Summation Notation: $(\sum \equiv \boldsymbol{sigma})$

$\sum_{i=1}^{n} x_i$ means $x_1 + x_2 + x_3 + \cdots + x_n$, the letter i is called the index of the summation. The letter i is dummy variable and it can be replaces by any letter, that is:

$$x_1 + x_2 + x_3 + \cdots + x_n = \sum_{i=1}^{n} x_i = \sum_{j=1}^{n} x_j = \sum_{k=1}^{n} x_k$$

Examples:

(1) $\sum_{i=1}^{3} x_i = x_1 + x_2 + x_3.$

(2) $\sum_{i=1}^{4} x_i^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2.$

(3) $\left(\sum_{i=1}^{3} x_i\right)^2 = (x_1 + x_2 + x_3)^2.$

(4) $\sum_{i=1}^{2} x_i - 2 = x_1 + x_2 - 2.$

(5) $\sum_{i=1}^{3} (x_i - 2) = (x_1 - 2) + (x_2 - 2) + (x_3 - 2).$

(6) $\sum_{i=1}^{3} \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \frac{x_3}{y_3}.$

(7) $\frac{\sum_{i=1}^{2} x_i}{\sum_{i=1}^{3} y_i} = \frac{x_1 + x_2}{y_1 + y_2 + y_3}$.

(8) $\sum_{i=1}^{4} x = x + x + x + x = 4x.$

(9) $\sum_{x=1}^{3} x = 1 + 2 + 3 = 6.$

(10) $\sum_{i=1}^{2} 3^{i+1} = 3^2 + 3^3 = 9 + 27 = 36.$

(11)

i	1	2	3	4
x_i	4	2	3	7
y_i	3	9	6	2

Find

(i) $\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4 = 4 + 2 + 3 + 7 = 16.$

(ii) $\sum_{i=1}^{3} x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 = 4 * 3 + 2 * 9 + 3 * 6$

$$= 12 + 18 + 18 = 48.$$

(iii) $(\sum_{i=1}^{4} x_i)(\sum_{i=1}^{4} y_i) = 16 * 20 = 320.$

(iv) $\sum_{i=2}^{3} ix_i y_{i+1} = 2x_2y_3 + 3x_3y_4 = 2 * 2 * 6 + 3 * 3 * 2 = 24 + 18 = 42.$

Some Important Sums:

(1) $\sum_{i=1}^{n} i = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$

(2) $\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$

(3) $\sum_{i=1}^{n} i^3 = 1^3 + 2^3 + \cdots + n^3 = \left[\frac{n(n+1)}{2}\right]^2.$

Some Basic Rules:

If c is any constant, then

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

(2) $\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i.$

(3) $\sum_{i=1}^{n}(x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i.$

Proof:

(1) $\sum_{i=1}^{n} c = c + c + \cdots + c = nc.$

(2) $\sum_{i=1}^{n} cx_i = cx_1 + cx_2 + \cdots + cx_n = c(x_1 + x_2 + \cdots + x_n) = c\sum_{i=1}^{n} x_i.$

(3) $\sum_{i=1}^{n}(x_i + y_i) = (x_1 + y_1) + (x_2 + y_2) + \cdots + (x_n + y_n)$

$$= (x_1 + x_2 + \cdots + x_n) + (y_1 + y_2 + \cdots + y_n)$$

$$= \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i.$$

Exercises:

(1) Simplify $\sum_{i=1}^{n}(i-1)^2$.

(2) If $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, then show that $\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} {x_i}^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2$.

Probability and Statistics

Description Statistical:

Description statistical concerned with statistical methods that used in describing a set of data. These methods include the procedure of collecting data as numerical measurements and then organizing, summarizing, presenting such data and calculating some statistical measurements in order to obtain meaningful information.

Statistical Inference:

Statistical inference concerned with statistical methods that used in the analysis of a subset of data leading to inferences about the entire set of data.

Definition: (Statistics)

Statistics is the science uses a scientific procedure in methods of collecting, organizing, summarizing, presenting, calculating and analyzing a set of data in order to reach to inferences and appropriate decisions.

Nature of Statistical Data:

Definition: (Observations)

Any recorded information about certain phenomena whether it is numerical or categorical called observations.

Remark:

We shall use the symbol x_i to denote the observation.

Definition: (Variable)

Any phenomena shows differences in its observation called variable.

There are two types of variables:

(1) Qualitative Variables.

(2) Quantitative Variables.

Definition: (Qualitative Variable)

Qualitative variable is that phenomena whose observations can not be measured by numerical numbers.

Examples:

(1) Tossing a coin: x=Head or Tail.

(2) Sex: x=Male or Female.

(3) Eyes color: x=Black or Brown or Blue or Green.

(4) Blood type: x=A or B or AB or D (with plus or minus signs)

Definition: (Quantitative Variable)

Quantitative variable is that phenomena whose observations can be measured by numerical numbers.

Examples:

(1) Casting a die: x=1, 2, 3, 4, 5, 6.

(2) Exam grades: x=0,1,2,…,100.

(3) Height: x=150, 100,….

Remark:

There are two types of quantitative variables:

(1) Discrete Quantitative Variables.

(2) Continuous Quantitative Variables.

Definition: (Discrete Quantitative Variable)

Discrete quantitative variable is the quantitative variable defined on a countable (finite or infinite) set of data.

Examples:

(1) Casting a die: x=1, 2, 3, 4, 5, 6 (countable (finite)).

(2) Selecting a natural number: x=1,2,… (countable (infinite)).

Definition: (Continuous Quantitative Variable)

Continuous quantitative variable is the quantitative variable defined on uncountable set of data.

Example:

Selecting a real number from the interval (0,1):$0<x<1$ (uncountable).

Population and Samples:

Definition: (Population)

A population consists of the totally of observations with which we are concerned under certain discussion.

Definition: (Sample)

A sample is a set of observations selected in such a way from a population, that is, a sample is a subset of a population.

Statistical Description of Data

When we face a large masses of statistical data about certain phenomena, we summarize and display such data in an appropriate tables,

charts and graphs so that the study and the analysis of such data become easier.

Tabular Presentation

In tabular presentation, we group the data into different classes and then determining the number of observations that fall in each class. Such arrangement in tabular form is called **frequency distribution.**

Definition: (Frequency Distribution Table)

It is a simple table consists of 4 columns: The first column represents class interval (C.I), the second column represents class boundaries (C.B), the third column represents class marks (x_i) and the fourth column represents frequencies (f_i).

The frequency distribution table with k classes can be pictured as follows:

class interval C.I	class boundaries C.B	class mark x_i	frequency f_i
$a_1 - b_1$	$L_1 - U_1$	x_1	f_1
$a_2 - b_2$	$L_2 - U_2$	x_2	f_2
⋮	⋮	⋮	⋮
$a_k - b_k$	$L_k - U_k$	x_k	f_k
Total			$\sum_{i=1}^{k} f_i$

Data presented in the form of frequency distribution are called **group data**, otherwise called **ungrouped data**.

To each class interval, there are two limits called respectively, **the lower and the upper limits** and denoted by a_i and b_i, i=1,2,…,k.

Definition: (Class Boundaries), (C.B)

Class boundaries are class intervals but they always carried out to one more decimal place than the record observation to ensure that no observation can fall precisely on class boundary.

Definition: (Class Mark), (x_i)

Class mark is the midpoint between the upper and lower class limits or boundaries, that is

$$x_i = \frac{1}{2}(a_i + b_i), i = 1,2, \dots, k$$

or

$$x_i = \frac{1}{2}(U_i + L_i), i = 1,2, \dots, k$$

Definition: (Class Frequency), (f_i)

Class frequency is the number of observations falling in a particular class interval boundary.

Definition: (Class Width), (C)

Class width is the numerical difference between the upper and lower class boundaries, that is

$$C = U_i - L_i, \qquad i = 1,2, \dots, k$$

or

$$C = a_{i+1} - a_i = b_{i+1} - b_i = L_{i+1} - L_i = U_{i+1} - U_i, i = 1,2, \dots, k$$

or

$$C = x_{i+1} - x_i, i = 1,2, \dots, k$$

Definition: (Range), (R)

Rang is the difference between the largest and the smallest number in the set of data.

The Steps in Grouping a Set of Data into Frequency Distribution Table

Step (1): Arrange the dat in ascending order.

Step (2): Find the range R.

Step (3): Decide the number of C.I (usually it is between 5 and 15).

Step (4): Divide the range R by the number of classes and then estimate the approximate class width C.

Step (5): For the first C.I, decide on the lower class limit and then the lower C.B. Then add C to the lower C.B to obtain the upper C.B. Write down the upper class limit.

Step (6): List all class limits and C.B by adding C to the class limits and C.B.

Step (7): Find the class number x_i for each interval by averaging the class limits.

Step (8): Count the number of observations falling in each class.

Step (9): Sum the frequency column and check the total number of observations.

Example:

The following data:

2.2	3.4	2.5	3.3	4.7	4.1	1.6
4.3						
3.1	3.8	3.5	3.1	3.4	3.7	3.2
4.5						

3.3 3.6 4.4 2.6 3.2 3.8 2.9
3.2
3.9 3.7 3.1 3.3 4.1 3.0 3.0
4.8
3.9 1.9 4.2 2.6 3.7 3.1 3.4
3.5

represents the lives of 40 similar car batteries recorded to the nearest tenth of a year. Choose 7 C.I with the lowest starting at 1.5 and set up the frequency distribution table.

Solution:

We arrange the data in ascending order:

1.6 1.9 2.2 2.5 2.6 2.6 2.9
3.0
3.0 3.0 3.1 3.1 3.1 3.1 3.2
3.2
3.3 3.3 3.3 3.4 3.4 3.4 3.5
3.5
3.6 3.7 3.7 3.7 3.8 3.8 3.9
3.9
4.1 4.1 4.2 4.3 4.4 4.5 4.7
4.8

Range$\equiv R = 4.8 - 1.6 = 3.2.$

Class Width$\equiv C = \frac{R}{number\ of\ classes} = \frac{3.2}{7} = 0.457 \cong 0.5.$

C.I	C.B	x_i	f_i
1.5−1.9	1.45−1.95	1.7	2
2.0−2.4	1.95−2.45	2.2	1
2.5−2.9	2.45−2.95	2.7	4

3.0−3.4	2.95−3.45	3.2	15
3.5−3.9	3.45−3.95	3.7	10
4.0−4.4	3.95−4.45	4.2	5
4.5−4.9	4.45−4.95	4.7	3
Total			$\sum_{i=1}^{7} f_i = 40$

Relative Frequency Distribution Table:

The relative frequency for each C.I obtained by dividing the class frequency by total frequency, that is

$$\text{relative frequency for each C.I} = \frac{f_i}{\sum_{i=1}^{k} f_i}.$$

Percentage Frequency Distribution Table:

The percentage frequency for each C.I obtained by multiplying the relative frequency for this C.I by 100, that is

$$\text{percentage frequency for each C.I} = \frac{f_i}{\sum_{i=1}^{k} f_i} \times 100$$

Remarks:

(1) A table listing relative frequencies is called **relative frequency distribution table**.

(2) A table listing percentage frequencies is called **percentage frequency distribution table**.

Example:

The following data:

2.2 3.4 2.5 3.3 4.7 4.1 1.6 4.3

3.1	3.8	3.5	3.1	3.4	3.7	3.2
4.5						
3.3	3.6	4.4	2.6	3.2	3.8	2.9
3.2						
3.9	3.7	3.1	3.3	4.1	3.0	3.0
4.8						
3.9	1.9	4.2	2.6	3.7	3.1	3.4
3.5						

represents the lives of 40 similar car batteries recorded to the nearest tenth of a year. Choose 7 C.I with the lowest starting at 1.5 and set up the relative and percentage frequency distribution table.

Solution:

C.I	C.B	x_i	f_i	relative f_i	percentage f_i
1.5−1.9	1.45−1.95	1.7	2	0.05	5
2.0−2.4	1.95−2.45	2.2	1	0.025	2.5
2.5−2.9	2.45−2.95	2.7	4	0.1	10
3.0−3.4	2.95−3.45	3.2	15	0.375	37.5
3.5−3.9	3.45−3.95	3.7	10	0.25	26
4.0−4.4	3.95−4.45	4.2	5	0.125	12.5
4.5−4.9	4.45−4.95	4.7	3	0.075	7.5
Total			$\sum_{i=1}^{7} f_i = 40$	1	100

This table is called relative and percentage frequency distribution table.

Definition: (Cumulative Frequency Distribution Table):

The total frequency of all values less than the upper C.B of a given C.I is called cumulative frequency up to and including that class and is denoted by F_i. A table listing cumulative frequencies is called cumulative frequency distribution table.

Example:

The following data:

2.2	3.4	2.5	3.3	4.7	4.1	1.6
4.3						
3.1	3.8	3.5	3.1	3.4	3.7	3.2
4.5						
3.3	3.6	4.4	2.6	3.2	3.8	2.9
3.2						
3.9	3.7	3.1	3.3	4.1	3.0	3.0
4.8						
3.9	1.9	4.2	2.6	3.7	3.1	3.4
3.5						

represents the lives of 40 similar car batteries recorded to the nearest tenth of a year. Choose 7 C.I with the lowest starting at 1.5 and set up the cumulative frequency distribution table.

Solution:

C. B	F_i
less than 1.45	0
less than 1.95	2
less than 2.45	3
less than 2.95	7
less than 3.45	22

less than 3.95	32
less than 4.45	37
less than 4.95	40

Definition: (Relative Cumulative Frequency Distribution Table):

The relative cumulative frequency each C.I obtained by dividing the class cumulative frequency by total frequency, that is,

$$\text{relative cumulative frequency each C.I} = \frac{F_i}{\sum_{i=1}^{k} f_i}$$

A table listing relative cumulative frequencies is called relative cumulative frequency distribution table.

Definition: (Percentage Cumulative Frequency Distribution Table):

The percentage cumulative frequency for each C.I obtained by multiplying the relative cumulative frequency for this C.I by 100, that is

$$\text{percentage cumulative frequency for each C.I} = \frac{F_i}{\sum_{i=1}^{k} f_i} \times 100$$

A table listing percentage cumulative frequencies is called percentage cumulative frequency distribution table.

Example:

The following data:

2.2 3.4 2.5 3.3 4.7 4.1 1.6 4.3

3.1 3.8 3.5 3.1 3.4 3.7 3.2 4.5

3.3 3.6 4.4 2.6 3.2 3.8 2.9 3.2

3.9 3.7 3.1 3.3 4.1 3.0 3.0 4.8

3.9 1.9 4.2 2.6 3.7 3.1 3.4 3.5

represents the lives of 40 similar car batteries recorded to the nearest tenth of a year. Choose 7 C.I with the lowest starting at 1.5 and set up the relative and percentage cumulative frequency distribution table.

Solution:

C. B	F_i	relative F_i	percentage F_i
less than 1.45	0	0	0
less than 1.95	2	0.05	5
less than 2.45	3	0.075	7.5
less than 2.95	7	0.175	17.5
less than 3.45	22	0.55	55
less than 3.95	32	0.8	80
less than 4.45	37	0.92	92
less than 4.95	40	1.00	100

Example:

Find C.B, class marks and class width for the data in the following table:

	C.I	C.B	class mark x	class width C
a	16-24	15.5−24.5	20	9
b	4.2−7.6	4.15−7.65	5.9	3.5
c	3.26−6.78	3.25−6.80	5.02	3.55
d	−8 − −2	−8.5− − 1.5	−5	7
e	−6.3−11.5	−6.35−11.55	2.6	17.9

Example:

Complete the following table:

C.I	C.B	x_i	f_i	relative f_i	percentage f_i
		4	2		
		9	5		
			10		
			25		
			8		
Total			$\sum_{i=1}^{5} f_i = 50$		

Solution:

Class Width≡ $C = x_{i+1} - x_i = x_2 - x_1 = 9 - 4 = 5.$

$$x_1 = \frac{1}{2}(U_1 + L_1) \Rightarrow 4 = \frac{1}{2}(U_1 + L_1) \Rightarrow$$

$U_1 + L_1 = 8$

(1)

$$C = U_1 - L_1 \Rightarrow$$

$U_1 - L_1 = 5$

(2)

By solving the system given by equations (1)-(2) we get

$$L_1 = 1.5$$

$$U_1 = 6.5$$

Therefore the above table becomes:

C.I	C.B	x_i	f_i	relative f_i	percentage f_i
2−6	1.5−6.5	4	2	0.04	4
7−11	6.5−11.5	9	5	0.1	10
12−16	11.5−16.5	14	10	0.2	20
17−21	16.5−21.5	19	25	0.5	50
22−26	21.5-26.5	24	8	0.16	16

Total	$\sum_{i=1}^{5} f_i = 50$	1	100

Graphical Representation:

The information provided by frequency distribution in tabular for is easier to understand if it presented graphically.

Frequency Histogram:

Frequency histogram constructed as follows:

Let the x-axis represent the C.B's with equal spaces and y-axis represent the frequency with different scales. Then draw to each one a rectangle in which its base represent the class width and its height represent the frequency.

Frequency Polygon:

Frequency polygon constructed by plotting class marks against the class frequencies and then join the consecutive points by a straight line.

A polygon is a many-sided closed figure. To close the frequency polygon an additional C.I must be added to both end of the distribution each with zero frequency.

Frequency Curve:

Frequency curve constructed by plotting class marks against the class frequencies and then join most of the consecutive points by smooth curve.

Exercises:

(1) Find C.B, class marks and class width for the data in the following table:

C.I	C.B	x_i	C
7−9			
-5−(-1)			
10.4−18.7			
0.346−0.418			
-2.75−1.35			
78.49−86.75			

(2) The following data represent the final examination grade for statistics course:

23	60	79	32	57	74	52	70	82	36
80	77	81	95	41	65	92	85	55	76

52	10	64	75	78	25	80	98	81	67
41	71	83	54	64	72	88	62	74	43
60	78	89	76	84	48	84	90	15	79
34	67	17	82	69	74	63	80	85	61

Use 9 C.I with the lowest starting at 10:

(i) Set up the frequency, relative frequency and percentage frequency distribution table.

(ii) Set up the cumulative frequency distribution table, relative cumulative frequency distribution table and percentage cumulative distribution table.

(iii) Construct the frequency Histogram, frequency polygon and frequency curve.



Statistical Measures of Data

(i) Measures of Central Tendency.

(ii) Measures of Variation.

(i) Measures of Central Tendency:

(1) Sample Mean.

(2) Weighted Mean.

(3) Geometric Mean.

(4) Harmonic Mean.

(5) Quadratic Mean.

(6) Median.

(7) Mode.

(1) Sample Mean:

If the set of data $x_1, x_2, \cdots, x_n$ represent a finite sample of size n then the sample mean denoted by $\bar{x}$ is defined by:

(a) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ for ungrouped data.

(b) $\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$ for grouped data,

where x_i is a class mark, f_i is a frequency and k is the number of C.I.

Some Properties of the Sample Mean:

(1) The sum of deviations of observations from their sample mean is zero, that is

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

Proof:

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

(2) The sum of squares of deviations of observations from their sample mean is minimum, that is

$\sum_{i=1}^{n}(x_i-\bar{x})^2 \le \sum_{i=1}^{n}(x_i-k)^2$, k is any constant

Proof:

$$\sum_{i=1}^{n}(x_i-k)^2=\sum_{i=1}^{n}[(x_i-\bar{x})+(\bar{x}-k)]^2$$

$$=\sum_{i=1}^{n}[(x_i-\bar{x})^2+2(x_i-\bar{x})(\bar{x}-k)+(\bar{x}-k)^2]$$

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

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

Since $n(\bar{x}-k)^2 \ge 0$, then $\sum_{i=1}^{n}(x_i-k)^2 \ge \sum_{i=1}^{n}(x_i-\bar{x})^2$.

Example:

Let $x_1=2, x_2=3, x_3=7, x_4=13, x_5=15$, then n=5.

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{5}(2+3+7+13+15)=8.$$

x_i	$x_i-\bar{x}$	x_i-6	$(x_i-\bar{x})^2$	$(x_i-6)^2$
2	-6	-4	36	16
3	-5	-3	25	9

7	-1	1	1	1
13	5	7	25	49
15	7	9	49	81
Sum	0	10	136	156

Therefore

$\sum_{i=1}^{5}(x_i-\bar{x})=0, \sum_{i=1}^{5}(x_i-\bar{x})^2=136$ and $\sum_{i=1}^{5}(x_i-6)^2=156.$ Hence

$$\sum_{i=1}^{5}(x_i-\bar{x})^2 \le \sum_{i=1}^{5}(x_i-6)^2.$$

(3) If $y_i = x_i + k, i = 1,2,\dots,n$, where k is any constant, then $\bar{y} = \bar{x} + k$.

Proof:

$$y_i = x_i + k \Rightarrow \sum_{i=1}^{n} y_i = \sum_{i=1}^{n}(x_i + k) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} k = \sum_{i=1}^{n} x_i + nk \Rightarrow$$

$$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\sum_{i=1}^{n} x_i + nk \Rightarrow \bar{y} = \bar{x} + k.$$

(4) If $y_i = k\,x_i\,, i = 1,2,\dots,n$, where k is any constant, then $\bar{y} = k\bar{x}$.

Proof:

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

Example:

Let n=6 and $x_1 = 2, x_2 = 4, x_3 = 4, x_4 = 8, x_5 = 5, x_6 = 7.$

(i) If $y_i = x_i - 2, i = 1,2,\dots,6$, then find $\bar{y}$.

(ii) If $y_i = 5\,x_i, i = 1,2,\dots,6$, then find $\bar{y}$.

Solution:

(i) $\bar{x} = \frac{1}{n}\sum_{i=1}^{6} x_i = \frac{1}{6}(2+4+4+8+5+7) = 5 \Rightarrow \bar{y} = \bar{x} - 2 \Rightarrow$

$$\bar{y} = 5 - 2 = 3.$$

(ii) $\bar{y} = 5\bar{x} \Rightarrow \bar{y} = 25.$

(2) Weigthed Mean:

When we wish to average k quantities $x_1, x_2, \cdots, x_k$ by attaching more significance to some of numbers than to other, we accomplish this by assigning weights $w_1, w_2, \cdots, w_k$ to the quantities, where the weights represent measures of their importance. The corresponding weighted mean denoted by $\bar{x}_w$ is defined by:

$$\bar{x}_w = \frac{\sum_{i=1}^{k} w_i x_i}{\sum_{i=1}^{k} w_i}.$$

Example:

Find the average for a student who received grades with significance as given in the following table:

Exam	Grade	Significance
1st	70	10%
2nd	75	20%
3rd	90	30%
4th	55	40%

Then

$$\bar{x}_w = \frac{\sum_{i=1}^{k} w_i x_i}{\sum_{i=1}^{k} w_i} = \frac{(10*70)+(20*75)+(30*90)+(40*55)}{10+20+30+40}$$

$$= 71.$$

(3) Geometric Mean:

If $x_1, x_2, \dots, x_n$ is a finite sample of size n, then the geometric mean denoted by $\bar{G}$ is defined by:

(a) $\bar{G} = \sqrt[n]{x_1 . x_2 . \dots . x_n}$ for ungrouped data.

(b) $\bar{G} = \sqrt[\sum_{i=1}^{k} f_i]{x_1^{f_1} . x_2^{f_2} . \dots . x_k^{f_k}}$ for grouped data,

where x_i is a class mark, f_i is a frequency and k is the number of C.I.

Remark:

In practice, we may write the geometric mean as follows:

(a) $\bar{G} = e^{\frac{1}{n}\sum_{i=1}^{n} lnx_i}$ for ungrouped data.

(b) $\bar{G} = e^{\frac{\sum_{i=1}^{k} f_i lnx_i}{\sum_{i=1}^{k} f_i}}$ for grouped data.

(4) Harmonic Mean:

If $x_1, x_2, \dots, x_n$ is a finite sample of size n, then the Harmonic mean denoted by $\bar{H}$ is defined by:

(a) $\bar{H} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$ for ungrouped data.

(b) $\bar{H} = \frac{\sum_{i=1}^{k} f_i}{\sum_{i=1}^{k} \frac{f_i}{x_i}}$ for grouped data,

where x_i is a class mark, f_i is a frequency and k is the number of C.I.

(5) Quadratic Mean:

If $x_1, x_2, \dots, x_n$ is a finite sample of size n, then the quadratic mean denoted by $\bar{Q}$ is defined by:

(a) $\bar{Q} = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$ for ungrouped data.

(b) $\bar{Q} = \sqrt{\frac{\sum_{i=1}^{k} f_i x_i^2}{\sum_{i=1}^{k} f_i}}$ for grouped data,

where x_i is a class mark, f_i is a frequency and k is the number of C.I.

(6) Meadian:

If $x_1, x_2, \dots, x_n$ is a finite sample of size n, then the median denoted by M_e is defined by:

$$M_e = \begin{cases} y_{\frac{n+1}{2}} & \text{if n is odd} \\ \frac{y_{\frac{n}{2}} + y_{\frac{n}{2}+1}}{2} & \text{if n is even} \end{cases}$$ for ungrouped data,

where $y_1, y_2, \dots, y_n$ is a rearrangement of the observations $x_1, x_2, \dots, x_n$ in an increasing or decreasing order.

Examples:

(a) Let n=7 and x_1=-3, x_2=-7, $x_3 = 0$, $x_4 = 4$, $x_5 = 11$, $x_6 = -1$, $x_7 = -2$. Find the median of this sample.

(b) Let n=10 and $x_1 = -1$, $x_2 = 5$, $x_3 = 0$, $x_4 = 4$, $x_5 = -2$, $x_6 = 8$, $x_7 = 3$, $x_8 = 8$, $x_9 = 6$, $x_{10} = 11$. Find the median of this sample.

Solution:

(a) $y_1 = -7, y_2 = -3, y_3 = -2, y_4 = -1, y_5 = 0, y_6 = 4, y_7 = 11$. Therefore

$$M_e = y_{\frac{n+1}{2}} = y_{\frac{7+1}{2}} = y_4 = -1.$$

(b) $y_1 = -2, y_2 = -1, y_3 = 0, y_4 = 3, y_5 = 4, y_6 = 5, y_7 = 6, y_8 = 8, y_9 = 8, y_{10} = 11$. Therefore

$$M_e = \frac{y_{\frac{n}{2}} + y_{\frac{n}{2}+1}}{2} = \frac{y_{\frac{10}{2}} + y_{\frac{10}{2}+1}}{2} = \frac{y_5 + y_6}{2} = \frac{4+5}{2} = \frac{9}{2} = 4.5.$$

(7) Mode:

If $x_1, x_2, \dots, x_n$ is a finite sample of size n, then the mode denoted by M_o is defined by:

for ungrouped data: M_o is the value that occurs upper most ,

Remark:

The mode of a finite sample may not exists or may have more than one mode.

Examples:

(a) Let n=6 and $x_1 = -2, x_2 = 11, x_3 = 4, x_4 = -3, x_5 = 7, x_6 = 13$. Find the mode (if it exists) of this sample.

(b) Let n=10 and $x_1 = 9, x_2 = 10, x_3 = 5, x_4 = 9, x_5 = 9, x_6 = 7, x_7 = 8, x_8 = 6, x_9 = 10, x_{10} = 4$. Find the mode (if it exists) of this sample.

(c) Let n=12 and $x_1 = 2, x_2 = 0, x_3 = 3, x_4 = 1, x_5 = 2, x_6 = 4, x_7 = 2, x_8 = 5, x_9 = 4, x_{10} = 0, x_{11} = 1, x_{12} = 4$. Find the mode (if it exists) of this sample.

Solution:

(a) There is no mode of this sample.

(b) $M_o = 9$.

(c) $M_o = 2,4$.

(ii) Measures of Variation

The measures of central tendency do not give a right description of our data, so we need to know how the observations spread out from the average. We may have two sets of observations with the same mean but there is a considerable different in the variabilities about the average. For example, the following two samples:

Sample A	17	20	23	18	19	21	22
Sample B	35	15	7	5	45	20	13

$\bar{x}_A = 20 = \bar{x}_B$. Therefore these samples have the same mean but we can see that the sample A is more uniform than the sample B.

Measures of variation provide a measure of variability among the observations.

(1) Range.

(2) Mean Deviation.

(3) Variance.

(4) Standard Deviation.

(1) Range:

The range R is the difference between the largest and the smallest numbers in the set of data, that is,

$$R = x_{max} - x_{min}.$$

Remark:

The range is a poor measure of variation if the size of the sample is large. It considers only the extreme values and tell us nothing about the distribution of numbers in between.

For example, the following two samples:

Sample A	3	4	5	6	8	9	10	12	15
Sample B	3	7	7	7	8	8	8	9	15

$R_A = 15 - 3 = 12 = \ R_B$ and $\bar{x}_A = \bar{x}_B = 8$. We see in sample B most values are closer to the center of data. Although the range fails to measure this variation between the upper and lower observations, but it does have useful applications.

(2) Mean Deviation:

If the set of data $x_1, x_2, \cdots, x_n$ represent a finite sample of size n then the mean deviation denoted by M. D is defined by:

(a) $\text{M. D} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$ for ungrouped data,

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

(b) $\text{M. D} = \frac{\sum_{i=1}^{n} f_i|x_i - \bar{x}|}{\sum_{i=1}^{n} f_i}$ for grouped data,

where x_i is a class mark, f_i is a frequency, k is the number of C.I and $\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$.

$$\text{M. D} = \frac{\sum_{i=1}^{k} f_i|x_i - \bar{x}|}{\sum_{i=1}^{k} f_i}$$

$$= \frac{1}{40}[2 * |1.7 - 3.4125| + 1 * |2.2 - 3.4125| + \cdots + 3 * |4.7 - 3.415|] =$$

(3) Variance:

If the set of data $x_1, x_2, \cdots, x_n$ represent a finite sample of size n then the sample variance denoted by S^2 is defined by:

(a) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2]$ for ungrouped data,

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

(b) $S^2 = \frac{\sum_{i=1}^{k} f_i(x_i-\bar{x})^2}{\sum_{i=1}^{k} f_i - 1}$ for grouped data,

where x_i is a class mark, f_i is a frequency, k is the number of C.I and $\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$.

Some Properties of Sample Variance:

(i) If $y_i = x_i \pm k, i = 1,2,\dots,n$, where k is any constant, then ${S^2}_y = {S^2}_x$.

(ii) If $y_i = kx_i,\ i = 1,2,\dots,n$, where k is any constant, then ${S^2}_y = k^2{S^2}_x$.

(4) Standard Deviation:

The standard deviation is defined to be the positive square root of the sample variance.

Example:

Verify that the variance of the sample 4, 9, 3, 6, 4 is 5.7. Use this fact to find the variance of the samples:

(a) 12, 27, 9, 18, 12.

(b) 9, 14, 8, 11,9.

Solution:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{26}{5} = 5.2 \Rightarrow$$

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

$$= \frac{1}{4}[158 - 5 * (5.2)^2] = \frac{1}{4}[158 - 135.2] = \frac{22.8}{4} = 5.7.$$

(a) Let $x_1 = 4, x_2 = 9,\ x_3 = 3,\ x_4 = 6, x_5 = 4$ and let $y_1 = 12,\ y_2 = 27,\ y_3 = 9,\ y_4 = 18, y_5 = 12,$ then $y_i = 3x_i, i = 1,2,3,4,5.$

Therefore

$${S^2}_y = 9{S^2}_x = 9 * 5.7 = 51.3.$$

(b) Let $x_1 = 4, x_2 = 9, x_3 = 3, x_4 = 6, x_5 = 4$ and let $y_1 = 9, y_2 = 14, y_3 = 8, \quad y_4 = 11, y_5 = 9,$ then $y_i = x_i + 5, i = 1,2,3,4,5.$ Therefore

$$S^2{}_y = S^2{}_x = 5.7.$$

Exercises:

(1) If $\bar{x} = 6$, $S^2{}_x = 10$. Find $\bar{y}$ and $S^2{}_y$ if

(a) $y_i = 2x_i + 3.$

(b) $y_i = \frac{1}{2}x_i - 5.$

(2) The number of incorrect answers on a true-false test for a sample of 15 students are 2, 1, 3, 0, 1, 3, 6, 0, 3, 3, 5, 2, 1, 4, 2. Find the sample mean, median, mode and variance.

Al-Nahrain University - College of Science 1988

Probability

Definition: (Random Experiment)

Radom experiment is an experiment having the following properties:

(i) The outcome can not be predicated with certainty.

(ii) The outcome can be described prior to its performance.

(iii) It can be repeated under the same conditions.

Examples:

(1) Tossing a coin.

(2) Casting a die.

(3) Drawing a card from a desk of 52 playing cards.

(4) Selecting a number from the interval (0,1).

Definition: (Sample Space), (S.S)

Sample space is a set of all possible outcomes of a random experiment and is denoted by $\mathcal{A}, \Omega, \mathcal{S}$.

Examples:

(1) The S.S of a tossing a coin is $\Omega = \{w: w = T, H\}$.

(2) The S.S of a casting a die is $\Omega = \{x: x = 1, 2, 3, 4, 5, 6\}$. .

(3) The S.S of selecting a number from the interval (0,1) is $\Omega = \{x: 0 < x < 1\}$.

Definition: (Event)

Event is a subset of sample space and is denoted by A, B, C,…

Example:

Consider an experiment of casting a die and let x be the number shown by a die.

(1) Describe the S.$\mathcal{S}$ $\mathcal{A}$ of this experiment.

(2) List all the elements corresponding to an event A where x is even.

(3) List all the elements corresponding to an event B where x is odd.

(4) List all the elements corresponding to an event C where $x \geq 5$.

(5) List all the elements corresponding to an event D where x>6.

Solution:

(1) $\mathcal{A} = \{x: x = 1, 2, 3, 4, 5, 6\}$.

(2) A=$\{x: x = 2, 4, 6\}$.

(3) B=$\{x: x = 1, 3, 5\}$.

(4) C=$\{x: x = 5, 6\}$.

(5) D=$\emptyset$, where $\emptyset$ denote the empty set.

Example:

An experiment of casting a pair of die, one red and one blue and let the outcomes be the ordered pair (x,y), where x is the number occur on the red die and y where x is the number occur on the blue die.

(1) Describe the S.S $\mathcal{A}$ of this experiment.

(2) List all the elements corresponding to an event A where x+y>9.

(3) List all the elements corresponding to an event B where 2 occur on either die.

(4) List all the elements corresponding to an event C where x divides 2 and y>5.

(5) List all the elements corresponding to an event D where x+y is divisible by 5.

Solution:

(1) $\mathcal{A}$ ={(x,y): (x,y)=(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}.

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

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

(4) C={(2,6), (4,6), (6,6)}.

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



Counting Sample Points:

Multiplication Rule:

If an operation can be performed in n_1 ways and if for each of these (n_1 ways), a second operation can be performed in n_2 ways and if for each of these (n_1 and n_2 ways), a third operation can be performed in n_3 ways and so on, then the sequence of k operations can be performed in $n_1.n_2.n_3 \ldots\ldots n_k$ ways.

Examples:

(1) How many sample points are in the sample space when a pair of dies is thrown once?

Solution:

The first die can be land in 6 different ways and the second die can be land in 6 different ways. Therefore a pair of dies can be land in 6*6=36 ways, or

6	6

number of points in S.S= =36.

(2) How many lunches are possible consisting of soup, a sandwich, dessert and a drink if one can select from 4 soups, 3 kinds of sandwich, 5 desserts and 4 drinks?.

Solution:

Total number of lunches =

4	3	5	4

= 240.

(3)

(a) How many three digit numbers can be formed from the digits 2, 1, 5, 6, 9 if each digit can be used only once?.

Solution:

Total of three digit numbers = | 5 | 4 | 3 | = 60.

(b) How many of these are even numbers?.

Solution:

Total of even three digit numbers | 4 | 3 | 2 | = 24.

(c) How many of these are greater than 900?.

Solution:

Total of three digit numbers greater than 900 = | 1 | 4 | 3 | = 12.

(4) In how many different ways can true-false test consisting of 5 questions be answered?.

Solution:

number of ways = | 2 | 2 | 2 | 2 | 2 | $=2^5 = 32$.

Bahrain University - College of Science 1986

Permutation

Definition: (permutation)

A permutation is an order arrangement of all or part of a set of objects.

Theorem:

The number of all permutations of n distinct objects is n!

Theorem:

The number of all permutations of n distinct objects taken r at a time is

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

Theorem:

The number of all permutations of n distinct objects arranged in a circle is $(n-1)!$.

Remark:

If r = n, then $P_n^n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1} = n!$.

Theorem:

The number of distinct permutations of n things of which n_1 are of one kind, n_2 are of second kind, …, n_k are of k-th kind is $\frac{n!}{n_1!\, n_2! \ldots n_k!}$.

Example:

Write all order arrangements of 3 books A, B, C in a row?.

Solution:

The number of all permutations of 3 books = $P_3^3 = 3! = 6$, or

by multiplication rule = | 3 | 2 | 1 | = 6.

The 3 books can be arranged as: ABC, ACB, BAC, BCA, CAB, CBA.

Example:

Write all permutations of 2 books selected from the 4 books A, B, C, D in a row ?.

Solution:

The number of all permutations = $P_2^4 = \frac{4!}{(4-2)!} = \frac{4!}{2!} = 4 * 3 = 12,$ or

by multiplication rule = | 4 | 3 | =12.

Example:

(a) How many ways you can write 5 digits numbers from the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 if each digit can be used only once?

(b) How many of these start with 3?.

Solution

(a) number of ways = | 9 | 9 | 8 | 7 | 6 | = 27216 , or

number of ways = $P_1^9 . P_4^9 = \frac{9!}{8!} . \frac{9!}{5!} = 9 * (9 * 8 * 7 * 6) = 27216.$

(b) number of ways = | 1 | 9 | 8 | 7 | 6 | = 3024 , or

number of ways = $1 . P_4^9 = \frac{9!}{5!} = 9 * 8 * 7 * 6 = 3024.$

Example:

A shipment of 10 different transistors, 8 of them which are good and 2 are faulty. Suppose that 3 transistors are selected at random from the shipment.

(a) How many points are there in the S.S?.
(b) In how many ways that there are two good and one faulty transistors?.

Solution:

(a) The S.S contain $P_3^{10} = \frac{10!}{7!} = 10 * 9 * 8 = 720.$

(b) Number of ways $= P_2^8 . P_1^2 = \frac{8!}{6!} . \frac{2!}{0!} = 8 * 7 * 2 = 112.$

Example:

How many different ways can 3 red, 4 yellow and 2 blue bulbs be arranged in a string with 9 sockets ?.

Solution:

number of ways $= \frac{9!}{3!*4!*2!} = \frac{9*8*7*6*5}{6*2} = 9 * 4 * 7 * 5 = 1260.$

Example:

Find the number of permutations that can be performed from word "statistics"?.

Solution:

number of permutations $= \frac{9!}{3!*3!*2!*1!*1!} = \frac{9*8*7*6*5*4*3}{6*2} = 50400.$

Definition:

The number of ways of partitioning a set of n objects into k subsets with n_1 elements in the first subset, n_2 elements in the second subset,…, n_k elements in the k-th subset is

$$\binom{n}{n_1\ n_2\ \dots\ n_k} = \frac{n!}{n_1!*n_2!*\dots*n_k},$$

where $n = \sum_{i=1}^{k} n_i$.

Example:

How many ways can 7 visitors be distributed in one triple and two double room?.

Solution:

number of ways = $\binom{7}{3\ 2\ 2} = \frac{7!}{3!*2!*2!} = 7 * 6 * 5 = 210.$

Al-Nahrain University - College of Science 1988

Combination

Definition (Combination):

Suppose that we are interested in partitioning a set of n objects into two subsets, one subset containing r elements and the other subset containing $(n-r)$ elements. This partition is called combination and the number of all combinations in this case is

$$\binom{n}{r \quad n-r} = \frac{n!}{r!(n-r)!},$$

and denoted by $\binom{n}{r}$ or C_r^n.

Definition:

The number of all combinations of n distinct objects taken r at a time is

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

Example:

From 9 persons, find the number of committee of 3 persons that can be formed?.

Solution:

number of committer $= \binom{9}{3} = \frac{9!}{3!(9-3)!} = \frac{9!}{3!*6!} = 84.$

Example:

From 4 men and 3 women, find the number of committee of 3 persons that can be formed with 2 men and 1 women?.

Solution:

number of committer = $\binom{4}{2}.\binom{3}{1} = \frac{4!}{2!(4-2)!}.\frac{3!}{1!*2!} = \frac{4!}{2!*2!}.\frac{3!}{1!*2!}$=18.

Exercise:

In how many ways can two committees of 4 persons be performed from 24 persons if the chosen persons can not be in more than one committee?.

Example:

12 TV sets where 3 sets are defective. In how many ways one can buy 5 sets and receive at least 2 sets of the defective?.

Solution:

number of ways = $\binom{3}{2}.\binom{9}{1} + \binom{3}{3}.\binom{9}{2}$

Exercise:

In how many ways we can select a committee of 8 persons from a group of 30 men and 20 women if the committee involve:

(a) 3 men only.
(b) at least 5 men and at least 2 women.

The Classical Definition of Probability:

If a random experiment with sample space $\mathcal{A}$ can result in N different equally likely outcomes and if exactly n of these outcomes corresponding to an event A, then the probability of A, denoted by P(A) is defined by

$$P(A) = \frac{number\ of\ elements\ in\ A}{number\ of\ elements\ in\ \mathcal{A}} = \frac{n}{N}.$$

Remark:

Equally likely means that each outcome (element) in $\mathcal{A}$ have the same chance of occurrence.

Example:

A coin is tossed once. What is the probability that a head occurs?.

Solution:

The S.S is $\mathcal{A} = \{x: x = \text{T or H}\}$. Let A be the event of obtaining head , then A=$\{x: x = \text{H}\}$. Therefore

$$P(A) = \frac{number\ of\ elements\ in\ A}{number\ of\ elements\ in\ \mathcal{A}} = \frac{1}{2}.$$

Example:

Two different coins are tossed once. What is the probability that at least one head occurs?

Solution:

The S.S is $\mathcal{A} = \{\text{TT}, \text{TH}, \text{HT}, \text{HH}\}$. Let A be the event of getting at least one head, then A=$\{\text{TH}, \text{HT}, \text{HH}\}$. Therefore

$$P(A) = \frac{number\ of\ elements\ in\ A}{number\ of\ elements\ in\ \mathcal{A}} = \frac{3}{4}.$$

Example:

Three different coins are tossed once.

(a) What is the probability that two heads occur?.

(b) What is the probability that at least two heads occur?.

Solution:

The S.S is $\mathcal{A} = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}$.

(a) Let A be the event of obtaining two heads, then A={THH, HTH, HHT}.

Therefore

$$P(A) = \frac{number\ of\ elements\ in\ A}{number\ of\ elements\ in\ \mathcal{A}} = \frac{3}{8}.$$

(b) Let B be the event of obtaining at least two heads, then B={THH, HTH, HHT, HHH}. Therefore

$$P(B) = \frac{number\ of\ elements\ in\ B}{number\ of\ elements\ in\ \mathcal{A}} = \frac{4}{8} = \frac{1}{2}.$$

Example:

A die is tossed once.

(a) What is the probability of obtaining even number?.

(b) What is the probability of obtaining prime number?.

(c) What is the probability of obtaining perfect square number?.

(d) What is the probability of obtaining numbers that divisible by 3?.

(e) What is the probability of obtaining numbers greater than 6?.

Solution:

The S.S is $\mathcal{A} = \{1, 2, 3, 4, 5, 6\}$.

(a) Let A be the event of obtaining even number, then A=$\{2, 4, 6\}$. Therefore

$$P(A) = \frac{number\ of\ elements\ in\ A}{number\ of\ elements\ in\ \mathcal{A}} = \frac{3}{6} = \frac{1}{2}.$$

(b) Let B be the event of obtaining prime number, then B=$\{2, 3, 5\}$. Therefore

$$P(B) = \frac{number\ of\ elements\ in\ B}{number\ of\ elements\ in\ \mathcal{A}} = \frac{3}{6} = \frac{1}{2}.$$

(c) Let C be the event of obtaining perfect square number, then C=$\{1, 4\}$. Therefore

$$P(C) = \frac{number\ of\ elements\ in\ C}{number\ of\ elements\ in\ \mathcal{A}} = \frac{2}{6} = \frac{1}{3}.$$

(d) Let D be the event of obtaining a number that is divisible by 3, then D=$\{3, 6\}$. Therefore

$$P(D) = \frac{number\ of\ elements\ in\ D}{number\ of\ elements\ in\ \mathcal{A}} = \frac{2}{6} = \frac{1}{3}.$$

(e) Let E be the event of obtaining a number greater than 6, then E=$\emptyset$. Therefore

$$P(E) = \frac{number\ of\ elements\ in\ E}{number\ of\ elements\ in\ \mathcal{A}} = \frac{0}{6} = 0.$$

Example:

A letter is chosen at random from English alphabet.

(a) What is the probability of obtaining vowel letter?.

(b) What is the probability of obtaining letter precedes the letter j?.

(c) What is the probability of obtaining letter follows the letter j?.

Solution:

The S.S is $\mathcal{A} = \{a, b, c, \ldots, z\}$.

(a) Let A be the event of obtaining vowel letter, then A=$\{a, e, i, o, u\}$. Therefore

$$P(A) = \frac{number\ of\ elements\ in\ A}{number\ of\ elements\ in\ \mathcal{A}} = \frac{5}{26}.$$

(b) Let B be the event of obtaining letter precedes the letter j, then B={a, b, c, d, e, f, g, h, i}. Therefore

$$P(B) = \frac{number\ of\ elements\ in\ B}{number\ of\ elements\ in\ \mathcal{A}} = \frac{9}{26}.$$

(c) Let C be the event of obtaining letter follows the letter j, then C={k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z}. Therefore

$$P(C) = \frac{number\ of\ elements\ in\ C}{number\ of\ elements\ in\ \mathcal{A}} = \frac{16}{26} = \frac{8}{13}.$$

Example:

A shipment of 12 T.V sets in which 3 sets are defective. What is the probability that one buy 5 T.V sets and receive at least 2 defective sets?.

Solution:

Let A be the event of receiving at least 2 defective sets. The S.S $\mathcal{A}$ consists of $\binom{12}{5}$ equally likely outcomes. Therefore

$\left[\binom{3}{2}\binom{9}{3} + \binom{3}{3}\binom{9}{2}\right] \in A$, then

$$P(A) = \frac{\binom{3}{2}\binom{9}{3} + \binom{3}{3}\binom{9}{2}}{\binom{12}{5}} = \frac{4}{11}.$$

Kirkuk University - College of Science

Finite Sample Space with or Without Equally likely Outcomes:

The mathematical theory of probability provides an alternative approach for finding the probability of an event, when the finite sample space of a random experiment with or without equally likely outcomes.

Let the sample space of a random experiment be $\mathcal{A} = \{x_1, x_2, \dots, x_N\}$ and let $A = \{x_1, x_2, \dots, x_n\}$ be an event, n≤N. To every element $x_i \in \mathcal{A}$, we assign a probability α_i, i=1,2,…,N such that $\sum_{i=1}^{N} \alpha_i = \alpha_1 + \alpha_2 + \cdots + \alpha_N = 1$. To find the probability of an event, we sum all the probability assigned to the elements in A.

Example:

Two different coins are tossed once. What is the probability that at least one head occurs?

Solution:

The S.S is $\mathcal{A} = \{\text{TT}, \text{TH}, \text{HT}, \text{HH}\}$. Let A be the event of getting at least one head, then A=$\{\text{TH}, \text{HT}, \text{HH}\}$. In this example, all elements in v have the same chance of occurrence (equally likely), Therefore

$\alpha + \alpha + \alpha + \alpha = 1 \Rightarrow 4\alpha = 1 \Rightarrow \alpha = \frac{1}{4}$. But P(A)= $\alpha + \alpha + \alpha = 3\alpha$.

Therefore P(A)= $\frac{3}{4}$.

Example:

A die is loaded in such a way than an even number is twice as likely to occur as an odd number. If a die is tossed once.

(a) What is the probability of obtaining even number?.

(b) What is the probability of obtaining prime number?.

(c) What is the probability of obtaining a number less than 4?.

Solution:

The S.S is $\mathcal{A} = \{1, 2, 3, 4, 5, 6\}$.

(a) Let A be the event of obtaining even number, then A=$\{2, 4, 6\}$. Let α be the probability assigned to an odd number. Therefore 2α be the probability assigned to even number. Hence

$\alpha + 2\alpha + \alpha + 2\alpha + \alpha + 2\alpha = 1 \Rightarrow 9\alpha = 1 \Rightarrow \alpha = \frac{1}{9}$.

$$\mathrm{P(A)} = 2\alpha + 2\alpha + 2\alpha = 6\alpha = \frac{6}{9} = \frac{2}{3}.$$

(b) Let B be the event of obtaining prime number, then B=$\{2, 3, 5\}$. Therefore

$$\mathrm{P(B)} = 2\alpha + \alpha + \alpha = 4\alpha = \frac{4}{9}.$$

(c) Let C be the event of obtaining a number less than 4, then C=$\{1, 2, 3\}$. Therefore

$$\mathrm{P(C)} = \alpha + 2\alpha + \alpha = 4\alpha = \frac{4}{9}.$$

Example:

A five-sided die with sides numbered 1,2,3,4,5 constructed so that 1 and 5 occur twice as often as 2 and 4. Also 2 and 4 occur three times as often as 3. If a die is tossed once.

(a) What is the probability of obtaining odd number?.

(b) What is the probability of obtaining a number that is divisible by 2?.

(c) What is the probability of obtaining perfect square number?.

(d) What is the probability of obtaining a number between 2 and 5 inclusive?.

(e) What is the probability of obtaining a number greater than 6?.

Solution:

The S.S is $\mathcal{A} = \{1, 2, 3, 4, 5\}$.

(a) Let A be the event of obtaining odd number, then A={1,3,5}. Let α_i be the probability assigned to an element $\alpha_i \in \mathcal{A}, i = 1,2,3,4,5$. Then we have

$$\alpha_1 = \alpha_5 = 2\alpha_2 = 2\alpha_4$$

and

$$\alpha_2 = \alpha_4 = 3\alpha_3.$$

Therefore

$$\alpha_1 = \alpha_5,$$

$$\alpha_2 = \frac{1}{2}\alpha_1 = \alpha_4$$

and

$$\alpha_3 = \frac{1}{3}\alpha_2 = \frac{1}{6}\alpha_1.$$

But

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = 1 \Rightarrow \alpha_1 + \frac{1}{2}\alpha_1 + \frac{1}{6}\alpha_1 + \frac{1}{2}\alpha_1 + \alpha_1 = 1$$

$$\Rightarrow \frac{19}{6}\alpha_1 = 1.$$

Hence

$$\alpha_1 = \frac{6}{19} = \alpha_5,$$

$$\alpha_3 = \frac{1}{6}\alpha_1 = \frac{1}{6} * \frac{6}{19} = \frac{1}{19}$$

and

$$\alpha_2 = \alpha_4 = \frac{1}{2}\alpha_1 = \frac{1}{2} * \frac{6}{19} = \frac{3}{19}.$$

Thus

$$\mathrm{P(A)} = \alpha_1 + \alpha_3 + \alpha_5 = \frac{6}{19} + \frac{1}{19} + \frac{6}{19} = \frac{13}{19}.$$

(b) Let B be the event of obtaining a number that is divisible by 2, then B=$\{2, 4\}$. Therefore

$$\mathrm{P(B)} = \alpha_2 + \alpha_4 = \frac{3}{19} + \frac{3}{19} = \frac{6}{19}.$$

(c) Let C be the event of obtaining perfect square number, then C=$\{1, 4\}$. Therefore

$$\mathrm{P(C)} = \alpha_1 + \alpha_4 = \frac{6}{19} + \frac{3}{19} = \frac{9}{19}.$$

(d) Let D be the event of obtaining a number between 2 and 5 inclusive, then D=$\{2,3,4,5\}$. Therefore

$$\mathrm{P(D)} = \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = \frac{3}{19} + \frac{1}{19} + \frac{3}{19} + \frac{6}{19} = \frac{13}{19}.$$

(e) Let E be the event of obtaining a number greater than 6, then E=$\emptyset$. Therefore

$$\mathrm{P(E)} = 0.$$

Al-Nahrain University - College of Science 1988

The Probability Set Function and Axioms:

Let $\mathcal{A}$ be a simple space of a random experiment and let $A \subset \mathcal{A}$ be an event. We define a set function P(A) as the probability that the outcome of a random experiment is an element of A. Furthermore P(A) satisfy the following axioms:

(1) $P(A) \geq 0, \forall A \subset \mathcal{A}$.

(2) $P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + \cdots$ provided $A_i \cap A_j = \emptyset \ \forall i \neq j$.

(1) $P(\mathcal{A}) = 1$.

Theorem (1):

$$P(A) = 1 - P(A^c), \forall A \subset \mathcal{A} .$$

Proof:

Since $\mathcal{A} = A \cup A^c,\ A \cap A^c = \emptyset$, then

$$P(\mathcal{A}) = P(A \cup A^c) = P(A) + P(A^c).$$

But $P(\mathcal{A}) = 1$. Therefore

$$1 = P(A) + P(A^c)$$

and this implies that

$$P(A) = 1 - P(A^c), \forall A \subset \mathcal{A} .$$

Theorem (2):

$$P(\emptyset) = 0.$$

Proof:

Since $\mathcal{A} = \mathcal{A} \cup \emptyset,\ \mathcal{A} \cap \emptyset = \emptyset$, then

$$P(\mathcal{A}) = P(\mathcal{A} \cup \emptyset) = P(\mathcal{A}) + P(\emptyset).$$

But $P(\mathcal{A}) = 1$. Therefore

$$1 = 1 + P(\emptyset)$$

and this implies that

$$P(\emptyset) = 0.$$

Theorem (3):

If A and B are events with $A \subset B$, then

$$P(A) \leq P(B) .$$

Proof:

Let $C=A^c \cap B$. Since $B=A \cup C$ and $A \cap C=\emptyset$ then

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

But $P(C) \geq 0$. Therefore

$$P(B) \geq P(A).$$

Theorem (4):

$$0 \leq P(A) \leq 1 , \forall A \subset \mathcal{A}.$$

Proof:

Since $\emptyset \subset A \subset \mathcal{A}$, then

$$P(\emptyset) \leq P(A) \leq P(\mathcal{A}).$$

But $P(\mathcal{A}) = 1$ and $P(\emptyset)=0$. Therefore

$$0 \leq P(A) \leq 1 , \forall A \subset \mathcal{A}.$$

Theorem (5):

If A and B are events with $A \subset B$, then

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

Proof:

Let $C=A^c \cap B$. Then $A \cup B = A \cup C$ and $A \cap C=\emptyset$ thus

$$P(A \cup B) = P(A \cup C) = P(A) + P(C) \quad (1)$$

Also $B=(A \cap B) \cup C$ and But $(A \cap B) \cap C = \emptyset$. Therefore

$$P(B) = P\big((A \cap B) \cup C\big) = P(A \cap B) + P(C) \quad (2)$$

From equations (1)-(2), we get

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

Example:

Three candidates A, B, C are seeking public office. A and B are given the same chance of winning, but C is given twice the chance of either A or B.

(a) What is the probability that C win?.

(b) What is the probability that A does not win?.

Solution:

From the hypothesis, we have

$$P(A)=P(B) \text{ and } P(C)=2P(A)=2P(B).$$

But

$$1 = P(A) + P(B) + P(C) \Rightarrow 1 = P(A) + P(A) + 2P(A) \Rightarrow 1 = 4P(A).$$

Therefore

$$P(A) = \frac{1}{4} = P(B).$$

(a) $P(C) = 2P(A) = 2 * \frac{1}{4} = \frac{1}{2}$.

(b) $P(A^c) = 1 - P(A) = 1 - \frac{1}{4} = \frac{3}{4}$.

Example:

The probability that a student passes mathematics test is $\frac{2}{3}$ and the probability that he passes statistics test is $\frac{4}{9}$. If the probability that he passes at least one test is $\frac{4}{5}$, then find the probability that he will pass both tests

Solution:

Let A be the event that the student pass mathematics test and let B be the event that the statistics test. Then P(A)= $\frac{2}{3}$, P(B)= $\frac{4}{9}$ and P(A ∪ B) = $\frac{4}{5}$. The problem here is to find P(A ∩ B). But

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

Therefore

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{2}{3} + \frac{4}{9} - \frac{4}{5} = \frac{14}{45}.$$

Example:

The probability that a car A in use is 0.3, the probability that a car B in use is 0.25 and the probability that both cars in use is 0.15.

(a) What is the probability that at least one car in use?.

(b) What is the probability that neither car a nor car B in use?.

(c) What is the probability that only one car in use?.

Solution:

From the hypothesis, we have

$$P(A)=0.3 , \ P(B)=0.25 \text{ and } P(A \cap B) = 0.15.$$

(a) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.25 - 0.15 = 0.4$.

(b) $P(A^c \cap B^c) = P((A \cup B)^c) = 1 - P(A \cup B) = 1 - 0.4 = 0.6.$

(c) $P\big((A \cap B^c) \cup (A^c \cap B)\big) = P(A \cap B^c) + P(A^c \cap B)$

$= P(A) - P(A \cap B) + P(B) - P(A \cap B)$

$= P(A \cup B) - P(A \cap B) = 0.4 - 0.15 =$ 0.25.

Exercises:

(1) If A and B are mutually exclusive events with P(A)=0.3 and P(B)=0.6. Find $P(A^c)$, $P(A \cup B)$, $P(A^c \cap B^c)$ and $P(A^c \cap B)$.

(2) Let the sample space $\mathcal{A} = \{x: x = 1, 2, 3, 4, 5\}$. For each event $A \subset \mathcal{A}$, we define the probability set function as

$$P(A) = \sum_{x \in A} \frac{x}{15}.$$

If A={x: x=1, 2, 3} and B={x: x=2, 3, 4}, then find $P(A)$, $P(A^c)$, $P(B)$, $P(B^c)$, $P(A \cap B)$ and $P(A \cup B)$.

(3) Let the sample space $\mathcal{A} = \{x: 0 < x < \infty\}$. For each event $A \subset \mathcal{A}$, we define the probability set function as

$$P(A) = \int_{x \in A} e^{-x} dx.$$

If $A=\{x: 0 < x < 1\}$ and $B=\left\{x: \frac{1}{2} < x < 3\right\}$, then find $P(A)$, $P(A^c)$, $P(B)$, $P(B^c)$, $P(A \cap B)$ and $P(A \cup B)$.

Al-Mahrain University - College of Science 1981

Conditional Probability and Stochastic Independence:

Definition: (Conditional Probability)

Let A and B be two events of the sample space $\mathcal{A}$, the condirtional probability of an event A given an event B, denoted b $\mathrm{P}(\mathrm{A}\backslash\mathrm{B})$ is defined by

$$\mathrm{P}(\mathrm{A}\backslash\mathrm{B}) = \frac{\mathrm{P}(\mathrm{A} \cap \mathrm{B})}{\mathrm{P}(\mathrm{B})}, \text{provided } \mathrm{P}(\mathrm{B}) \neq 0.$$

Similarly

$$\mathrm{P}(\mathrm{B}\backslash\mathrm{A}) = \frac{\mathrm{P}(\mathrm{A} \cap \mathrm{B})}{\mathrm{P}(\mathrm{A})}, \text{provided } \mathrm{P}(\mathrm{A}) \neq 0.$$

Theorem (1):

(1) $0 \leq \mathrm{P}(\mathrm{A}\backslash\mathrm{B}) \leq 1 \; \forall A, B \subset \mathcal{A}$.

(2) $\mathrm{P}(\bigcup_{i=1}^{\infty} A_i \backslash \mathrm{B}) = \sum_{i=1}^{\infty} P(A_i \backslash B)$, $\forall \mathrm{A_i}, \mathrm{A_j} \subset \mathcal{A}$ and $\mathrm{A_i} \cap A_j = \emptyset \; \forall i \neq j$.

(3) $\mathrm{P}(\mathrm{A}\backslash\mathrm{A}) = 1, \; \emptyset \neq A \subset \mathcal{A}$.

(4) $\mathrm{P}(\mathrm{A}\backslash\mathrm{B}) = 1 - \mathrm{P}(\mathrm{A}^c\backslash\mathrm{B}) \; \forall A, B \subset \mathcal{A}$.

(5) $\mathrm{P}(\emptyset\backslash\mathrm{A}) = 0 \; \forall A \subset \mathcal{A}$.

(6) If $\mathrm{A} \subset \mathrm{B} \subset \mathcal{A}$, then $\mathrm{P}(\mathrm{A}\backslash\mathrm{C}) \leq \mathrm{P}(\mathrm{B}\backslash\mathrm{C})$.

(7) $\mathrm{P}(\mathrm{A} \cup \mathrm{B}\backslash\mathrm{C}) = \mathrm{P}(\mathrm{A}\backslash\mathrm{C}) + \mathrm{P}(\mathrm{B}\backslash\mathrm{C}) - \mathrm{P}(\mathrm{A} \cap B\backslash\mathrm{C}) \; \forall A, B, C \subset \mathcal{A}$.

Example:

Let D be the event that a plane departs on time and A be the event that a plane arrives on time. If P(D)=0.83, P(A)=0.92 and $\mathrm{P}(\mathrm{A} \cap D) = 0.78$.

(a) What is the probability that a plane arrives on time given it departed on time?.

(b) What is the probability that a plane departs on time given it arrived on time?.

Solution:

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

(b) $\mathrm{P(D\backslash A)} = \frac{\mathrm{P(A \cap D)}}{\mathrm{P(A)}} = \frac{0.78}{0.92} = 0.85.$

Example:

A box contains 8 chips, 3 chips are red and 5 chips are blue, two chips are drawn successfully at random without replacement. What is the probability that the first draw is red chip and the second draw is blue chip?.

Solution:

Let R be the event that the first draw is red chip and let B be the event that the second draw is blue chip. The problem is to find $\mathrm{P(R \cap B)}$.

From the hypothesis, we have $\mathrm{P(R)} = \frac{3}{8}$ and $\mathrm{P(B\backslash R)} = \frac{5}{7}$.

Since

$$\mathrm{P(B\backslash R)} = \frac{\mathrm{P(B \cap R)}}{\mathrm{P}(R)}$$

then

$$\mathrm{P(B \cap R)} = \mathrm{P(B\backslash R)} * \mathrm{P}(R) = \frac{5}{7} * \frac{3}{8} = \frac{15}{56}.$$

Product Rule:

From the definition of the conditional probability, we have:

$$\mathrm{P(A \cap B)} = \mathrm{P}(A)\mathrm{P(B\backslash A)} \text{ provided } \mathrm{P(A)} \neq 0.$$

Therefore

$$P(A \cap B \cap C) = P\big((A \cap B) \cap C\big) = P(A \cap B)P(C\backslash A \cap B)$$
$$= P(A)P(B\backslash A)P(C\backslash A \cap B)$$

provided $P(A) \neq 0$ and $P(A \cap B) \neq 0$.

So, in general, we have

$$P(A_1 \cap A_2 \cap \dots \cap A_n)$$
$$= P(A_1)P(A_2\backslash A_1)\, P(A_3\backslash A_1 \cap A_2) \dots P(A_n\backslash A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

provided $P(A_1) \neq 0$ and $P(A_1 \cap A_2) \neq 0$, $P(A_1 \cap A_2 \cap A_3) \neq 0, \dots, P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \neq 0$.

Example:

Three cards are drawn in successive without replacement from desk of 52 playing cards. Find the probability that the first card is red Ace and the second card is 10 or Jack and the third number is between 3 and 7 excusive?.

Solution:

Let A be the event that the first card is red Ace, let B be the event that the second card is 10 or Jack and let C be the event that the third card is a number between 3 and 7 excusive. The problem is to find $P(A \cap B \cap C)$.

From the hypothesis, we have

$$P(A)= \frac{2}{52}\,,\ P(B\backslash A) = \frac{8}{51} \text{ and } P(C\backslash A \cap B) = \frac{12}{50}.$$

Therefore

$$P(A \cap B \cap C) = P(A)P(B\backslash A)P(C\backslash A \cap B) = \frac{2}{52} * \frac{8}{51} * \frac{12}{50} = 0.002.$$

Bayes Theorem:

Let $A_1, A_2, \dots, A_n$ be mutually exclusive events of a sample space $\mathcal{A}$, (that is $A_i \cap A_j = \emptyset$ for $i \neq j$) and let B an event of $\mathcal{A}$. If $B \subset \bigcup_{i=1}^{n} A_i$, then

(i) $P(B) = \sum_{i=1}^{n} P(B\backslash A_i)\, P(A_i)$.

(ii) $\mathrm{P}(\mathrm{A}_i \backslash \mathrm{B}) = \frac{\mathrm{P}(\mathrm{B} \backslash \mathrm{A}_i)\mathrm{P}(\mathrm{A}_i)}{\sum_{i=1}^{n} \mathrm{P}(\mathrm{B} \backslash \mathrm{A}_i)\mathrm{P}(\mathrm{A}_i)}, i = 1,2, \dots, n.$

Proof:

(i) Since $\mathrm{A}_1, \mathrm{A}_2, \dots, \mathrm{A}_n$ be mutually exclusive events and $\mathrm{A}_1 \cap B \subset \mathrm{A}_1, \mathrm{A}_2 \cap B \subset \mathrm{A}_2, \dots, \mathrm{A}_n \cap B \subset \mathrm{A}_n$ then $\mathrm{A}_1 \cap B, \mathrm{A}_2 \cap B, \dots, \mathrm{A}_n \cap B$ are also mutually exclusive events. Since $B \subset \bigcup_{i=1}^{n} \mathrm{A}_i$ then we can write B as

$$\mathrm{B} = (\mathrm{A}_1 \cap B) \cup (\mathrm{A}_2 \cap B) \cup \dots \cup (\mathrm{A}_n \cap B).$$

Therefore

$$\mathrm{P}(\mathrm{B}) = P((\mathrm{A}_1 \cap B) \cup (\mathrm{A}_2 \cap B) \cup \dots \cup (\mathrm{A}_n \cap B))$$
$$= P(\mathrm{A}_1 \cap B) + P(\mathrm{A}_2 \cap B) + \cdots + P(\mathrm{A}_n \cap B) = \sum_{i=1}^{n} P(\mathrm{A}_i \cap B).$$

But

$$\mathrm{P}(\mathrm{B} \backslash \mathrm{A}_i) = \frac{\mathrm{P}(\mathrm{B} \cap \mathrm{A}_i)}{\mathrm{P}(\mathrm{A}_i)}, \text{ provided } \mathrm{P}(\mathrm{A_i}) \neq 0, \mathrm{i} = 1,2, \dots, \mathrm{n}.$$

This implies that

$$\mathrm{P}(\mathrm{B}) = \sum_{i=1}^{n} \mathrm{P}(\mathrm{B} \backslash \mathrm{A}_i)\, \mathrm{P}(\mathrm{A}_i).$$

(ii) $P(\mathrm{A}_i \backslash \mathrm{B}) = \frac{\mathrm{P}(\mathrm{A}_i \cap \mathrm{B})}{\mathrm{P}(\mathrm{B})} = \frac{\mathrm{P}(\mathrm{B} \backslash \mathrm{A}_i)\mathrm{P}(\mathrm{A}_i)}{\sum_{i=1}^{n} \mathrm{P}(\mathrm{B} \backslash \mathrm{A}_i)\mathrm{P}(\mathrm{A}_i)}, \ i = 1,2, \dots, n,$

provided $\mathrm{P}(\mathrm{B}) \neq 0$ and

$\mathrm{P}(\mathrm{A_i}) \neq 0, \mathrm{i} = 1,2, \dots, \mathrm{n}.$

Example:

There are three boxes:

Box I contain 3 red balls and 7 blue balls.

Box II contain 6 red balls and 4 blue balls.

Box III contain 2 red balls and 8 blue balls.

Box is selected at random and then one ball is drawn from the selected box.

(a) What is the probability that the ball is red?.

(b) If the ball is red, what is the probability that it drawn from Box II?.

Solution:

From the hypothesis, we have

$$P(I) = P(II) = P(III) = \frac{1}{3}, P(R\backslash I) = \frac{3}{10}, P(R\backslash II) = \frac{6}{10} \text{ and } P(R\backslash III) = \frac{2}{10}.$$

(a) $P(R) = P(R\backslash I)P(I) + P(R\backslash II)P(II) + P(R\backslash III)P(III)$

$$= \frac{3}{10} * \frac{1}{3} + \frac{6}{10} * \frac{1}{3} + \frac{2}{10} * \frac{1}{3} = \frac{11}{30}.$$

(b) $P(II\backslash R) = \frac{P(R\backslash II)P(II)}{P(R)} = \frac{\frac{6}{10} * \frac{1}{3}}{\frac{11}{30}} = \frac{6}{11}.$

Al Mahrain University - College of Science 1988

Independence

Definition: (Independence)

Two events A and B are said to be independent iff

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

Remark:

If A and B are independent events then

$$\mathrm{P(A\backslash B)} = \frac{\mathrm{P(A \cap B)}}{\mathrm{P(B)}} = \frac{P(A)P(B)}{\mathrm{P(B)}} = P(A), \text{provided } \mathrm{P(B)} \neq 0$$

and

$$\mathrm{P(B\backslash A)} = \frac{\mathrm{P(B \cap A)}}{\mathrm{P(A)}} = \frac{P(B)P(A)}{\mathrm{P(A)}} = P(B), \text{provided } \mathrm{P(A)} \neq 0.$$

Theorem :

If A and B are independent events then

(i) A and $\mathrm{B^c}$ are independent events.

(ii) $\mathrm{A^c}$ and B are independent events.

(iii) $\mathrm{A^c}$ and $\mathrm{B^c}$ are independent events.

Proof:

(i) To prove A and $\mathrm{B^c}$ are independent events, we must prove:

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

Since

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

Therefore

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

This implies that

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B)$$

$$= P(A)\big(1 - P(B)\big)P(A)P(B^c).$$

(ii) To prove A^c and B are independent events, we must prove:

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

(iii) To prove A^c and B^c are independent events, we must prove:

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

Now,Since

$P(A^c \cap B^c) = P((A \cup B)^c) = 1 - P(A \cup B) \quad = 1 - [P(A) + P(B) - P(A \cap B)]$

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

$$= P(A^c) - P(B)[1 - P(A)] = P(A^c) - P(B)P(A^c)$$

$$= P(A^c)[1 - P(B)] = P(A^c)P(B^c).$$

Example:

Let A and B be two events of the sample space $\mathcal{A}$ such that P(A)=0.5 and $P(A \cup B) = 0.6$.

(a) Find P(B) if A and B are mutually exclusive.

(b) Find P(B) if A and B are independent.

(c) Find P(B) if P(A\B)=0.4.

Solution:

(a) Since A and B are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$. Therefore

$$P(B) = P(A \cup B) - P(A) = 0.6 - 0.5 = 0.1.$$

(b) Since A and B are independent, then $P(A \cap B) = P(A)P(B)$. Therefore

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

$$= P(A) + P(B) - P(A)P(B) \Rightarrow$$

$$0.6 = 0.5 + P(B) - 0.5P(B) \Rightarrow 0.1 = 0.5P(B) \Rightarrow P(B) = \frac{0.1}{0.5}$$

$$= 0.2.$$

(c) $\mathrm{P(A\backslash B)} = \frac{\mathrm{P(A\cap B)}}{\mathrm{P(B)}} = \frac{P(A)+P(B)-P(A\cup B)}{\mathrm{P(B)}} \Rightarrow 0.4 = \frac{0.5+P(B)-0.6}{\mathrm{P(B)}} \Rightarrow$

$$0.4P(B) = P(B) - 0.1 \Rightarrow 0.1 = 0.6P(B) \Rightarrow P(B) = \frac{0.1}{0.6} = \frac{1}{6}.$$

Exercises:

(1) If A and B are independent events with $P(B) = \frac{1}{4}$ and $P(B) = \frac{1}{3}$. Find $P\big((\mathrm{A} \cap B^c) \cup (\mathrm{A^c} \cap B)\big)$.

(2) Let A and B be two events with $P(B) \neq 1$. Show that

$$P(A\backslash B^c) = \frac{P(A) - P(\mathrm{A} \cap B)}{1 - P(B)}.$$

(3) How many ways can 4 boys and 5 girls sit in a row if the boys and girls must alternate?

Ans: 2880

(4) (a) How many 3-digit numbers can be formed from the digits 0, 1, 2, 3, 4, 5, 6 if each digit used only once?.

(b) How many of these are odd numbers?.

(c) How many of these are greater than 330?. Ans:(a) 180, (b)75, (c) 105

(5) Prove that, if P(A\B)=P(B|A), then P(A)=P(B).

(6) If A and B are independent events with $P(A) = \frac{1}{3}$ and $P(B^c) = \frac{1}{4}$. Find $P(\mathrm{A^c} \cup B)$.

(7) Four married couples have bought 8 seats in a row for a concert. In how many different ways can they seated:

(a) With no restriction.

(b) If each couple is sit together.

(c) If all men sit together to the right of all the women?.

Ans:(a) 40320, (b) 384, (c) 576

(8) There are 5 boxes numbers 1 to 5. Each box contains 10 balls. Box I has I defective and 10-I non defective balls, i=1, 2, 3, 4, 5. First a box is selected at random and then a ball drawn from the selected box.

(a) What is the probability that a defective ball will be obtained?.

(b) If the obtained ball is defective, what is the probability that it comes from box number 4?.

Ans: (a) $\frac{3}{10}$, (b) $\frac{4}{15}$

(9) The probability that a man watches a T.V show is 0.4 and the probability that his wife watches the show is 0.5, the probability that the man watches the show given his wife does is 0.7.

(a) Find the probability that a couple watches the show.

(b) Find the probability that a wife watches the show given that her husband does.

(c) Find the probability that at least one person of the couple will watch the show. Ans: (a) 0.35, (b) 0.875, (c) 0.55

(10) In how many different ways can a true-false test consists of 7 questions be answered?.

Ans: 128.

(11) In a medical study patients classified according to whether they have blood type AB^+, AB^-, $0^+, 0^-$, A^+, A^-, B^+, B^- and also according to whether their blood pressure is low, normal and high. Find the number of ways in which a patient can be classified.

Ans: 24

(12) In an experiment consisting of throwing a die and then drawing a letter at random from English alphabet. How many points are in the sample space?.

Ans: 156

(13) Two digits are selected at random from the digits 1,2,…,12 if the sum is even then find the probability that both numbers are even?.

Ans. 0.5

Therefore

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

This implies that

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B)$$
$$= P(A)\big(1 - P(B)\big)P(A)P(B^c).$$

(iv)

Al Mahrain University - College of Science 1988