne Marie gave her 200Frw for buying two pens, however, she wants to participate in the lottery to get more money before buying pens. What can you advise her?

Hint: *Use the following events: A: selecting a number divisible by 4; B: selecting the number divisible by 3; C: selecting the number divisible by 5, and D: Selecting the number divisible by 2.*

REFERENCE

Allan G. B. (2007). Elementary statistics: a step by step approach, seventh edition, *Von Hoffmann Press*, New York.

David R. (2000). Higher GCSE Mathematics, revision and Practice. Oxford University Press, UK.

Elliot M. (1998). Schaum's outline series of Calculus. MCGraw-Hill Companies, Inc. USA.

Geoff Mannall & Michael Kenwood, Pure Mathematics 2, Heinemann Educational Publishers 1995

George B. Thomas, Maurice D. Weir & Joel R. Hass, Thomas' Calculus Twelfth Edition, Pearson Education, Inc. 2010

Gilbert J.C. et all. (2006). Glencoe Advanced mathematical concepts, MCGraw-Hill Companies, Inc. USA.

Glencoe. (2006). Advanced mathematical concepts, Precalculus with Applications.

J. CRAWSHAW and J. CHAMBERS 2001. A concise course in Advanced Level Statistics with worked examples 4^{th} Edition. Nelson Thornes Ltd, UK.

J. Sadler, D. W. S. Thorning: Understanding Pure Mathematics, Oxford University Press 1987.

J.K. Backhouse, SPTHouldsworth B.E.D. Copper and P.J.F. Horril. Pure Mathematics 2. Longman, third edition 1985, fifteenth impression 1998.

JF Talber & HH Heing, Additional Mathematics 6th Edition Pure & Applied, Pearson Education South Asia Pte Ltd 1995

John bird. (2005). Basic engineering mathematics. 4^{th} Edition. Linacre House, Jordan Hill, Oxford OX2 8DP

K.A. Stroud. (2001). Engineering mathematics. 5^{th} Edition. Industrial Press, Inc, New York

Michael Sullivan, 2012. Algebra and Trigonometry 9^{th} Edition. Pearson Education, Inc

Ngezahayo E.(2016). Subsidiary Mathematics for Rwanda secondary Schools, Learners' book 4, Fountain publishers, Kigali.

Paule Faure- Benjamin Bouchon, Mathématiques Terminales F. Editions Nathan, Paris 1992.

Peter S. (2005). Mathematics HL&SL with HL options, Revised edition. Mathematics Publishing PTY. Limited.

REB. (2015). Subsidiary Mathematics Syllabus, MINEDUC, Kigali, Rwanda.

REB. (2019). Mathematics Syllabus for TTC-Option of LE, MINEDUC, Kigali Rwanda.

Robert A. Adms & Christopher Essex, Calculus A complete course Seventh Edition, Pearson Canada Inc., Toronto, Ontario 2010

Ron Larson and David C (2009). Falvo. Brief Calculus, An applied approach. Houghton Mifflin Company.

Sadler A. J & Thorning D.W. (1997). Understanding Pure mathematics, Oxford university press, UK.