# Chapter 2
# REGRESSION AND CORRELATIONS

## Rank Correlation

This is the approach used to determine the degree of the relationship between two variables by ranking them

**There are two methods employed**:

1. Spearman's rank correlation coefficient.( $\rho$ )
2. Kendall's rank correlation coefficient ( $\tau$ )

***However, we shall strictly use only Spearman's rank Correlation approach***

***Spearman's rank Correlation Coefficient***

With this approach, the variables are ranked according to the level of importance or magnitudes of the scores.

**For example**, given the following scores, 90, 80, 85, 75, 60. Rank them with the highest score taking the rank of 1.

| Scores | 90 | 80 | 85 | 75 | 60 |
|---|---|---|---|---|---|
| Ranks | 1 | 3 | 2 | 4 | 5 |

Note: in cases where some scores appear more than once, we give positions to the scores and then add the positions and divide by the number of times the score(s) appear.

**For example**, given the following scores 20,30,15,40,30,15,10,30,20,45, rank them by giving a rank of 1 to the highest score

| Scores | 20 | 30 | 15 | 40 | 30 | 15 | 10 | 30 | 20 | 45 |
|---|---|---|---|---|---|---|---|---|---|---|
| Positions | 6 | 3 | 8 | 2 | 4 | 9 | 10 | 5 | 7 | 1 |
| Rank | 6.5 | 4 | 8.5 | 2 | 4 | 8.5 | 10 | 4 | 6.5 | 1 |

***Note***: For the rank of 30, we consider the positions 3, 4 and 5, hence the rank of 30

$$= \frac{3+4+5}{3} = \frac{12}{3} = 4$$

For the rank of 20, we consider the positions, 6 and 7, hence the rank of 20 = $\frac{6+7}{2} = 6.5$

For the rank of 15, we consider the positions, 8 and 9, hence the rank of 15 = $\frac{8+9}{2} = 8.5$

The spearman's rank correlation coefficient denoted by $\rho$ is defined as

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

Where $d$ = the difference between the rankings of the scores

$n$ = total number of pairs of the scores.

***Examples***

**1.** The table below shows the marks obtained by 8 students in math and physics tests.

| Math(M) | 60 | 80 | 75 | 85 | 68 | 90 | 95 | 78 |
|---|---|---|---|---|---|---|---|---|
| Physics(P) | 70 | 75 | 80 | 78 | 85 | 90 | 96 | 83 |

Find the spearman's rank correlation coefficient and comment on your result.

***Solution***

| Math(M) | Physics(P) | $R_M$ | $R_P$ | D | $D^2$ |
|---|---|---|---|---|---|
| 60 | 70 | 8 | 8 | 0 | 0 |
| 80 | 75 | 4 | 7 | -3 | 9 |
| 75 | 80 | 6 | 5 | 1 | 1 |
| 85 | 78 | 3 | 6 | -3 | 9 |
| 68 | 85 | 7 | 3 | 4 | 16 |
| 90 | 90 | 2 | 2 | 0 | 0 |
| 95 | 96 | 1 | 1 | 0 | 0 |
| 78 | 83 | 5 | 4 | 1 | 1 |
| TOTAL | | | | | 36 |

$$\rho = 1 - \frac{6\times 36}{8(64-1)} = 1 - \frac{216}{504} = 0.57$$

**Comment** $\rho$ = 0.57 shows that there is a moderate positive correlation between Math and Physics

**2.** The table below shows the marks obtained by 10 students in History and Geography tests

| History | 80 | 80 | 70 | 60 | 65 | 80 | 68 | 90 | 95 | 50 |
|---|---|---|---|---|---|---|---|---|---|---|
| Geography | 50 | 45 | 70 | 80 | 70 | 90 | 70 | 80 | 70 | 95 |

Find the spearman's rank correlation coefficient and comment on your result.

***Solution***

| History(H) | Geography(G) | $R_H$ | $R_G$ | D | $D^2$ |
|---|---|---|---|---|---|
| 70 | 50 | 5.5 | 9 | -3.5 | 12.25 |
| 80 | 45 | 3.5 | 10 | -6.5 | 42.25 |
| 70 | 70 | 5.5 | 6.5 | -1.0 | 1 |
| 60 | 80 | 9 | 3.5 | 5.5 | 30.25 |
| 65 | 70 | 8 | 6.5 | 1.5 | 2.25 |
| 80 | 90 | 3.5 | 2 | 1.5 | 2.25 |
| 68 | 70 | 7 | 6.5 | 0.5 | 0.25 |
| 90 | 80 | 2 | 3.5 | -1.5 | 2.25 |
| 95 | 70 | 1 | 6.5 | -5.5 | 30.25 |
| 50 | 95 | 10 | 1 | 9 | 81 |
| TOTAL | | | | | 204 |

$$\rho = 1 - \frac{6\times 204}{10(100-1)}$$

$$= 1 - \frac{1224}{990} = {}^{-}0.2364$$

**Comment**: ρ = -0.2364 shows that there is a low/weak negative correlation between history and geography

**3.** An examination body carried out a research to find out the relationship between the mock results and end of year examination results for senior 4 candidates. Ten students of a certain school were sampled. The table below shows scores of the students in grade (aggregate) form for the best done eight subjects.

| | Mock exam | Final exam |
|---|---|---|
| A | 15 | 21 |
| B | 20 | 16 |
| C | 54 | 40 |
| D | 36 | 35 |
| E | 40 | 16 |
| F | 35 | 20 |
| G | 16 | 13 |
| H | 36 | 20 |

| | | |
|---|---|---|
| I | 18 | 30 |
| J | 40 | 25 |

Calculate the rank correlation coefficient for the performance of students in the two examinations. Comment on your result.

***Solution***

$$\rho = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 102}{10\times 99}$$

$$= 1 - \frac{612}{990} = 0.38$$

**Comment**: There is a weak positive relationship between mock exams and final exams

## *Conclusions based on the calculated value and the statistical value given from the statistical tables basing on significance levels ($\alpha$)*

## Important points to note

- If the calculated value $r_C$ is greater than the statistical value from the statistical tables, $r_T$
  I.e if $r_C > r_T$, we conclude that the relationship is significant at the given level of significance.

- On the other hand, If the calculated value $r_C$ is less than the statistical value from the statistical tables, $r_T$
  I.e if $r_C < r_T$, we conclude that the relationship is insignificant at the given level of significance.
- The statistical values from the tables may be given. But if not given, then we read them directly from the mathematical tables. This is done by taking into consideration the number of pairs, n and the level of significance.
  For example if number of pairs, n = 8 and ***$\alpha$ =5%*** , here spearman's rank correlation coefficient (table value) = 0.71 and the Kendall's rank correlation coefficient (table value) = 0.64

***Note.** The proof of the above conclusions is outside the scope of our coverage.*

**Examples**:

**1.** Applicants for a job with a company are interviewed by two of the personnel staff. After the interviews, each applicant is awarded a mark by each of the interviewers. The marks are given as below.

**Candidates**

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Interviewer1 | 22 | 27 | 24 | 17 | 20 | 22 | 16 | 13 |
| Interviewer2 | 28 | 23 | 25 | 14 | 26 | 17 | 20 | 15 |

a) Calculate to 2 d.p, the spearman's rank correlation coefficient between the two sets of marks
b) Using a 5% level of significance, interpret your result.

***Solution***

| | Interviewer 1 | Interviewer 2 | $R_1$ | $R_2$ | d | $d^2$ |
|---|---|---|---|---|---|---|
| A | 22 | 28 | 3.5 | 1 | 2.5 | 6.25 |
| B | 27 | 23 | 1 | 4 | -3 | 9 |
| C | 24 | 25 | 2 | 3 | -1 | 1 |
| D | 17 | 14 | 6 | 8 | -2 | 4 |
| E | 20 | 26 | 5 | 2 | 3 | 9 |
| F | 22 | 17 | 3.5 | 6 | -2.5 | 6.25 |
| G | 16 | 20 | 7 | 5 | 2 | 4 |
| H | 13 | 15 | 8 | 7 | 1 | 1 |
| Total | | | | | | 40.5 |

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

$$= 1 - \frac{6\times 40.5}{8\times 63} = \frac{261}{504} = 0.52$$

b) Reading the critical value of $r_T$ from the tables $\frac{59}{120}$

$r_T = 0.71$

Here, $\rho = 0.52 < r_T = 0.71$

Hence we conclude that the relationship is insignificant at 5% level of significance.

**2.** The course work grades ranging from A to G and examination marks of 8 candidates are given below.

| Course work grade | Examination mark |
|---|---|
| A | 92 |
| C | 75 |
| D | 63 |
| B | 54 |
| F | 48 |
| C | 45 |
| G | 34 |
| E | 18 |

a) Calculate the spearman's rank correlation coefficient for the two sets of data
b) Comment on the relationship between course work grades and examination grades at 5% significance level

***Solution***

| Course work grade (C) | Examination Mark (E) | $R_C$ | $R_E$ | D | $D^2$ |
|---|---|---|---|---|---|
| A | 92 | 1 | 1 | 0 | 0 |
| C | 75 | 3.5 | 2 | 1.5 | 2.25 |
| D | 63 | 5 | 3 | 2 | 4 |
| B | 54 | 2 | 4 | -2 | 4 |
| F | 48 | 7 | 5 | 2 | 4 |
| C | 45 | 3.5 | 6 | -2.5 | 6.25 |
| G | 34 | 8 | 7 | 1 | 1 |
| E | 18 | 6 | 8 | 2 | 4 |
| TOTAL | | | | | 25.5 |

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

$$= 1 - \frac{6\times 25.5}{8(64-1)}$$

$$= 1 - \frac{153}{504} = \frac{351}{504} = 0.696$$

b) From the tables, $r_T = 0.71$

***Conclusion***:

Since $\rho$ = 0.696 < 0.711, we conclude that the relationship between the two is insignificant at 5% level of significance.

**3.** Bird abundance may be accused in several ways. In one long -term study in a nature reserve, two independent surveys (A and B) are carried out. The data show the number if fowl territories recorded (survey A) and the number of adult of fowls trapped in a fine mesh net (survey B) over a number of years.

| Survey A | Survey B |
|---|---|
| 16 | 11 |
| 19 | 12 |
| 27 | 15 |
| 50 | 18 |
| 60 | 22 |
| 70 | 35 |
| 79 | 35 |
| 79 | 71 |
| 84 | 46 |
| 85 | 53 |
| 97 | 52 |

a) Calculate the spearman's rank correlation coefficient
b) Using a 5% level of significance, interpret your result.

***Solution***

| Survey A | Survey B | $R_A$ | $R_B$ | $D^2$ |
|---|---|---|---|---|
| 16 | 11 | 11 | 11 | 0 |
| 19 | 12 | 10 | 10 | 0 |
| 27 | 15 | 9 | 9 | 0 |
| 50 | 18 | 8 | 8 | 0 |
| 60 | 22 | 7 | 7 | 0 |
| 70 | 35 | 6 | 5.5 | 0.25 |
| 79 | 35 | 4.5 | 5.5 | 1 |
| 79 | 71 | 4.5 | 1 | 12.25 |
| 84 | 46 | 3 | 4 | 1 |
| 85 | 53 | 2 | 2 | 0 |
| 97 | 52 | 1 | 3 | 4 |
| TOTAL | | | | 18.5 |

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

$$=1-\frac{6\times 18.5}{11(121-1)}=1-\frac{111}{1320}$$

$$=\frac{1209}{1320}=0.916$$

b) From the tables, $r_T = 0.60$

***Conclusion***

Since $\rho = 0.916 > 0.6$, we conclude that the relationship between the two is significant at 5% level of significance.

# Examination Questions

**1.** (a) The price of Matooke is found to depend on the distance the market is away from the nearest town. The table below gives the average price of Matooke from markets around Kampala City

| Distance, $d$ (km) | Price, $P$ (shs) |
|---|---|
| 40 | 120 |
| 8 | 160 |
| 17 | 140 |
| 20 | 130 |
| 24 | 135 |
| 30 | 125 |
| 10 | 150 |
| 28 | 130 |
| 16 | 145 |
| 36 | 125 |

i) Plot these data on a scatter diagram.
ii) Draw the line of best fit on your diagram.
iii) Find the equation of your line in the form $p=\alpha+\beta d$

Where $\alpha$ and $\beta$ are constants

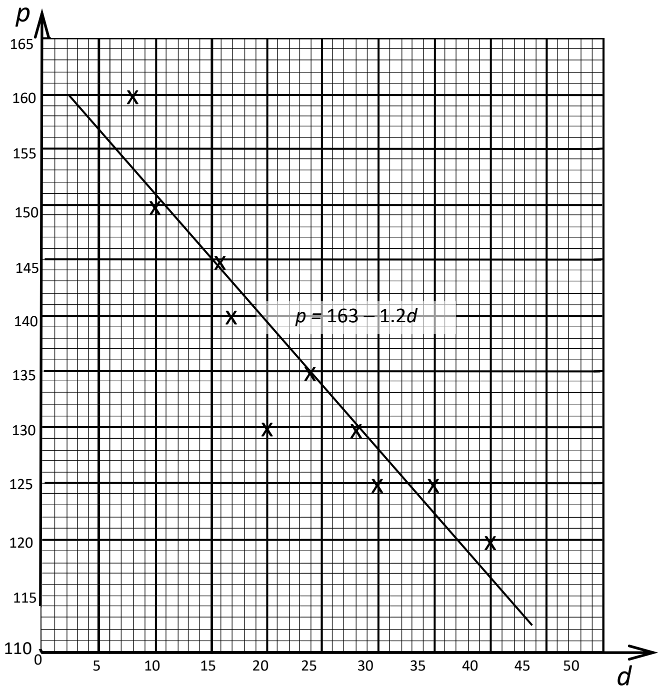
Hence estimate the price of matooke when d = 5

b) The following table gives the order in which six candidates were ranked in two tests $x$ and $y$

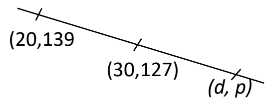
$x$: E, C, B, F, D, A
$y$: F, A, D, E, C, C

Calculate the coefficient of rank correlation and comment on your result. **(1988 No. 13)**

***Solution***




(iii) Picking any two points lying on the line, (20,139) and (30,127)




Using gradient approach

$$\frac{127-139}{30-20}=\frac{p-127}{d-30}$$

$$\frac{-12}{10}=\frac{p-127}{d-30}$$

$$p=-\frac{12}{10}(d-30)+127$$

$$p=-1.2d+163$$

Substituting for d = 5

$P$ = -1.2(5) + 163

$P$ = 157

**Using Spearman's method**

| | A | B | C | D | E | F | TOTAL |
|---|---|---|---|---|---|---|---|
| $X$ | 6 | 3 | 2 | 5 | 1 | 4 | |
| $Y$ | 2 | 5.5 | 5.5 | 3 | 4 | 1 | |
| d | 4 | -2.5 | -3.5 | 2 | -3 | 3 | |
| $d^2$ | 16 | 6.25 | 12.25 | 4 | 9 | 9 | 56.5 |

$$\rho = 1 - \frac{6\sum d^2}{n\left(n^2-1\right)} = 1 - \frac{6\times 56.5}{6\left(36-1\right)}$$

$$= 1 - \frac{339}{210} = \frac{-129}{210} = {}^{-}0.614$$

**Comment**: There is moderate (substantial) negative correlation between tests $X$ and $Y$.

**2.** The pairs of observations have been made on two random variables $X$ and $Y$. The ten $(x, y)$ values are (0, 20), (-7, 12), (-10, 15), (-12, 22), (-17, 5), ($^-$30,$^-$5), ($^-$32, 13), (10, 30), (15, 40), and ($^-$12, 8)

a) Plot these results on a scatter diagram

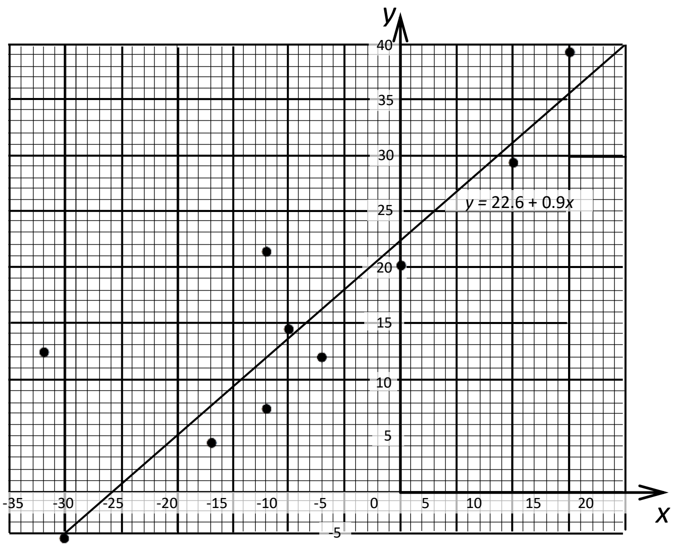
b) Draw on a scatter diagram the line of best fit for predicting $Y$ from $X$

c) Estimate the expected value of $Y$ corresponding to

$X$ = -7

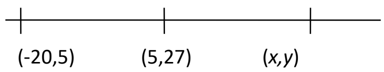
d) Calculate the rank correlation coefficient for these date (**1990 No. 12 modified)**

***Solution***




(b) Finding equation of $y$ on $x$:

Picking any two points on the line, (-20,5) and (5,27)




Using gradient approach

$$\frac{27-5}{5-{}^{-}20} = \frac{y-27}{x-5}$$

$$\frac{22}{25} = \frac{y-27}{x-5}$$

$$y = 27 + \frac{22}{25}(x-5)$$

$$y = 22.6 + 0.88x$$

(c) Substituting for $x$ = -7

$y$ = 22.6 + 0.88(-7)

$y$ = 16.44

d) Finding spearman's rank correlation coefficient

| $X$ | $Y$ | $R_X$ | $R_Y$ | $D$ | $D^2$ |
|---|---|---|---|---|---|
| 0 | 20 | 3 | 4 | -1 | 1 |
| -7 | 12 | 4 | 6 | -2 | 4 |
| -10 | 15 | 5 | 5 | 0 | 0 |
| -12 | 22 | 6.5 | 3 | 3.5 | 12.25 |
| -17 | 5 | 8 | 8 | 0 | 0 |
| -30 | -5 | 9 | 9 | 0 | 0 |
| -32 | -13 | 10 | 10 | 0 | 0 |
| 10 | 30 | 2 | 2 | 0 | 0 |
| 15 | 40 | 1 | 1 | 0 | 0 |
| -12 | 8 | 6.5 | 7 | 0.5 | 0.25 |
| TOTAL | | | | | 17.5 |

$$\rho = 1 - \frac{6\sum d^2}{n\left(n^2-1\right)} = 1 - \frac{6\times 17.5}{10\times 99} = \frac{885}{990} = 0.894$$

**3.** Three examiners $X$, $Y$ and $Z$ each marked the script of ten candidates who sat for a mathematics examination. The table below shows the examiners' ranking of the candidates.

Candidates

| Examiner | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 8 | 5 | 9 | 2 | 10 | 1 | 7 | 6 | 3 | 4 |
| $Y$ | 5 | 3 | 6 | 1 | 4 | 7 | 2 | 10 | 8 | 9 |
| $Z$ | 6 | 3 | 7 | 2 | 5 | 4 | 1 | 10 | 9 | 8 |

Calculate the coefficient of rank correlation of the rankings between

i) $X$ and $Y$

ii) $Y$ and $Z$

State with reason whether there is a significant difference between rankings of the three examiners

**(1991 No. 12)**

(i)

| $R_X$ | $R_Y$ | $D$ | $D^2$ |
|---|---|---|---|
| 8 | 5 | 3 | 9 |
| 5 | 3 | 2 | 4 |
| 9 | 6 | 3 | 9 |
| 2 | 1 | 1 | 1 |
| 10 | 4 | 6 | 36 |
| 1 | 7 | -6 | 36 |
| 7 | 2 | 5 | 25 |
| 6 | 10 | -4 | 16 |
| 3 | 8 | -5 | 25 |
| 4 | 9 | -5 | 25 |
| TOTAL | | | 186 |

$$\rho = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 186}{10\times 99} = \frac{-126}{99} = {}^{-}0.127$$

(ii)

| $R_Y$ | $R_Z$ | D | $D^2$ |
|---|---|---|---|
| 5 | 6 | -1 | 1 |
| 3 | 3 | 0 | 0 |
| 6 | 7 | -1 | 1 |
| 1 | 2 | -1 | 1 |
| 4 | 5 | -1 | 1 |
| 7 | 4 | 3 | 9 |
| 2 | 1 | 1 | 1 |
| 10 | 10 | 0 | 0 |
| 8 | 9 | -1 | 1 |
| 9 | 8 | 1 | 1 |
| TOTAL | | | 16 |

$$\rho = 1 - \frac{6\times 16}{10\times 99} = \frac{894}{990} = 0.903$$

**Giving reasons basing on statistical proofs**: Considering level of significance, $\alpha$ = 5%

From statistical tables, $\rho$ = 0.65

(i) $|\rho| = 0.127 < 0.65$ Hence, we conclude that at 5% level, there is insignificant relationship between the rankings made by X and Y.

ii) $|\rho| = 0.903 > 0.65$. Hence we conclude that there is sufficient evidence at 5% level of significance to show that there is significant relationship between the rankings made by Y and Z.

**4.** Three weighing scales from three different stalls, W,X and Y in Owino market were used to weigh 10 bags of beans A,B,C…J and the results (in kgs) were as given in the table below

One of the scales was known to be in good working condition.

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| Scale W | 65 | 68 | 70 | 63 | 64 | 62 | 73 | 75 | 72 | 78 |
| Scale X | 63 | 68 | 68 | 60 | 65 | 60 | 72 | 73 | 70 | 66 |
| Scale Y | 63 | 74 | 78 | 75 | 64 | 73 | 79 | 70 | 67 | 79 |

Determine rank correlation coefficient for the performances of the scales

i) *W* and *X*

j) *X* and *Y*.

Which of the three scales: W, X and Y were in good working conditions. **(1992 No. 13)**

***Solution***

| $R_W$ | $R_X$ | D | $D^2$ |
|---|---|---|---|
| 7 | 8 | -1 | 1 |
| 6 | 4.5 | 1.5 | 2.25 |
| 5 | 4.5 | 0.5 | 0.25 |
| 9 | 9.5 | -0.5 | 0.25 |
| 8 | 7 | 1 | 1 |
| 10 | 9.5 | 0.5 | 0.25 |
| 3 | 2 | 1 | 1 |
| 2 | 1 | 1 | 1 |
| 4 | 3 | 1 | 1 |
| 1 | 6 | -5 | 25 |
| TOTAL | | | 33 |

$$\rho = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 33}{10\times 99} = \frac{792}{990} = 0.8$$

(i)

| $R_X$ | $R_Y$ | D | $D^2$ |
|---|---|---|---|
| 8 | 10 | -2 | 4 |
| 4.5 | 5 | -0.5 | 0.25 |
| 4.5 | 3 | 1.5 | 2.25 |
| 9.5 | 4 | 5.5 | 30.25 |
| 7 | 9 | -2 | 4 |
| 9.5 | 6 | 3.5 | 12.25 |
| 2 | 1.5 | 0.5 | 0.25 |
| 1 | 7 | -6 | 36 |
| 3 | 8 | -5 | 25 |
| 6 | 1.5 | 4.5 | 20.25 |
| TOTAL | | | 134.5 |

$$\rho = 1 - \frac{6\times 134.5}{10\times 99} = \frac{183}{990} = 0.185$$

The scales in good working conditions are *W* and *X* as they have high positive correlation.

**5.** Ten shops in Kampala which attract a similar number and type of customers are ranked in terms of quality of service, size of veranda and price of items. Rank 1 indicates best service, largest veranda and lowest price of commodities. The results, including monthly average sales are given as below.

| Shop | Quality Of service | Size of veranda | Price of commodities | Sales (in terms of kg) |
|---|---|---|---|---|
| A | 3 | 3 | 6 | 20 |
| B | 7 | 5 | 10 | 10 |
| C | 4 | 10 | 7 | 31 |
| D | 6 | 7 | 2 | 47 |
| E | 8 | 2 | 4 | 37 |
| F | 2 | 1 | 5 | 38 |
| G | 5 | 8 | 3 | 38 |
| H | 9 | 6 | 8 | 15 |
| I | 10 | 4 | 10 | 21 |
| J | 1 | 9 | 1 | 42 |

**a)** By calculation, determine whether the price of commodities or size of the veranda is the more important factor affecting sales. **(1993 No. 14)**

**b)** Is there any evidence that the size of the veranda influences the quality of service?

**c)** Is there evidence that a shop with quality lower priced commodities offer poor quality services (e.g. by employing few sales people)

***Solution***

Let $q$ = quality of services

$V$ = size of veranda ,

$C$ = price of the commodities

$S$ = sales

a) **Considering size of the veranda and sales**

| Shop | V | S | $R_V$ | $R_S$ | D | $D^2$ |
|---|---|---|---|---|---|---|
| A | 3 | 20 | 3 | 8 | -5 | 25 |
| B | 5 | 10 | 5 | 10 | -5 | 25 |
| C | 10 | 31 | 10 | 6 | 4 | 16 |
| D | 7 | 47 | 7 | 1 | -6 | 36 |
| E | 2 | 37 | 2 | 5 | -3 | 9 |
| F | 1 | 38 | 1 | 3.5 | -2.5 | 6.25 |
| G | 8 | 38 | 8 | 3.5 | -4.5 | 20.25 |
| H | 6 | 15 | 6 | 9 | -3 | 9 |
| I | 4 | 21 | 4 | 7 | -3 | 9 |
| J | 9 | 42 | 9 | 2 | 7 | 49 |
| TOTAL | | | | | | 204.5 |

Spearman's rank correlation coefficient

$$\rho = 1 - \frac{6\times 204.5}{10\times 99} = 1 - \frac{1227}{990} = {}^{-}0.2394$$

**Considering price of commodities and sales**

| **Shop** | **C** | **S** | **$R_C$** | **$R_S$** | **D** | **$D^2$** |
|---|---|---|---|---|---|---|
| A | 6 | 20 | 6 | 8 | -2 | 4 |
| B | 10 | 10 | 10 | 10 | 0 | 0 |
| C | 7 | 31 | 7 | 6 | 1 | 1 |
| D | 2 | 47 | 2 | 1 | 1 | 1 |
| E | 4 | 37 | 4 | 5 | -1 | 1 |
| F | 5 | 38 | 5 | 3.5 | 1.5 | 2.25 |
| G | 3 | 38 | 3 | 3.5 | -0.5 | 0.25 |
| H | 8 | 15 | 8 | 9 | -1 | 1 |
| I | 10 | 21 | 10 | 7 | 3 | 9 |
| J | 1 | 42 | 1 | 2 | -1 | 1 |
| TOTAL | | | | | | 20.5 |

Spearman's rank correlation coefficient

$$\rho = 1 - \frac{6 \times 20.5}{10 \times 99} = 1 - \frac{123}{990} = 0.8758$$

Since there is a highly positive correlation between low prices of commodities and sales, therefore, it is the prices of the commodities other than size of the veranda that affects sales

(b) **Considering size of the veranda and quality of services**

| Shop | $R_V$ | Rq | D | $D^2$ |
|---|---|---|---|---|
| A | 3 | 3 | 0 | 0 |
| B | 5 | 7 | -2 | 4 |
| C | 10 | 4 | 6 | 36 |
| D | 7 | 6 | 1 | 1 |
| E | 2 | 8 | -6 | 36 |
| F | 1 | 2 | -1 | 1 |
| G | 8 | 5 | 3 | 9 |
| H | 6 | 9 | -3 | 9 |
| I | 4 | 10 | -6 | 36 |
| J | 9 | 1 | 8 | 64 |
| TOTAL | | | | 196 |

$$\rho = 1 - \frac{6 \times 196}{10 \times 99} = 1 - \frac{1176}{990} = {}^{-}0.188$$

Considering 5% level of significance,

From tables, $|\rho| = 0.65$

*Since* $|\rho| = 0.188 < 0.65$,

***Conclusion*** *We therefore, conclude that there is insignificant relationship that the size of the veranda influences the quality of service*

NB: Try using 1% level of significance

(c) **Considering prices and quality of services**

| Shop | $R_P$ | Rq | D | $D^2$ |
|---|---|---|---|---|
| A | 6 | 3 | 3 | 9 |
| B | 10 | 7 | 3 | 9 |
| C | 7 | 4 | 3 | 9 |
| D | 2 | 6 | -4 | 16 |
| E | 4 | 8 | -4 | 16 |
| F | 5 | 2 | 3 | 9 |
| G | 3 | 5 | -2 | 4 |
| H | 8 | 9 | -1 | 1 |
| I | 10 | 10 | 0 | 0 |
| J | 1 | 1 | 0 | 0 |
| TOTAL | | | | 73 |

$$\rho = 1 - \frac{6 \times 73}{10 \times 99} = 1 - \frac{438}{990} = 0.558$$

Considering 5% level of significance,

From tables, $|\rho| = 0.65$

*Since* $|\rho| = 0.558 < 0.65$

*We therefore, conclude that, there is insignificant relationship to conclude that a shop with lower priced commodities offer poor quality services*

**6.** (a) In many Government institutions, officers complain about typing errors. A test was designed to investigate the relationship between typing speed and errors made. Twelve typists A, B, C, D,.….L were picked at random to type the text. The table below shows the rankings of the typists according to speed and errors made. (N.B, lowest ranking in errors implies least errors made).

| Typist | Speed | Errors |
|---|---|---|
| A | 3 | 2 |
| B | 4 | 6 |
| C | 2 | 5 |
| D | 1 | 1 |
| E | 8 | 10 |
| F | 11 | 9 |
| G | 10 | 8 |
| H | 6 | 3 |
| I | 7 | 4 |
| J | 12 | 12 |
| K | 5 | 7 |
| L | 9 | 11 |

Calculate the rank correlation coefficient. Test the assertion made by the officers and comment on your result. ($\rho$ = 0.71 at 1% level of significance based on 12 observations.)

(b). The cost of travelling a certain distance a way from the city centre is found to depend on the route and the distance a given place is a way from the centre. The table below gives the average rates of travel charged for distances to be travelled a way from the city centre.

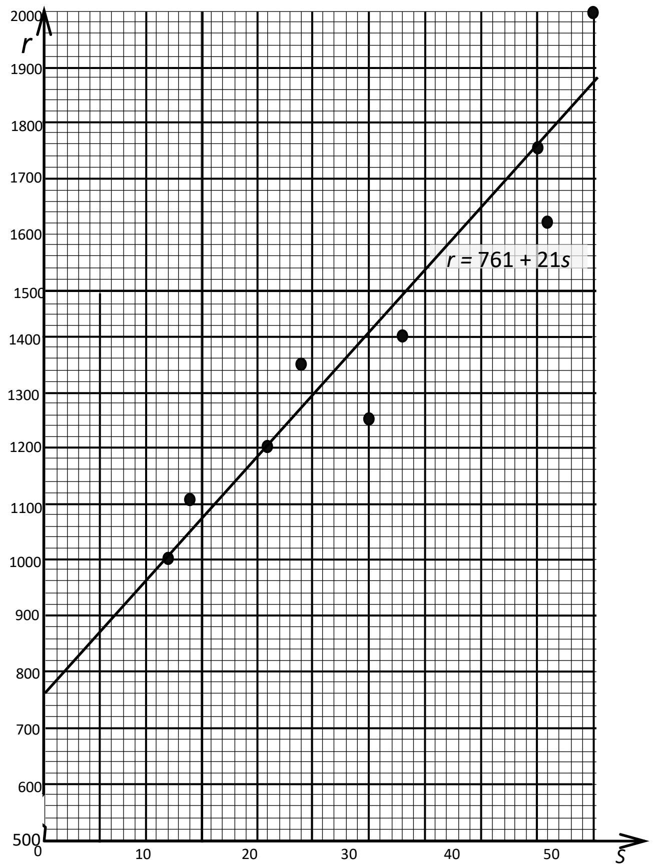
| Distance, S (km) | Rate charged, r (shs) |
|---|---|
| 9 | 750 |
| 12 | 1000 |
| 14 | 1150 |
| 21 | 1200 |
| 24 | 1350 |
| 30 | 1250 |
| 33 | 1400 |
| 45 | 1750 |
| 46 | 1600 |
| 50 | 2000 |

(i) Plot the above data on the a scatter diagram and draw a line of best fit through the points of the scatter diagram

(ii) Determine the equation of the line in (i) above in the form $r = \beta s + \alpha$, where β and α are constants. Use your result to estimate to the nearest shilling the cost of travelling a distance of 40km. **(1994 No. 14)**

**Answer: (a)** ρ = 0.8182,

***Conclusion***: high speed causes errors at 1% level of significance.




(ii) $r = 761 + 21s$

Cost = 1601/-

**Note**: *the line of best fit may vary from candidate to candidate*

**7.** (a) In a certain commercial institution a speed and error typing examination was administered to 12 randomly selected candidates A, B, C, …….L of the institution. The table below shows their speeds ($Y$) in seconds and the number of errors in their typed scripts ($X$)

| | No of errors ($X$) | Speed ($Y$) in seconds |
|---|---|---|
| A | 12 | 130 |
| B | 24 | 136 |
| C | 20 | 124 |
| D | 10 | 120 |
| E | 32 | 153 |
| F | 30 | 160 |
| G | 28 | 155 |
| H | 15 | 142 |
| I | 18 | 145 |
| J | 40 | 172 |
| K | 27 | 140 |
| L | 35 | 157 |

i) Plot the data on a scatter diagram

ii) Draw the line of best fit on your diagram and comment on the association between speed and the errors made.

iii) Determine the equation of the line in the form $y = kx + b$, where $k$ and $b$ are constants.

iv) By giving rank 1 to the fastest student and the student with the fewest errors, rank the above data, and use them to calculate the rank correlation coefficient. Comment on your result. **(1995 No. 13)**

***Answers***: iii) $y = 1.5x + 107.5$

iv) $\rho = 0.84$

High positive correlation between typing speed and errors made.

**8.** The following table gives the marks obtained in Calculus, Physics and Statistics by seven students

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Calculus | 72 | 50 | 60 | 55 | 35 | 48 | 82 |
| Physics | 61 | 55 | 70 | 50 | 30 | 50 | 73 |
| Statistics | 50 | 40 | 62 | 70 | 40 | 40 | 60 |

Draw scatter diagrams and determine the rank correlation coefficients between the performances of the students in

i) Calculus and physics

ii) Calculus and Statistics

Give interpretations to your results. **(1996 No. 16)**

**Answers:** i) $\rho = 0.9$ ii) $\rho = 0.64$

**9.** Given the variables $x$ and $y$ below

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $x$ | 80 | 75 | 86 | 60 | 75 | 92 | 86 | 50 | 64 | 75 |
| $y$ | 62 | 58 | 60 | 45 | 68 | 68 | 81 | 48 | 50 | 70 |

Obtain a rank correlation coefficient between the variables $x$ and $y$. comment on your result.

**(1999 No. 8)**

**Answers:** $\rho = 0.715$

**10.** The table below shows the percentage of sand $y$, in the soil at different depths $x$, (in cm)

| Soil depth ($x$), (cm) | Percentage of sand, ($y$) |
|---|---|
| 35 | 86 |
| 65 | 70 |
| 55 | 84 |
| 25 | 92 |
| 45 | 79 |
| 75 | 68 |
| 20 | 96 |
| 90 | 58 |
| 51 | 86 |
| 60 | 77 |

(a) i) Plot a scatter diagram for the data. Comment on the relationship between the depth of the soil and the percentage of sand in the soil.

ii) Draw a line of best fit through the points of the scatter diagram. Use your result to estimate the;

- Percentage of sand in the soil at depth of 31cm.
- Depth of soil with 54% sand.

(b) Calculate the rank correlation coefficient between the percentage of sand in soil and the depth of the soil

**(2003 No. 15)**

***Answers*** (a) ii) 92%, 96%

(b) $\rho$ = - 0.95

**11.** Eight applicants for a certain job obtained the following marks in aptitude and written tests

| Applicant | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Aptitude test | 33 | 45 | 15 | 42 | 45 | 35 | 40 | 48 |
| Written test | 57 | 60 | 40 | 75 | 58 | 48 | 54 | 68 |

Calculate the rank correlation coefficient of the applicants' performance in the two tests. Comment on your result.

**(2004 No. 7)**

**Answer:** $\rho = 0.78$

**12.** Below are marks scored by 8 students A,B,C,D,E,F,G and H in Mathematics, Economics and Geography in the end of term examinations

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Math | 52 | 75 | 41 | 60 | 81 | 31 | 65 | 52 |
| Econ | 50 | 60 | 35 | 65 | 66 | 45 | 69 | 48 |
| Geog | 35 | 40 | 60 | 54 | 63 | 40 | 55 | 72 |

Calculate the rank correlation coefficients between the performances of the students in

**i)** Mathematics and Economics

**ii)** Geography and Mathematics

Comment on the significance of Mathematics in the performance of Economics and Geography. [Spearman, $\rho$ = 0.86, based on 8 observations at 1% level of significance.]

**(2007 No. 12)**

***Answers*** **(i)** $\rho$ = 0.85 (ii) $\rho$ = 0.19

Not significant at 1% level

**13.** The heights and masses of ten students are given in the table below.

| Height (cm) | Mass (kg) |
|---|---|
| 156 | 62 |
| 151 | 58 |
| 152 | 63 |
| 146 | 58 |
| 160 | 70 |
| 157 | 6O |
| 149 | 55 |
| 142 | 57 |
| 158 | 68 |
| 141 | 56 |

(a) (i) Plot the data on a scatter diagram.

(ii) Draw the line of best fit. Hence estimate the mass corresponding to a height of 155cm.

(b) (i)Calculate the rank correlation coefficient for the data.

(ii) Comment on the significance of the heights on masses of the students. [Spearman's $\ell$ = 0.79 and Kendall's $\tau$ = 0.64 at 1% level of significance based on 10 observations.]

**(2011 No. 12)**

**Answers:** (b)(i)0.87 (ii) there is significant relationship at 1% level

**14.** A teacher gave two tests in chemistry. Five students were graded as follows:

| | GRADE | | | | |
|---|---|---|---|---|---|
| Test I | A | B | C | D | E |
| Test II | B | A | C | D | E |

Determine the rank correlation coefficient between the two tests. Comment on your result.**(2012 No 5)**

**Answer:** $\rho = 0.9$ *or* $\tau = 0.8$

There is very high positive correlation between the two tests.

**Or** not significant at 5%

**15.** The heights (cm) and ages (years) of a random sample of ten farmers are given in the table below;

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Heights (cm)** | 156 | 151 | 152 | 160 | 146 | 157 | 149 | 142 | 158 | 140 |
| **Ages (years)** | 47 | 38 | 44 | 55 | 46 | 49 | 45 | 30 | 45 | 30 |

(a) (i) Calculate the rank correlation coefficient.

(ii) Comment on your result.

(b) Plot a scatter diagram for the data. Hence draw a line of best fit.

(c) Use your diagram in (b) to find

(i) $y$ when $x = 147$

(ii) $x$ when $y = 43$

**(2013 No. 9)**

**Answers**

(c) (i) 37 (ii) 151

**16.** The table below shows scores if students in Mathematics and English tests.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Math | 72 | 65 | 82 | 54 | 32 | 74 | 40 | 53 |
| English | 58 | 50 | 86 | 35 | 76 | 43 | 40 | 60 |

Calculate the rank correlation coefficient for the students performance in the two subjects.**(2014 No.7)**

**Answers:** $\rho = \frac{1}{7}$ *or* $\tau = \frac{1}{14}$

**17.** The table below gives the points awarded to eight schools by three judges $J_1$, $J_2$ and $J_3$ during a music competition. $J_1$ was the chief judge.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $J_1$ | 72 | 50 | 50 | 55 | 35 | 38 | 82 | 72 |
| $J_2$ | 60 | 55 | 70 | 50 | 50 | 50 | 73 | 70 |
| $J_3$ | 50 | 40 | 62 | 70 | 40 | 48 | 67 | 67 |

(a) Determine the rank correlation coefficient between the judgments of

(i) $J_1$ and $J_2$.

(ii) $J_1$ and $J_3$.

(b)Who of the two other judges had a better correlation with the chief judge? Give a reason.

**(2015 No.12)**

**Answer:** (a)(i)0.7440 (ii) 0.7023

(b) $J_2$ had a better correlation with the chief Judge

**Reason**: It has slightly a higher value with chief Judge

## EXERCISE

**1.** Seven army recruits (A, B, ..., G) were given two separate aptitude tests. Their orders of merit in each test were:

| Order of merit | $1^{st}$ | $2^{nd}$ | $3^{rd}$ | $4^{th}$ | $5^{th}$ | $6^{th}$ | $7^{th}$ |
|---|---|---|---|---|---|---|---|
| $1^{st}$ test | G | F | A | D | B | C | E |
| $2^{nd}$ test | D | F | E | B | G | C | A |

Find Spearman's coefficient of rank correlation between the two orders and comment briefly on the correlation obtained.

**2.** A doctor asked ten of his patients who were smokers how many years they had smoked. In addition, for each patient, he gave a grade between 0 to 100 indicating the extent of their lung damage. The following table shows the results:

| Patient | Number of years smoking | Lung damage grade |
|---|---|---|
| A | 15 | 30 |
| B | 22 | 50 |
| C | 25 | 55 |
| D | 28 | 30 |
| E | 31 | 57 |
| F | 33 | 35 |
| g | 36 | 60 |
| H | 39 | 72 |
| I | 42 | 70 |
| J | 48 | 75 |

(a) Calculate Spearman's coefficient of rank correlation between the number of years of smoking and the extent of lung damage.
(b) Comment on the figure which you obtain

**3.** The table below shows the original marks of six candidates in two examinations.

| Candidate | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| English | 38 | 62 | 56 | 42 | 59 | 48 |
| History | 64 | 84 | 84 | 60 | 73 | 69 |

(a) Calculate a coefficient of rank correlation and comment on the value of your results.

(b) The History papers are re-marked and one of the six candidates is awarded five additional marks. Given that the other marks, and the coefficient of rank correlation, are unchanged, state, with reasons, which candidate received the extra marks.

**4.** A teacher selects one boy and one girl at random from her class, and asks them to rearrange 11 types of food in order of preference. The food types are labelled A to K and the results are given in the following table.

| Boys' order | Girls' order |
|---|---|
| E | F |
| K | K |
| F | E |
| C | C |
| B | B |
| I | I |
| D | H |
| A | D |
| G | A |
| J | J |
| H | G |

a) Calculate spearman's rank correlation coefficient for the data
b) Test at 1% level of significance whether or not there is evidence of a positive correlation
c) Interpret your conclusion to the test in part (b) above

**5.** The position in a league of 8 hockey clubs at the end of a season are shown in the table. Shown also are the average attendances (in hundreds) at home matches during that season. Calculate a coefficient of rank correlation between.

| Club | Position | Average attendance |
|---|---|---|
| A | 1 | 30 |
| B | 2 | 32 |
| C | 3 | 12 |
| D | 4 | 19 |
| E | 5 | 27 |
| F | 6 | 18 |
| G | 7 | 15 |
| H | 8 | 25 |

a) Calculate the spearman rank correlation coefficient between position in the league and average home attendances
b) Interpret your result at 5% level of significance

**6.** Ten architects each produced a new design for a new building and two judges A and B independently awards a mark $x$ and $y$ respectively to the 10 designs as given in the table below.

| Design | Judge A ($x$) | Judge B ($y$) |
|---|---|---|
| 1 | 50 | 46 |
| 2 | 35 | 26 |
| 3 | 55 | 48 |
| 4 | 60 | 44 |
| 5 | 85 | 62 |
| 6 | 25 | 28 |
| 7 | 65 | 30 |
| 8 | 90 | 60 |
| 9 | 45 | 34 |
| 10 | 40 | 42 |

a) Calculate the spearman's rank correlation coefficient for the data
b) Test at the 5% level of significance whether there is correlation between the marks awarded by the two judges

**7.** Six students were graded as follows in two subjects Mathematics and Economics by UNEB

| | | | | | | |
|---|---|---|---|---|---|---|
| Math | A | B | C | D | E | F |
| Economics | C | F | B | A | E | D |

Calculate the Kendall's rank correlation coefficient and comment on the results

**8.** Twelve chemistry students were each given a theory and practical examination. Their positions in the two examinations were as follows

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Theory | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Practical | 1 | 4 | 7 | 3 | 5 | 2 | 9 | 8 | 10 | 6 | 12 | 12 |

Calculate the Spearman's correlation coefficient for the data and test for the significance of the coefficient at 1%.

**9.** The course work grades ranging from A to G and examination marks of 8 candidates are given below.

| Coursework grade | Examination mark |
|---|---|
| A | 92 |
| C | 75 |
| D | 63 |
| B | 54 |
| F | 48 |
| C | 45 |
| G | 34 |
| E | 18 |

**Note:** Grade A is the highest grade.
(a) Calculate the spearman's rank correlation coefficient for the two sets of data
(b) Comment on the significance of the relationship between the two at 5% level. [Spearman, $\rho$ = 0.71 based on 8 observations at 5% level of significance]

# Chapter 3
# PROBABILITY THEORY

Outline
- *Definition of probability*
- *Sample space and how it is generated*
- *Probability situations*
- *Selection and transfer of items*
- *Total theorem and Baye's rule*

**Definition of probability**
The probability of an event occurring is a measure of the likelihood that it will happen. This is given a numerical scale from 0 to 1
The numbers representing probabilities can be written as decimals, fractions or percentages. For example the probability of $\frac{1}{4}$ may as well be represented as 0.25, or 25%
The probability of an event is zero if it is certain that an event is impossible to occur. For example the probability that a girl will one day become a boy, the chance that a person aged 20 years will one day become a 5 years child.

The probability of an event is 1 if it is certain that such an event will occur. For example the chance that a person will die is one (even though the time is not specified)

**Types of probability**
**a) Experimental probability**
This is the probability arising after an experiment has been performed. For example the probability of obtaining two heads when a coin is tossed three times.

**b) Equally likely probability**
This is probability of an event whose outcomes are equally likely to occur for example if a fair coin is tossed once, the probability of a head showing up is $\frac{1}{2}$

***Note***: Here the probability is thought of even before carrying out an experiment.
The probability that a pregnant mother will give birth on a Monday is $\frac{1}{7}$ because there is only one Monday out of the seven days of a week

c) **Subjective Probability**
This is probability arising due to a person's own judgment according to the experience. For example the probability that a particular car make will be stolen during a certain period is based on one's past experience. However this method is highly subject to errors and therefore cannot be tested or treated mathematically.

## Sample Space
The set of all possible outcomes of an experiment performed is called sample space, denoted by *S*. each of the individual outcome in the sample space is called a sample point. For example when a coin is tossed once, the sample space is denoted by

$S$ = ($H$, $T$); where $H$ represents a head turning uppermost and $T$ represents a tail turning uppermost

***Note***: $T$ and $H$ are sample points
When a die is thrown once, the outcomes are numbers 1 to 6; hence the sample space $S$ is

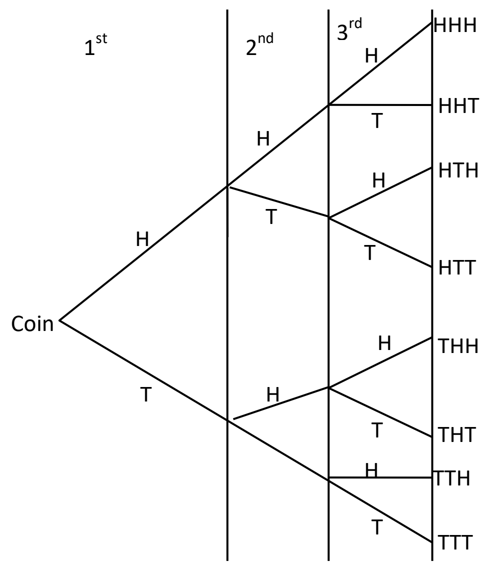
$$S = (1, 2, 3, 4, 5, 6)$$

## *Generation of Sample Space*
There are various methods used to list down the sample points. Some of which include:
- Tree diagram
- Table of outcomes
- Permutations and combinations

a) **Tree Diagram**
This is usually used to generate sample points for an experiment which results in two possible outcomes for every trial

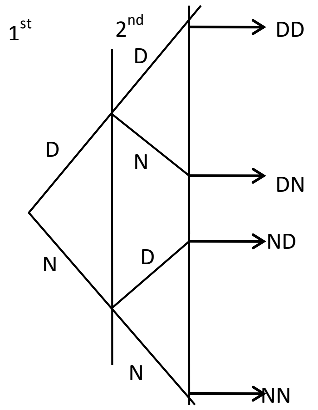
***Note***:
i) This is used when an experiment is performed at least twice
ii) The method is used regardless of whether the selection is done with or without replacement

***Examples***
Suppose three coins are tossed simultaneously once. The sample points are obtained from the tree diagram as follows:




Sample space $S$ is given by:
S = (HHH, HHT, HTH, THH, THT, TTH, TTT)

**2.** Suppose that 2 items are selected at random from a manufacturing process. If each item inspected is classified as Defective (D) on Non defective (N). to list the sample points of the sample space may be done using a tree diagram as shown below.




Sample space $S$ is given by: $S = (DD, DN, ND, NN)$

**Table of outcomes**

The table of outcomes is used to generate sample spaces where sample points appear in pairs for every trial

***Example***

Suppose the faces of each of two dice are numbered from 1 to 6 and that the dice are tossed once, we can use a table of outcomes to generate the sample outcomes as follows:

**Note**: Here we fix the outcome of one of the dice and let the outcome of the other die vary from 1 to 6

A

| dice | 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 |

### c) Permutations and Combinations

We can use the permutation theory to know the number of sample points in a sample space, however this will not enable us obtain the sample points themselves.

Permutation is an arrangement of all or part of a set of objects

***Note***:

i) The number of permutations of $n$ distinct objects is $n!$ (Read as $n$ factorial, which is defined as $n! = n(n-1)(n-2) \ldots \times 1$). For example the number of arrangements which can be made out of letters $a$, $b$, $c$ and $d$ is $4!$ $= 4 \times 3 \times 2 \times 1 = 24$

ii) The number of permutations of $n$ distinct objects taken $r$ at a time is given by ${}^{n}P_{r} = \dfrac{n!}{(n-r)!}$

**Example**

Two rotary tickets are drawn from 20 for the 1st and 2nd prices. Find the number of sample points in the sample space(s)

$$S = \frac{20}{(20-2)!} = \frac{20!}{18!}$$
$$= \frac{20 \times 19 \times 18!}{18!} = 20 \times 19$$
$$= 380$$

The number of distinct permutations of n things of which $n_1$, are of one kind, $n_2$ of second kind, and $n_k$ of the $k^{th}$ kind is given by $S = \dfrac{n!}{n_1!n_2!\ldots.n_k!}$

**For example**

In how many different ways can 3 red, 4 yellow and 2 blue bulbs be arranged in a string of $x$-mass tree lights with 9 sockets?

The total number of distinct arrangements is given by

$$\frac{9!}{3!4!2!} = \frac{9 \times 8 \times 7 \times 6 \times 5 \times 4!}{3 \times 2 \times 2 \times 4!}$$
$$= \frac{9 \times 8 \times 7 \times 6 \times 5 \times}{12} = 1260$$

## Combinations

The number of ways of selecting $r$ things from n unlike things is denoted by:

$\begin{pmatrix} n \\ r \end{pmatrix}$ or ${}^{n}C_{r}$

Or it denotes the number of combinations (selections) of $n$ unlike things taken $r$ at a time.

This is defined as: $\begin{pmatrix} n \\ r \end{pmatrix} = \dfrac{n!}{(n-r)!r!}$

***Example***:

1. Find the number of selecting 11 players from 14 members

Number of selecting all players = $\begin{pmatrix} 14 \\ 11 \end{pmatrix}$

$$= \frac{14 \times 13 \times 12 \times 11!}{11! \times 3!}$$
$$= \frac{14 \times 13 \times 12}{3 \times 2} = 14 \times 13 \times 2$$
$$= 364$$

**Note**: The number of ways of choosing 11 players from 14 members is the same as the number of ways of selecting 3 non-players from 14

Number of selecting 3 non members $= \begin{pmatrix} 14 \\ 3 \end{pmatrix}$

$$= \frac{14 \times 13 \times 12 \times 11!}{11! \times 3!} = \frac{14 \times 13 \times 12}{6} = 364$$

Hence $\begin{pmatrix} n \\ r \end{pmatrix} = \begin{pmatrix} n \\ n-r \end{pmatrix}$

2. From 4 republicans and 3 democrats, find the number of committee of 3 that can be formed with 2 republicans and 1 democrat.

***Solution***

Number of ways of selecting 2 republicans from 4 is

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{4\times3\times2\times1}{2\times1\times2\times1}$$
$$= 3\times2 = 6$$

Number of ways of selecting 1 democrat out of 3 is:

$$\binom{3}{1} = \frac{3!}{2!1!} = \frac{3\times2}{2\times1} = 3$$

Therefore the number of committees that can be formed is
6 × 3 = 18

Or: This could be explained by saying number of

selecting 3 members $= \binom{4}{2}\cdot\binom{3}{1}$

$$= \frac{4!}{2!2!}\cdot\frac{3!}{2!1!} = 6\times3 = 18$$

**Note**: *Permutations and Combinations are covered in details in Pure Mathematics*

**Application of Combinations in Probabilities**

1. A box contains 10 radio valves, all apparently, although 4 of them are actually sub-standard. Find the probability that if two of the valves are taken from the box, they are both sub-standard.

***Solution***

P(Both values are substandard)

$$= \frac{\text{number of selecting 2 from 4 sub-standard}}{\text{number of selecting 2 from the 10}}$$

$$= \frac{\binom{4}{2}}{\binom{10}{2}} = \frac{6}{15} = \frac{2}{15}$$

Note: $\binom{n}{r}$ can be read off directly from mathematical tables.

2. If two cards are picked from a well shuffled pack of 52 playing cards, what is the probability that:

   a) they are both Aces?
   b) neither of them is an Ace
   c) at least one of them is an Ace

***Solution***

(a) $P$(both are Aces) = $P$(1st is an Ace, 2nd is an Ace)
= $P$(1st is an Ace) × P(2nd is an Ace)

$$= \frac{4}{52}\times\frac{3}{51} = \frac{1}{221}$$

***Alternatively***,

$P$(both Aces)

$$= \frac{\text{no of ways of choosing 2 cards from 4 aces in pack}}{\text{no of ways of choosing 2 out of 52 in pack}}$$

$$= \frac{\binom{4}{2}}{\binom{52}{2}} = \frac{6}{1326} = \frac{1}{221}$$

ii) $P$(neither is an ace)
= $P$(1st is not an Ace) × P(2nd is not an Ace)

$$= \frac{48}{52}\times\frac{47}{51} = \frac{188}{221}$$

***Alternatively***;

$P$(Neither is an Ace)

$$= \frac{\text{no of ways of selecting 2 cards from 48 which are not Aces}}{\text{no of ways of choosing 2 out of 52 in cards}}$$

$$= \frac{\binom{48}{2}}{\binom{52}{2}} = \frac{1128}{1326} = \frac{188}{221}$$

iii) $P$(At least one is an Ace)
= 1 − $P$(neither is an Ace)

$$= 1 - \frac{188}{221} = \frac{33}{221}$$

3. Seven cards labelled A, B, C, D, E, F, G are thoroughly shuffled and dealt out face upwards on a table. Find the probability that:
   a) the first three cards to appear are labelled A, B, C in that order
   b) the first three cards to appear are labelled A, B, C but in any order
   c) the seven cards appear in their original order: A, B, C, D, E, F, G

***Solution***

(a)With specific order, we use permutations

Number of ways of arranging 3 letters from seven

$$= {}^7P_3 = \frac{7!}{(7-3)!} = \frac{7!}{4!}$$

$$= \frac{7!}{4!} = 7\times6\times5 = 210$$

$\therefore P(1^{st}$ three letters are A, B, C in that order$) = \frac{1}{210}$

b) With the order not specific, we use combinations

number of ways of arranging three letters from seven

$$= {}^7C_3$$

$$= \frac{7!}{4!3!} = \frac{7\times6\times5}{3\times2} = 35$$

$\therefore P(1^{st}$ three letters are A, B, C) in any order)

$$= \frac{1}{35}$$

c) $P$(Number of ways of arranging 7 letters) = 7!

= 5040

$\therefore P$(7 cards appear in their order) = $\frac{1}{5040}$

4. Three letters are selected at random from the word BIOLOGY. Find the probability that the selection
   a) doesn't contain the letter O
   b) contains both of the letters O

***Solution***

With this question, we need to observe the following:

- Find *the number of selections without the letter O*
- *Find the number of selections with one letter O*

- *Find the number of selections with two letters O*
- *Find the total number of selections by adding the above three*

Now the number of selections without the letter O
= number of ways of choosing three letters from B, I, L, G, Y
$= {}^5C_3 = 10$

Number of selections with one letter O
= number of ways of choosing two letters from B, I, L, G, Y
$= {}^5C_2 = 10$

Number of selections with two letters O
= number of ways to choosing one letter from B, I, L, G, Y
$= {}^5C_1 = 5$

Total number of selections of choosing three letters from BIOLOGY = 10 + 10 + 5 = 25

a) P(selection doesn't contain letter O) $= \frac{10}{25} = \frac{2}{5}$

b) P(Selection contains two letters O) $= \frac{5}{25} = \frac{1}{5}$

## Events

An event is a sub-set of a sample space. For example when a affair coin is tossed three times, the sample space is
***S*** = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
If we are interested in picking at least two heads out of the three tosses, this constitutes an event E, where
E = {HHH, HHT, HTH, THH}
**Note**: Probability of an event is defined as

$$P(E) = \frac{n(E)}{n(S)}$$

In the example above, $P(E) = \frac{4}{8} = \frac{1}{2}$

### *Intersection of events:*

For two events A and B, the probability that A and B occur is denoted by $P(A \cap B)$

*See more details about the "AND" situation under probability situations a head.*

### *Union of Events*

For two events A and B, the probability that A or B or both occur is denoted by $P(A \cup B)$ which is defined as
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

*See more details about the "OR" situation under probability situations a head.*

***Types of Events***

There are several types of events, however in this book we shall strictly look at the following:

- Independent events
- Mutually exclusive events
- Exhaustive events
- Undefined events

***Independent events***

Two events are said to be independent if the occurrence of one event does not affect the other.
If events A and B are independent, then

i) $P(A \cap B) = P(A) \times P(B)$

ii) $P(A \cup B) = P(A) + P(B) - P(A) \times P(B)$

Also the compliments of A and B denoted by A' and B' are also independent.
$\Rightarrow P(A' \cap B') = P(A') + P(B')$
Where $P(A') = 1 - P(A)$ and $P(B') = 1 - P(B)$

***Examples***

**1.** Events A and B are independent such that $P(A \cup B)$ = 0.8, P(A) = 0.5. Find:

i) P(B)
ii) $P(A' \cap B)$
iii) $P(A' \cup B)$

*Solution*
*By definition:*

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

0.8 = 0.5 + $P$(B) − 0.5 × $P$(B)
0.3 = $P$(B) − 0.5P(B)
0.3 = 0.5$P$(B)

$P(B) = \frac{0.3}{0.5}$ = 0.6

ii) $P(A' \cap B) = P(A') \times P(B)$

$= [1 - P(A)] \times P(B) = 0.5 \times 0.6 = 0.3$

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

= 0.5 + 0.6 − 0.3 = 0.8

**2.** A and B are said to be independent events such that P(A) = 0.4, $P(A \cup B')$ = 0.9. Find:

i) P(B)
ii) P(A$\cap$B)
iii) $P(A' \cap B')$
iv) $P(A' \cup B')$

***Solution***

$$P(A \cup B') = P(A) + P(B') - P(A) \cdot P(B')$$
$$0.9 = 0.4 + P(B') - 0.4P(B')$$
$$0.5 = 0.6P(B')$$
$$P(B') = \frac{0.5}{0.6} = \frac{5}{6}$$
$$P(B) = 1 - P(B')$$
$$= 1 - \frac{5}{6} = \frac{1}{6}$$

(ii) $P(A \cap B) = P(A) \cdot P(B)$

$$= 0.4 \times \frac{1}{6}$$
$$= \frac{2}{5} \times \frac{1}{6} = \frac{2}{30} = \frac{1}{15}$$

(iii) $P(A' \cap B') = P(A') \cdot P(B')$

$$= 0.6 \times \frac{5}{6} = \frac{6}{10} \times \frac{5}{6} = \frac{1}{2}$$

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

$= 0.6 + \frac{5}{6} - \frac{1}{2}$

$= \frac{3}{5} + \frac{5}{6} - \frac{1}{2} = \frac{18+25-15}{30}$

$= \frac{28}{30} = \frac{14}{15}$

**3. A** and B are independent events such that P(A $\cup$B) = 0.8, P(A $\cap$B) = 0.1. Find the possible values of P(A) and P(B)

***Solution***

By definition,

P(A $\cap$B) = P(A) × P(B)

0.1 = P(A) × P(B)

$P(A) = \frac{0.1}{P(B)}$

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

$0.8 = P(A) + P(B) - 0.1$

$0.9 = P(A) + P(B)$

Substituting for P(A),

$0.9 = \frac{0.1}{P(B)} + P(B)$

Let P(B) = $x$

$\Rightarrow 0.9x = 0.1 + x^2$

$x^2 - 0.9x + 0.1 = 0$

$x = \frac{0.9 \pm \sqrt{(0.9)^2 - 4 \times 0.1}}{2}$

$x = \frac{0.9 \pm 0.64}{2}$

Either $x = \frac{0.9+0.64}{2} = 0.77$ hence $P(B)$ = 0.77

Or $x = \frac{0.9-0.64}{2} = 0.13$ Hence $P(B)$ = 0.13

If P(B) = 0.77, $P(A) = \frac{0.1}{0.77} = 0.13$

If P(B) = 0.13, $P(A) = \frac{0.1}{0.13} = 0.77$

## *Mutually Exclusive Events*

Two events are said to be mutually exclusive if they do not have any element in common. For example even numbers and odd numbers are mutually exclusive.

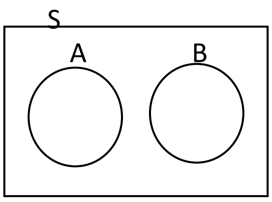
***Note***: mutually exclusive events are represented by disjoint sets.

If A and B are mutually exclusive events, then

$A \cap B = \{\ \}$

$n(A \cap B) = 0$

Hence $P(A \cap B) = 0$

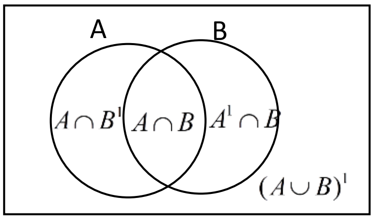



From $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,

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

**Important properties**:

Consider the Venn diagram shown below




From the Venn diagram above,

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

But $P(A \cap B) = 0$

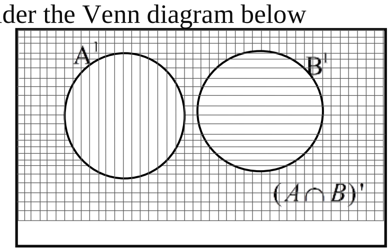
Hence $P(A) = P(A \cap B')$

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

But $P(A \cap B) = 0$

Hence $P(B) = P(A' \cap B)$

Also consider the Venn diagram below




By shading the regions for A' and B',

$A' \cap B'$ = region shaded twice

= $(A \cup B)'$

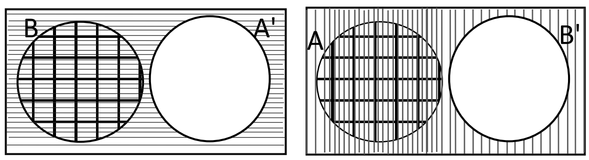
Hence $P(A' \cap B') = P(A \cup B)'$

$A' \cup B'$ = any shaded region i.e. whether shaded once or twice

Hence P(A' $\cup$B') = 1

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

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




Hence in summery,

P(A $\cap B^1$) = P(A)

P($A^1 \cap$B) = P(B)

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

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

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

P($A^1 \cap B^1$) = P(A $\cup$B)$^1$

***Note***: Total area under the Venn diagram = 1

### *Examples*

1. A and B are mutually exclusive events such that $P(A \cup B) = 0.8$, $P(A) = 0.4$,

Find: i) $P(B)$, ii) $P(A \cap B')$

iii) $P(A' \cap B)$ iv) $P(A' \cap B')$

v) $P(A' \cup B')$

***Solution***

i) *For mutually exclusive events,*

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

$0.8 = 0.4 + P(B)$

$P(B) = 0.4$

ii) $P(A\cap B') = P(A) = 0.4$

iii) $P(A'\cup B) = P(A')$

$= 1 - P(A) = 1 - 0.4$

$= 0.6$

iv) $P(A'\cap B') = P(A\cup B)'$

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

$= 1 - 0.8 = 0.2$

v) $P(A'\cup B') = P(A\cap B)'$

$= 1 - P(A\cap B)$

$= 1 - 0 = 1$

## Some important laws/rules

a) Laws of induction

For events A and B,

i) $P(A)+P(A') = 1$

ii) $P(B)+P(B') = 1$

iii) $P(A\cap B)+P(A\cap B)' = 1$

iv) $P(A\cup B)+P(A\cup B)' = 1$

Hence $P(X)+P(X') = 1$

b) Demorgan's rule

For any events A and B,

i) $P(A'\cap B') = P(A\cup B)'$

ii) $P(A'\cup B') = P(A\cap B)'$

**Note:** These laws/rules are applicable to all types of events

2. Events A and B are mutually exclusive such that $P(A'\cap B) = 0.3$, $P(A'\cup B) = 0.45$

Find: i) P(B), ii) P(A) iii) $P(A\cap B)$

iv) $P(A\cap B')$ v) $P(A'\cap B')$

vi) $P(A'\cup B')$

***Solution***

i) $P(A'\cap B) = P(B)$

Hence $P(B) = 0.3$

ii) $P(A'\cup B) = P(A')$

$P(A') = 0.45$

$P(A) = 1 - 0.45 = 0.55$

iii) $P(A\cap B) = 0$

iv $P(A\cap B') = P(A) = 0.55$

v) $P(A'\cap B') = P(A\cup B)'$ *(Demorgan's rule)*

$= 1 - P(A\cup B) = 1 - [0.3 + 0.55]$

$= 1 - 0.85 = 0.15$

vi) $P(A'\cup B') = P(A\cap B)'$ (*Demorgan's rule*)

$= 1 - P(A\cap B)$

$= 1 - 0 = 1$

**3**. A and B are mutually exclusive events such that $P(A\cup B) = 0.9$, $P(A\cup B') = 0.6$

Find: i) $P(B)$ ii) $P(A)$

iii) $P(A'\cup B)$ iv) $P(A'\cap B')$

v) $P(A'\cup B')$

***Solution***

i) $P(A\cup B') = P(B')$

$P(B') = 0.6$

$P(B) = 1 - P(B')$

$= 1 - 0.6 = 0.4$

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

$0.9 = P(A) + 0.4$

$P(A) = 0.5$

iii) $P(A'\cup B) = P(A')$

$= 1 - P(A)$

$= 1 - 0.5 = 0.5$

iv) $P(A'\cap B') = P(A\cup B)'$ *Demorgan's rule*

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

$= 1 - 0.9 = 0.1$

v) $P(A'\cup B') = P(A\cap B)'$ *Demorgan's rule*

$= 1 - P(A\cap B)$

$= 1 - 0 = 1$

***Exhaustive events***

Two events are said to be exhaustive if: -

i) they do not have any element(s) in common
ii) the sum of their probabilities is one

***Note***: A and A' are exhaustive events because

$P(A) + P(A') = 1$

$A\cap B$ and $(A\cap B)'$ are also exhaustive events because

$P(A\cap B)+P(A\cap B)' = 1$

***Examples***

**1.** Given that A and B are exhaustive events such that P(A) = 0.6. Find

i) $P$(B) ii) $P(A\cup B)$

iii) $P(A'\cup B')$

***Solution***

$P(A) + P(B) = 1$

$P(B) = 1 - 0.6$

$= 0.4$

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

$= 0.6 + 0.4 = 1$

iii) $P(A'\cup B') = P(A\cap B)'$

$= 1 - P(A\cap B)$

$= 1 - 0 = 1$

**2**. A, B and C are exhaustive events such that $P(A) = 0.1$, $2P(B) = P(C)$

Find (i) $P$(B) (ii) $P$(C) iii) $P(A\cup B)$

***Solution***

i) $P(A\cup B\cup C) = 1$

$P(A) + P(B) + P(C) = 1$

$0.1 + P(B) + 2P(B) = 1$

$3P(B) = 0.9$

$P(B) = \frac{0.9}{3} = 0.3$

ii) $P(C) = 2 \times 0.3 = 0.6$

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

$= 0.1 + 0.6 = 0.7$

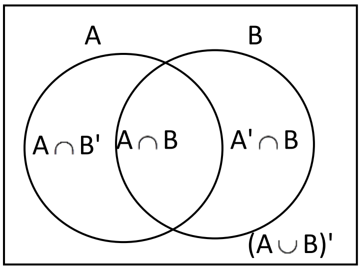
## *Undefined Events*

For undefined events *A* and *B*,

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

Where $P(A\cap B)$ has no restrictions unlike for mutually exclusive and independent events.

By considering the Venn diagram below,




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

The above results are summarized in a square table known as a contingency table.
For events A and B, we have:

***Contingency Table***

| 1 | B | B' | 1 |
|---|---|---|---|
| A | $A\cap B$ | $A\cap B'$ | A |
| A' | $A'\cap B$ | $A'\cap B'$ | A' |
| 1 | B | B' | |

From the table above,

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

***Examples***

1. Events A and B are such that $P(A) = 0.4$, $P(A\cup B) = 0.9$, $P(A\cap B) = 0.1$

Find: i) $P(B)$ ii) $P(A\cap B')$
iii) $P(A\cup B')$ iv) $P(A'\cap B')$
v) $P(A'\cup B')$

***Solution***

i) $P(A\cup B) = P(A) + P(B) - P(A\cap B)$
$0.9 = 0.4 + P(B) - 0.1$
$0.9 - 0.3 = P(B)$
$P(B) = 0.6$

ii) By using the contingency table;

| 1 | B | B' | 1 |
|---|---|---|---|
| A | $A\cap B$ | $A\cap B'$ | A |
| A' | $A'\cap B$ | $A'\cap B'$ | A' |
| 1 | B | B' | 1 |

From the table above;

$P(A) = P(A\cap B) + P(A\cap B')$
$0.4 = 0.1 + P(A\cap B')$
$P(A\cap B') = 0.3$

iii) $P(A\cup B') = P(A) + P(B') - P(A\cap B')$
$= 0.4 + 0.4 - 0.3 = 0.5$

iv) $P(A'\cap B') = P(A\cup B)'$ *By Demorgan's rule*
$= 1 - P(A\cup B)$ (*By induction*)
$= 1 - 0.9 = 0.1$

**OR** from the table above,

$P(B') = P(A\cap B') + P(A'\cap B')$
$0.4 = 0.3 + P(A'\cap B')$
$P(A'\cap B') = 0.1$

v) $P(A'\cup B') = P(A\cap B)'$ *By Demorgan's rule*
$= 1 - P(A\cup B)$ (*By induction*)
$= 1 - 0.1$
$= 0.9$

OR: $P(A'\cup B') = P(A') + P(B') - P(A'\cap B')$
$= 0.6 + 0.4 - 0.1 = .9$

2. Events A and B are such that $P(A'\cap B) = 0.3$, $P(A\cup B) = 0.8$, $P(A\cap B) = 0.2$

a) Find: i) $P(B)$ ii) $P(A)$
` iii) $P(A'\cup B)$ iv) $P(A'\cap B')$
v) $P(A'\cup B')$

b) Show whether A and B are mutually exclusive or independent events

***Solution***

i) From the contingency table, (It is advisable to draw it whenever referring to it)
$P(B) = P(A\cap B) + P(A^1\cap B)$
$= 0.2 + 0.3 = 0.5$

ii) $P(A\cup B) = P(A) + P(B) - P(A\cap B$
$0.8 = P(A) + 0.5 - 0.2$
$P(A) = 0.8 - 0.3 = 0.5$

iii) $P(A'\cup B) = P(A') + P(B) - P(A'\cap B)$
$= 0.5 + 0.5 - 0.3 = 0.7$

iv) $P(A'\cap B') = P(A\cup B)'$ *By Demorgan's rule*
$= 1 - P(A\cup B)$
$= 1 - 0.8 = 0.2$

v) $P(A'\cup B') = P(A\cap B)'$ *By Demorgan's rule*
$= 1 - P(A\cap B)$
$= 1 - 0.2 = 0.8$

b) For mutually exclusive events, $P(A\cap B) = 0$
But $P(A\cap B) = 0.2$. Hence the events are not mutually exclusive

Or; For mutually exclusive events,
$P(A\cup B) = P(A) + P(B)$
$0.8 \neq (0.5 + 0.5 = 1)$

For independent events, $P(A\cap B) = P(A) \times P(B)$
Now $P(A\cap B) = 0.2$
$P(A) \times P(B) = 0.5 \times 0.5 = 0.25$
$P(A\cap B) \neq P(A) \times P(B)$
$\therefore$ The events are not independent.
Hence the events are neither mutually exclusive nor independent

3. For three events A, B and C, prove that
$P(A\cup B\cup C) = P(A) + P(B) + P(C) - P(A\cap B) - P(A\cap C) - P(B\cap C) + P(A\cap B\cap C)$

***Solution***

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

Substituting for $X$;
$P(A\cup B\cup C) = P(A) + P(B\cup C) - P(A\cap B\cup C)$
$= P(A) + P(B) + P(C) - P(B\cap C) - P(A\cap B\cup A\cap C)$
$= P(A) + P(B) + P(C) - P(B\cap C) -$
$$\left[P(A\cap B)+P(A\cap C)-P(A\cap B\cap C)\right]$$
$= P(A) + P(B) + P(C) - P(B\cap C) - P(A\cap B) - P(A\cap C) + P(A\cap B\cap C)$.
$= P(A) + P(B) + P(C) - P(A\cap B) - P(A\cap C) - P(B\cap C) + P(A\cap B\cap C)$.

4. Events A, B and C are such that P(A) = 0.4, P(B) = 0.3, P(C) = 0.5, $P(A\cap C) = 0.1$ and $P(B\cap C) = 0.2$.

If A and B are mutually exclusive, find:
i) $P(A\cap B\cap C)$ ii) $P(A\cup B\cup C)$

***Solution***

i) Since A and B are mutually exclusive, $P(A\cap B) = 0$, hence $P(A\cap B\cap C) = 0$

ii) $P(A\cup B\cup C) = P(A) + P(B) + P(C) - P(A\cap C) - P(B\cap C)$
$= 0.4 + 0.3 + 0.5 - 0.1 - 0.2 = 0.9$

## *Conditional Probability*

This is a case where one event occurs given that the other has already occurred. The probability that event A occurs given that B has already occurred denoted by $P(A/B)$ is defined as $P(A/B) = \dfrac{P(A\cap B)}{P(B)}$

***Note***: i) If the events are independent, then

$$P(A/B) = \frac{P(A)\cdot P(B)}{P(B)} = P(A)$$

ii) If events are mutually exclusive, then

$$P(A/B) = \frac{P(A\cap B)}{P(B)} = \frac{0}{P(B)} = 0$$

iii) $P(A/B) + P(A'/B) = 1$

**Proof**:

$$P(A/B) + P(A'/B) = \frac{P(A\cap B)}{P(B)} + \frac{P(A'\cap B)}{P(B)}$$
$$= \frac{P(A\cap B) + P(A'\cap B)}{P(B)}$$
$$= \frac{P(B)}{P(B)} = 1$$

***Examples***

**1.** A and B are events such that $P(A/B) = 0.4$, P(B) = 0.25 and P(A) = 0.2.

Find: (i) $P(A\cap B)$ (ii) $P(B/A)$
(iii) $P(A\cup B)$

***Solution***

(i) $P(A/B) = \dfrac{P(A\cap B)}{P(B)}$

$$P(A/B) = \frac{P(A\cap B)}{P(B)}$$
$$0.4 = \frac{P(A\cap B)}{0.25}$$
$P(A\cap B) = 0.4 \times 0.25 = 0.1$

(ii) $P(B/A) = \dfrac{P(B\cap A)}{P(A)}$

$$= \frac{P(A\cap B)}{P(A)} = \frac{0.1}{0.2} = 0.5$$

(iii) $P(A\cup B) = P(A) + P(B) - P(A\cap B)$
$= 0.2 + 0.25 - 0.1$
$= 0.35$

**2**. Events A and B are such that $P(A\cup B) = 0.8$, $P(A/B) = 0.2$ and $P(A'\cap B) = 0.4$

Find i) P (B) (ii) $P(A\cap B)$ (iii) P (A)
(iv) $P(A/B')$ (v) $P(A'/B')$

***Solution***

(i) $P(A/B) = \dfrac{P(A\cap B)}{P(B)}$

$$0.2 = \frac{P(A\cap B)}{P(B)}$$
$$0.2P(B) = P(A\cap B)$$

But $P(B) = P(A\cap B) + P(A'\cap B)$
$P(B) = 0.2P(B) + 0.4$
$0.8P(B) = 0.4$

$$P(B) = \frac{0.4}{0.8} = 0.5$$

ii) $P(A\cap B) = 0.2P(B)$
$= 0.2 \times 0.5 = 0.1$

iii) $P(A\cup B) = P(A) + P(B) - P(A\cap B)$
$0.8 = P(A) + 0.5 - 0.1$
$0.8 = P(A) + 0.4$
$P(A) = 0.4$

iv) $P(A/B') = \dfrac{P(A\cap B')}{P(B')}$

But $P(A) = P(A\cap B) + P(A\cap B')$
$0.4 = 0.1 + P(A\cap B')$
$P(A\cap B') = 0.3$

Hence $P(A/B') = \dfrac{0.3}{1-0.5}$

$$= \frac{0.3}{0.5} = 0.6$$

v) $P(A'/B') = \dfrac{P(A'\cap B')}{P(B')}$

$$= \frac{P(A\cup B)'}{P(B')} = \frac{1-P(A\cup B)}{P(B')}$$
$$= \frac{1-0.8}{0.5} = \frac{0.2}{0.5} = 0.4$$

**3.** The probability that a regular scheduled flight departs on time is 0.83 and the probability that it arrives on time is 0.92. The probability that it departs on time and arrives on time is 0.78. Find the probability that the plane:

a) arrives on time given that it departs on time
b) departs on time given that it arrives on time.

***Solution***

Let $T_A$ = event that the flight arrives on time
$T_D$ = Event that the flight departs on time

Now $P(T_A) = 0.83$
$P(T_D) = 0.92$
$P(T_D\cap T_A) = 0.78$

a) $P(T_A/T_D) = \dfrac{P(T_D\cap T_A)}{P(T_D)} = \dfrac{0.78}{0.83} = 0.94$

b) $P(T_D/T_A) = \dfrac{P(T_D\cap T_A)}{P(T_A)} = \dfrac{0.78}{0.92} = 0.85$

**4**. In a large group of people it is known that 10% have a hot breakfast, 20% have a hot lunch and 25% have a hot breakfast or a hot lunch. Find the probability that a person chosen at random from this group:

a) has a hot breakfast and a hot lunch
b) has a hot lunch, given that he had a hot breakfast.

***Solution***

Let B = Event that a person has a hot breakfast
L = Event that a person has a hot lunch

Now $P(B) = 10\% = 0.1$
$P(L) = 20\% = 0.2$
$P(B \cup L) = 25\% = 0.25$

a) $P(B \cup L) = P(B) + P(L) - P(B \cap L)$
$0.25 = 0.1 + 0.2 - P(B \cap L)$
$P(B \cap L) = 0.3 - 0.25 = 0.05$

b) $P(L/B) = \dfrac{P(L \cap B)}{P(B)} = \dfrac{0.05}{0.1} = 0.5$

**5**. In a group of 100 people, 40 own a cat, 25 own a dog and 15 own a cat and a dog. Find the probability that a person chosen at random:

i) owns a dog or a cat
ii) owns a dog or a cat but not both
iii) owns a dog given that he owns a cat
iv) doesn't own a cat, given that he owns a dog

***Solution***

Let $C$ = event that a person owns a cat
$D$ = event that a person owns a dog

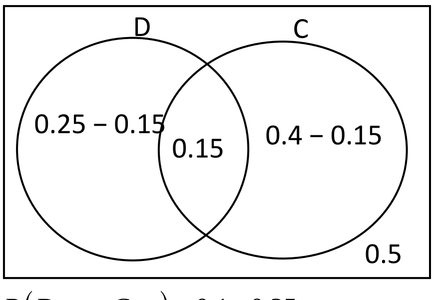
$$P(C) = \frac{40}{100} = 0.4$$

$$P(D) = \frac{25}{100} = 0.25$$

$$P(C \cap D) = \frac{15}{100} = 0.15$$

(i) $P(D \text{ or } C) = P(D \cup C)$
$P(D \cup C) = P(D) + P(C) - P(D \cap C)$
$= 0.4 + 0.25 - 0.15$
$= 0.5$

(ii) $P(D_{only} \cup C_{only}) = P(D_{only}) + P(C_{only})$




$$P(D_{only} \cup C_{only}) = 0.1 + 0.25$$
$$= 0.35$$

(iii) $P(D/C) = \dfrac{P(D \cap C)}{P(C)} = \dfrac{0.15}{0.40} = 0.375$

(iv) $P(C'/D) = 1 - P(C/D)$

$$= 1 - \frac{P(C \cap D)}{P(D)} = 1 - \frac{0.15}{0.25}$$
$$= 1 - 0.6 = 0.4$$

## *Probability Situations*

The possible situations that will be considered are:

- The 'AND' situation
- The 'OR' situation
- The 'AND and ONLY' situation

### *The "AND" Situation*

This arises under two circumstances

i) when we are interested in the joint occurrence of events
ii) when we are dealing with a sequence of events

**(i) Consider a case of joint occurrence of events:**

Here we could use Venn diagrams for illustrations

***Example***

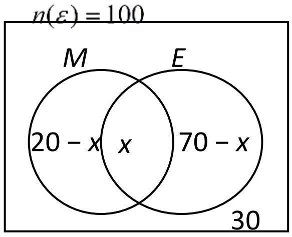
**1**. At a certain school, students take either Maths or Economics or none of these two. There are 100 students of whom 20 take Maths, 70 take Economics and 30 take neither subject. What is the probability that a student chosen at random takes both Economics and Maths?

***Solution***

Let $M$ = Event that a student takes Maths,
$E$ = Event that a student takes Economics

Let $n(M \cap E) = x$

$n(\varepsilon) = 100$



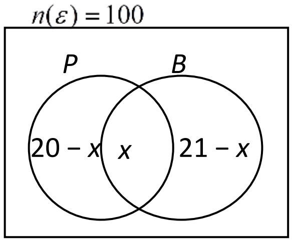

From the Venn diagram above,
$20 - x + 70 - x + x + 30 = 100$
$120 - x = 100$
$x = 20$

$$P(M \cap E) = \frac{n(M \cap E)}{n(\varepsilon)} = \frac{20}{100} = 0.2$$

**2.** In a group of 30 students, all study at least one of the subjects Physics and biology, 20 attend the Physics class and 21 attend the Biology class. Find the probability that a student chosen at random studies both Physics and Biology

***Solution***

$n(\varepsilon) = 100$




From the Venn diagram above,
$20 - x + 21 - x + x = 30$
$41 - x = 30$
$x = 11$

$$P(P \cap B) = \frac{n(P \cap B)}{n(\varepsilon)} = \frac{11}{40} = 0.275$$

**(ii) Considering sequence of events**

Suppose we have events A, B and C, the probability that they occur in the order A, B, C is given by $P(A \cap B \cap C)$.
If the events are independent, then
$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

***Examples***

**1**. A bag contains 8 red, 2 white and 6 blue balls. Three balls are drawn at random from the bag. Find the probability that the balls are drawn in this order Red, White and Blue if the:

i) balls are drawn with replacement
ii) balls are drawn without replacement

***Solution***

Let $R_1$ = a red ball is drawn on first draw

$W_2$ = a white ball is drawn on second draw

$B_3$ = a blue ball is drawn on third draw

| Red | White | Blue | Total |
|---|---|---|---|
| 8 | 2 | 6 | 16 |

i) Drawing with replacement

$P$(Red, White, Blue) = P( $R_1$ and $W_2$ and $B_3$ )

$$= P(R_1 \cap W_2 \cap B_3)$$
$$= P(R_1) \cdot P(W_2) \cdot P(B_3)$$
$$= \frac{8}{16} \times \frac{2}{16} \times \frac{6}{16} = \frac{96}{4096}$$
$$= 0.0234$$

ii) Drawing without replacement

P(Red, White, Blue) = P( $R_1$ and $W_2$ and $B_3$ )

$$= P(R_1 \cap W_2 \cap B_3)$$
$$= P(R_1) \cdot P(W_2/R_1) \times P(B_3/R_1 \cap W_2)$$
$$= \frac{8}{16} \times \frac{2}{15} \times \frac{6}{14}$$
$$= \frac{96}{3360} = 0.02857$$

**2**. Three boys A, B and C take part in a swimming competition. The respective probabilities of hitting the target are $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{4}$. If they all throw at the target at once, find the probability that the target will be hit by all of them.

***Solution***

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{3}, \quad P(C) = \frac{1}{4}$$

$P$(target will be hit by all)

$$= \text{P}(A_{hits}) \text{ and } \text{P}(B_{hits}) \text{ and } \text{P}(C_{hits})$$
$$= P(A \cap B \cap C)$$
$$= P(A) \times P(B) \times P(C)$$
$$= \frac{1}{2} \times \frac{1}{3} \times \frac{1}{4} = \frac{1}{24}$$

## *The OR situation*

This is a situation in which we are interested in the occurrence of either one or two or three or ….. or all of the events in question.

If A and B are two events, the probability that either A or B or even both occur is denoted by $P(A \cup B)$ which is defined as

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

***Example***

1. In an experiment, two fair dice each numbered from 1 to 6 are simultaneously tossed, determine the probability that:

i) a sum of either 7 or a sum which is a prime number is obtained
ii) a sum of either a composite number or even number is obtained

***Solution***

Table of sums

| | | First die | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| Second die | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

Let A = sum of prime numbers
B = sum of 7 numbers
E = a sum of even numbers
C = Composite numbers

(i) n(A) = $P(A) = \frac{\text{n(A)}}{\text{n(S)}} = \frac{6}{36} = \frac{1}{6}$

B = {2, 3, 5, 7, 11}

When you count them in the table, they are 15.
So n(B) = 15

$$P(B) = \frac{n(B)}{n(S)} = \frac{15}{36}$$

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

P(A or B) = $\frac{6}{36} + \frac{15}{36} - \frac{6}{36}$

$$= \frac{15}{36} = \frac{5}{12}$$

(ii)**Note**: A composite number is a number that has got at least two distinct factors apart from 1 and itself.

Now C = {6, 8, 10, 12}
n(C) = 14

$$P(C) = \frac{n(C)}{n(S)} = \frac{14}{15}$$

E = {2, 4, 6, 8, 10, 12}
n(E) = 18

P(E) = $\frac{n(E)}{n(S)} = \frac{18}{36}$

$\text{C} \cap \text{E}$ = {6, 8, 10, 12}

$\text{n}(\text{C} \cap \text{E})$ = 14

$$\text{P}(\text{C or E}) = P(C \cup E)$$
$$= P(C) + P(E) + P(C \cap E)$$
$$= \frac{14}{36} + \frac{18}{36} - \frac{14}{36} = \frac{18}{36} = \frac{1}{2}$$

**2**. An experiment involves randomly drawing two balls without replacement from a bag consisting of 10 white and 6 red balls. Determine the probability that the second ball is white

***Solution***

| White | Red | Total |
|---|---|---|
| 10 | 6 | 16 |

Let $W_1$ = a white ball is picked on the first draw

$R_1$ = a red ball is picked on the 1st draw

$W_2$ = a white ball is drawn on the second draw

So the second ball will be white when either the first ball drawn is white or red

So $P(W_2) = P(R_1 \cap W_2)$ or $P(W_1 \cap W_2)$

$$= \frac{6}{16}\times\frac{10}{15} + \frac{10}{16}\times\frac{9}{15}$$

$$= \frac{60}{240} + \frac{90}{240} = \frac{90}{240} = 0.625$$

## *The 'ONE' and 'ONLY' Situation*

This is a situation in which we may be interested in only one of a series of events to happen , while others fail, or two of them to occur while others fail or none of them to occur. If A, B, C are events with corresponding probabilities of occurrence P(A), P(B), P(C) and probabilities of non-occurrence P(A'), P(B') and P(C')

Assuming independent occurrence of events,

i) P(All of them occur) = P(A $\cap$ B $\cap$ C)
= P(A).P(B).P(C)

ii) P(None of them occur) = P(A' $\cap$ B' $\cap$ C')
= P(A').P(B').P(C')

iii) P(Only one of them occurs)
=P(A $\cap$ B' $\cap$ C') or P(A' $\cap$ B' $\cap$ C') **or** P(A' $\cap$ B' $\cap$ C')
=P(A).P(B').P(C') + P(A').P(B).P(C') + P(A').P(B').P(C)

iv) P(Only two of them occur)
= P(A' $\cap$ B $\cap$ C) + P(A $\cap$ B' $\cap$ C) + P(A $\cap$ B $\cap$ C')
=P(A')×P(B)×P(C)+P(A)×P(B')×P(C)+P(A)×P(B)×P(C')

v) P(At most two of them occur)
= 1 − P(three of them occur)
= 1 − P(A $\cap$ B $\cap$ C)
= 1 − (P(A).P(B).P(C))

***Examples***

1. In a shooting contest, three marks men A, B and C are said to participate. Their respective chances of hitting the target are $\frac{1}{3}$, $\frac{1}{7}$ and $\frac{1}{9}$. If all the three fire at once, find the probability that the target will be hit by:

i) none of them
ii) all of them
iii) only one of them
iv) only two of them
v) at least one of them
vi) at most two of them

***Solution***

(i) P(None hits the target) = P(A' $\cap$ B' $\cap$ C')
=P(A').P(B').P(C')

$$= \frac{2}{3}\times\frac{6}{7}\times\frac{8}{9} = \frac{96}{189} = 0.5079$$

(ii) P(all hit the target)
= P(A $\cap$ B $\cap$ C)

$$= \frac{1}{3}\times\frac{1}{7}\times\frac{1}{9} = \frac{1}{189} = 0.0053$$

(iii) P(only one hits the target)
= P(A $\cap$ B' $\cap$ C') or P(A' $\cap$ B $\cap$ C') or P(A' $\cap$ B' $\cap$ C)
= P(A).P(B').P(C') + P(A').P(B).P(C')+ P(A').P(B').P(C)

$$= \frac{1}{3}\times\frac{6}{7}\times\frac{8}{9} + \frac{2}{3}\times\frac{1}{7}\times\frac{8}{9} + \frac{2}{3}\times\frac{6}{7}\times\frac{1}{9}$$

$$= \frac{48}{189} + \frac{16}{189} + \frac{12}{189} = \frac{76}{189}$$

= 0.4021

(iv) P(two hit the target)
= P(A $\cap$ B $\cap$ C') + P(A $\cap$ B' $\cap$ C) + (A' $\cap$ B $\cap$ C)
=P(A).P(B).P(C') + P(A).P(B').P(C) + (A').P(B).P(C)

$$= \frac{1}{3}\times\frac{1}{7}\times\frac{8}{9} + \frac{1}{3}\times\frac{6}{7}\times\frac{1}{9} + \frac{2}{3}\times\frac{1}{7}\times\frac{1}{9}$$

$$= \frac{8}{189} + \frac{6}{189} + \frac{2}{189} = \frac{16}{189} = 0.0842$$

(v) P(At least one hits the target)
= P(only one hits) + P(two hits) + P(All hit)

$$\frac{76}{189} + \frac{16}{189} + \frac{1}{189} = \frac{93}{189} = 0.4921$$

**OR:** P(At least one hits the target)
= 1 – P(none hits the target)

$$= 1 - \frac{96}{189} = \frac{93}{189}$$

**2**. Three boys John, Patrick and Tom take part in a shooting competition. Their respective probabilities of hitting the target are $\frac{1}{3}$, $\frac{1}{4}$ and $\frac{1}{5}$. If all of them fire at one, find the probability that the target will be hit.

***Solution***

The target may be either hit by John, Patrick or Tom

Let *J* = Event that John hits the target

*P* = Event that Patrick hits the target

*T* = Event that Tom hits the target

Now P(J or P or T) = P(J $\cup$ P $\cup$ T)

P(J $\cup$ P $\cup$ T) = P(J) + P(P) + P(T) – P(J $\cap$ P) – P(J $\cap$ T) – P(P $\cap$ T) + P(J $\cap$ P $\cap$ T)

$$= \frac{1}{3}+\frac{1}{4}+\frac{1}{5}-\frac{1}{3}\times\frac{1}{4}-\frac{1}{3}\times\frac{1}{5}-\frac{1}{4}\times\frac{1}{5}+\frac{1}{3}\times\frac{1}{4}\times\frac{1}{5}$$

$$= \frac{20+15+12}{60} - \frac{1}{12} - \frac{1}{15} - \frac{1}{20} + \frac{1}{60}$$

$$= \frac{47-5-4+1}{60}$$

$$= \frac{36}{60} = 0.6$$

## *Cases involving competing in turns;*

In the previous case, all the participants fire at the target at once. Now in this case, participants take turns to hit at the target and whoever hits the target first becomes the winner and the competition ends

**Examples**

1. Two boys A and B aim at hitting the target. The probability that A hits the target is $\frac{1}{3}$ and that B hits the

target is $\frac{1}{4}$. It is also known that if anyone of the two hits the target first, becomes the winner and the game ends. Suppose that A starts first to hit at the target followed by B, find the probability that:

i) A wins on the second trial
ii) B wins on the $1^{st}$ trial
iii) A wins
iv) B wins

***Solution***

$$P(A)=\frac{1}{3},\ \ P(A')=\frac{2}{3}$$

$$P(B)=\frac{1}{4},\ \ P(B')=\frac{3}{4}$$

**i**) P(A wins on $2^{nd}$ trial) = P(A misses on $1^{st}$ trial and B misses on $1^{st}$ trial and A wins on $2^{nd}$ trial)

$$=P(A_1'\cap B_1'\cap A_2)$$

$$=P(A_1')\cdot P(B_1')\cdot P(A_2)$$

$$=\frac{2}{3}\times\frac{3}{4}\times\frac{1}{3}=\frac{1}{6}$$

***Note***: Since it is A who starts first to hit at the target, once he hits the target, then he becomes the winner and the game ends. So there is no need to proceed to B on the $2^{nd}$ trial.

(ii) P(B wins on the $1^{st}$ trial) $=P(A_1^1\cap B_1)$

$$=P(A_1^1)\cdot P(B_1)$$

$$=\frac{2}{3}\times\frac{1}{4}=\frac{1}{6}$$

(iii) P(A wins) = P(either he wins on the $1^{st}$ trial or $2^{nd}$ or $3^{rd}$ or ..... $n^{th}$ trial)

= sum of a Geometric Progression to infinity

$$=\frac{a}{1-r};$$

Where $a$ = $1^{st}$ term and $r$ = common ratio

Now P(A wins on the $1^{st}$ trial) = P(A) = $\frac{1}{3}$

P(A wins on the $2^{nd}$ trial) $=P(A_1'\cap B_1'\cap A_2)$

$$=\frac{2}{3}\times\frac{3}{4}\times\frac{1}{3}=\frac{1}{6}$$

P(A wins on $3^{rd}$ trial) $=P(A_1'\cap B_1'\cap A_2\cap B_2'\cap A_3)$

$$=\frac{2}{3}\times\frac{3}{4}\times\frac{2}{3}\times\frac{3}{4}\times\frac{1}{3}=\frac{1}{12}$$

Now $a=\frac{1}{3}$ and $r=\frac{1}{6}\div\frac{1}{3}=\frac{1}{2}$

**OR:** $r=\frac{1}{12}\div\frac{1}{6}=\frac{1}{2}$

P(A wins) $=\frac{1/3}{1-\frac{1}{2}}=\frac{1}{3}\div\frac{1}{2}$

$$=\frac{1}{3}\times\frac{2}{1}=\frac{2}{3}$$

(iv) P(B wins) = P(B wins on $1^{st}$ trial) or P(B wins on $2^{nd}$ trial) or P(B wins on $3^{rd}$ trial) or ... P(B wins on $n^{th}$ trial)

$$=P(A_1'\cap B)+P(A_1'\cap B_1'\cap A_2'\cap B)+........$$

$$=\frac{2}{3}\times\frac{1}{4}+\frac{2}{3}\times\frac{3}{4}\times\frac{2}{3}\times\frac{1}{4}+....$$

$$=\frac{1}{6}+\frac{1}{12}+....$$

$a=\frac{1}{6}$ and $r=\frac{1}{12}\times\frac{6}{1}=\frac{1}{2}$

P(B wins) $=\frac{1/6}{1-\frac{1}{2}}$

$$=\frac{1}{6}\div\frac{1}{2}=\frac{1}{6}\times\frac{2}{1}=\frac{1}{3}$$

2. Three participants A, B and C take part in a shooting competition. Their respective probabilities of hitting the target are $\frac{1}{4}$, $\frac{1}{3}$ and $\frac{1}{5}$. It is known that if any one of the three hits the target first becomes the winner and the competition ceases. If it is A who starts first followed by C and then B. Find the probability that:
   (i) C wins on the $2^{nd}$ attempt
   (ii) B wins

***Solution***

P(A) = $\frac{1}{4}$, P(A') = $\frac{3}{4}$

P(B) = $\frac{1}{3}$, P(B') = $\frac{2}{3}$

P(C) = $\frac{1}{5}$, P(C') = $\frac{4}{5}$

(i) P(C wins on $2^{nd}$ attempt) =

$$P\left(A_1^1\cap C_1^1\cap B_1'\cap A_2'\cap C_2'\cap B_2\right)$$

= P($A_1'$) · P($C_1'$) · P($B_1'$) · P($A_2'$) · P($C_2$)

$$=\frac{3}{4}\times\frac{4}{5}\times\frac{2}{3}\times\frac{3}{4}\times\frac{1}{5}=\frac{3}{50}$$

(ii) P(B wins) = P(B wins on $1^{st}$ trial or P(B wins on $2^{nd}$ trial) or …

$$=P\left(A_1'\cap C_1^1\cap B_1\right)+P\left(A_1^1\cap C_1^1\cap B_1'\cap A_2'\cap C_2'\cap B_2\right)+....$$

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

$$=\frac{1}{5}+\frac{2}{25}+...$$

$$=\frac{a}{1-r}$$

$a=\frac{1}{5}$, $r=\frac{2}{25}\times\frac{5}{1}=\frac{2}{5}$

P(B wins) = $\frac{1/5}{1-\frac{2}{5}}$

$$\frac{1}{5}\div\frac{3}{5}$$

$$= \frac{1}{5} \times \frac{5}{3} = \frac{1}{3}$$

## *Use of Probability tree Diagrams*

This is used when dealing with:-

(a) Selection of items out of the lot given
(b) Baye's rule

We shall use probability tree diagrams for Baye's rule when looking at it in the next section

**Selection of items:**

Here there must be at least two possible outcomes for every trial of the experiment performed.

**1**. Three items are randomly picked from a bag containing 10 defective (D) and 30 non defective (N) items. Find the probability that $\frac{2}{3}$ of these items are defective if:

(i) the items are picked with replacement
(ii) the items are picked without replacement

***Solution***

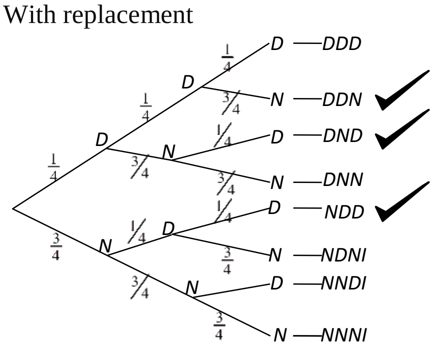
| Defective | Non defective | Total |
|---|---|---|
| 10 | 30 | 40 |

$$P(D) = \frac{10}{40} = \frac{1}{4}$$

$$P(N) = \frac{30}{40} = \frac{3}{4}$$

Since $n$ = 3, we need two defective items out of 3.

(i) With replacement

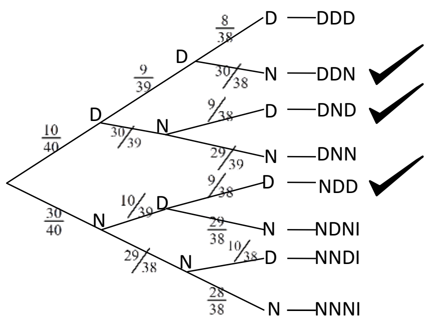



P(Two items are defective) = P(DDN) + P(DND) + P(NDD)

$$= \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} + \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} + \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4}$$

$$= \frac{3}{64} + \frac{3}{64} + \frac{3}{64}$$

$$= \frac{9}{64}$$




P(Two items are defective) = P(DDN) + P(DND) + P(NDD)

$$= \frac{10}{40} \times \frac{9}{39} \times \frac{30}{38} + \frac{10}{40} \times \frac{30}{39} \times \frac{9}{38} + \frac{30}{40} \times \frac{10}{39} \times \frac{9}{38}$$

$$= \frac{2700}{59280} + \frac{2700}{59280} + \frac{2700}{59280}$$

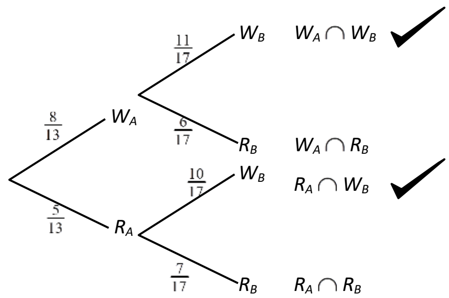
$$= \frac{8100}{59280} = 0.1366$$

2. Two baskets A and B contain similar balls. A contains 8 white and 5 red balls where as B contains 10 white and 6 red balls. A ball is randomly drawn from basket A and transferred to basket B. a ball is then randomly drawn from B, find the probability that it will be white.

| | White | Red | Total |
|---|---|---|---|
| A | 8 | 5 | 13 |
| B | 10 | 6 | 16 |

Let $W_A$ = White ball is drawn from A
$W_B$ = White ball is drawn from B
$R_A$ = Red ball is drawn from A
$R_B$ = Red ball is drawn from B

When a ball is transferred from A to B, the total number of balls in B will now be 17




P(White ball is picked from B)

$= \text{P}(W_A \cap W_B)$ or $\text{P}(R_A \cap W_B)$

$= \text{P}(\text{W}_\text{A}) \cdot \text{P}(\text{W}_\text{B}/\text{W}_\text{A}) + \text{P}(\text{R}_\text{A}) \cdot \text{P}(\text{W}_\text{B}/\text{R}_\text{A})$

$$= \frac{8}{13} \times 11 + \frac{5}{13} \times \frac{10}{17}$$

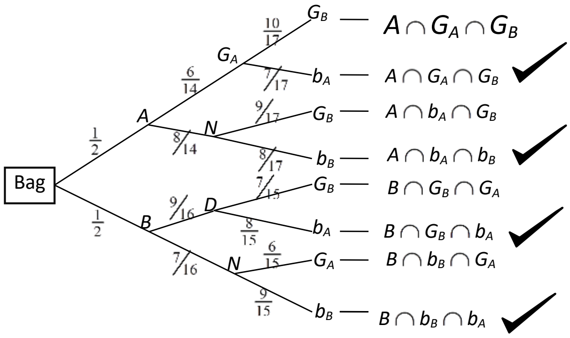
$$= \frac{88}{221} + \frac{50}{221} = \frac{138}{221} = 0.6244$$

**3.** Two bags A and B contain similar marbles. A contains 6 green and 8 blue marbles where as B contains 9 green and 7 blue marbles. A bag is picked at random unseen and from it one marble is picked and placed in the other bag. If a marble is finally picked from this second bag, find the probability that the marble selected is blue

| | Green | Blue | Total |
|---|---|---|---|
| A | 6 | 8 | 14 |
| B | 9 | 7 | 16 |

Let $G_A$ = Green marble is picked from bag A
$b_A$ = blue marble is picked from bag A
$G_B$ = green marble is picked from bag B
$b_B$ = blue marble is picked from B

Since a bag is picked at random from the two bags available, it means that the probability of picking either bag is $\frac{1}{2}$




P(Marble picked from 2nd bag is blue)

$= P(A \cap G_A \cap b_B) + P(A \cap b_A \cap b_B) + P(B \cap G_B \cap b_B) + P(b \cap b_B \cap b_B)$

$$= \frac{1}{2}\times\frac{6}{14}\times\frac{7}{17}+\frac{1}{2}\times\frac{8}{14}\times\frac{8}{17}+\frac{1}{2}\times\frac{9}{16}\times\frac{8}{15}+\frac{1}{2}\times\frac{7}{16}\times\frac{9}{15}$$

$$= \frac{42}{476}+\frac{64}{476}+\frac{72}{480}+\frac{63}{480}$$

$$= \frac{106}{476}+\frac{135}{480}$$

$= 0.222689 + 0.28125$

$= 0.5040$ (4 dp)

**4.** Two boxes P and Q contain similar balls. Box Q contains 6 red and 4 blue balls while box P contains 3 red and 5 blue balls. A box is randomly drawn and from it a ball is randomly drawn and put into the other. A ball is then drawn from this latter box. Find the probability that

i) both balls are red

ii) the first ball drawn is blue

iii) the first ball drawn is red given that the second ball drawn is blue

***Solution***

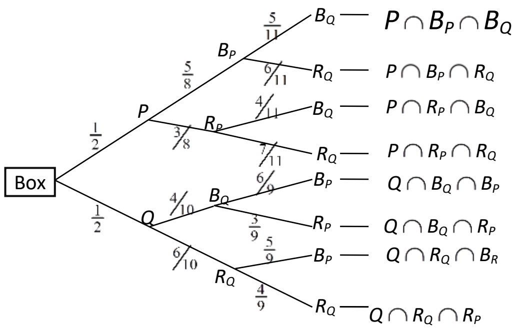
| | Red | Blue | Total |
|---|---|---|---|
| P | 3 | 5 | 8 |
| Q | 6 | 4 | 10 |

Let $B_P$ = a blue ball is drawn from box P

$R_P$ = red ball is drawn from box P

$B_Q$ = blue ball is drawn from box Q

$R_Q$ = red ball is drawn from box Q




(i) P(both balls are red) $= P(P \cap R_P \cap R_Q) + P(Q \cap R_Q \cap R_P)$

$$= \frac{1}{2}\times\frac{3}{8}\times\frac{7}{11}+\frac{1}{2}\times\frac{6}{10}\times\frac{4}{9}$$

$$= \frac{21}{176}+\frac{24}{180}$$

$= 0.119318 + 0.133333$

$= 0.252651$

(ii) P(1st ball drawn is blue)

$= P(P \cap B_P \cap B_Q) + P(P \cap B_P \cap R_Q) + P(Q \cap B_Q \cap B_P) + P(Q \cap B_Q \cap R_P)$

$$= \frac{1}{2}\times\frac{5}{8}\times\frac{5}{11}+\frac{1}{2}\times\frac{5}{8}\times\frac{6}{11}+\frac{1}{2}\times\frac{4}{10}\times\frac{6}{9}+\frac{1}{2}\times\frac{4}{10}\times\frac{3}{9}$$

$$= \frac{25}{176}+\frac{36}{176}+\frac{24}{180}+\frac{12}{180}$$

$$= \frac{55}{176}+\frac{36}{180} = 0.3125+0.2$$

$= 0.5125$

(iii) P(1st ball is red given the second ball is blue)

Let A = first ball drawn is red

B = second ball drawn is blue

Now $P(A/B) = \dfrac{P(A\cap B)}{P(B)}$

$P(A \cap B)$ = P(1st ball is red and 2nd ball is blue)

$= P(P \cap R_P \cap B_Q) + P(Q \cap R_Q \cap B_P)$

$$= \frac{1}{2}\times\frac{3}{8}\times\frac{4}{11}+\frac{1}{2}\times\frac{6}{10}\times\frac{5}{9}$$

$$= \frac{12}{176}+\frac{30}{180}$$

$= 0.2348484848$

P(B) = P(2nd ball is blue)

$= P(P \cap B_P \cap B_Q) + P(P \cap R_P \cap B_Q) + P(Q \cap B_Q \cap B_P) + P(Q \cap R_Q \cap B_P)$

$$= \frac{1}{2}\times\frac{5}{8}\times\frac{5}{11}+\frac{1}{2}\times\frac{3}{8}\times\frac{4}{11}+\frac{1}{2}\times\frac{4}{10}\times\frac{6}{9}+\frac{1}{2}\times\frac{6}{10}\times\frac{5}{9}$$

$$= \frac{25}{176}+\frac{12}{176}+\frac{24}{180}+\frac{30}{180}$$

$$= \frac{37}{176}+\frac{54}{180}$$

$= 0.5102272727$

$$P(A/B) = \frac{0.2348484848}{0.5102272727} = 0.46028$$

**5.** Three bags A, B and C contain respectively 3 white and 2 red balls, 4 white and 4 red, 5 white and 2 red balls. A ball is drawn unseen from A and placed in B. a ball is then drawn from B and placed in C. find the probability that if a ball is now drawn from C, it will be red.

***Solution***

| | White | Red | Total |
|---|---|---|---|
| A | 3 | 2 | 5 |
| B | 4 | 4 | 8 |
| C | 5 | 2 | 7 |

Let $R_A$ = red ball is drawn from bag A

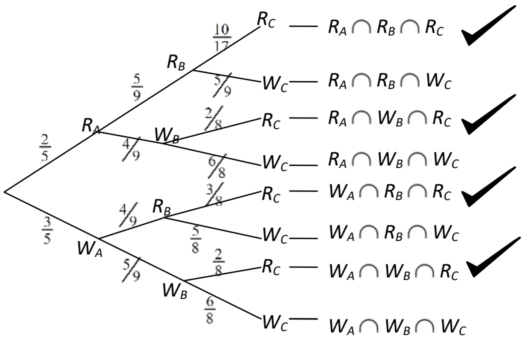
$R_B$ = red ball is drawn from bag B

$R_C$ = red ball is drawn from bag C

$W_A$ = white ball is drawn from bag A

$W_B$ = white ball is drawn from bag B

$W_C$ = white ball is drawn from bag C




P(ball drawn from C is red) =
$P(R_A \cap R_P \cap R_C) + P(R_A \cap W_B \cap R_C) + P(W_A \cap R_B \cap R_C) + P(W_A \cap W_B \cap R_C)$

$$= \frac{2}{5}\times\frac{5}{9}\times\frac{3}{8}+\frac{2}{5}\times\frac{4}{9}\times\frac{2}{8}+\frac{3}{5}\times\frac{4}{9}\times\frac{3}{8}+\frac{3}{5}\times\frac{5}{9}\times\frac{2}{8}$$

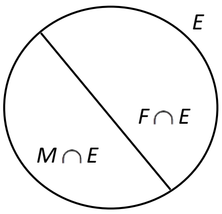
$$= \frac{30}{360}+\frac{16}{360}+\frac{36}{360}+\frac{30}{360}$$

$$= \frac{112}{360} = 0.3111$$

**Note:** When using tree diagrams, you make sure that the sum of probabilities for every branch is 1

## *Total Theorem*

Suppose that at St. Mary's College Kisubi there are two categories of teachers; males (M) and Females (F). Among these two categories, there are UNEB Examiners (E). If in a survey of these teachers at the school we are interested in randomly picking an Examiner from the school, then the examiner will either be a male or a female. This is illustrated in the diagram below.




The examiners are sub-divided into two non-intersecting groups called partitions

**Note**: We say a person is either a male and an examiner or a female and an examiner.

So E = $M \cap E$ or $F \cap E$

$P(E) = P(M \cap E) + P(F \cap E)$

$= P(M)\cdot P(E/M) + P(F)\cdot P(E/F)$

This is called the theorem of total probability.

In general, if $B_1, B_2, B_3 \ldots B_n$ constitute partitions of a sample space *S*, then for any event A in S

$A = (B_1 \cap A)$ or $(B_2 \cap A)$ or $(B_3 \cap A)$ or … or $(B_n \cap A)$

$P(A) = P(B_1 \cap A) + P(B_2 \cap A) + P(B_3 \cap A) + \ldots + P(B_n \cap A)$

$= P(B_1)\cdot P(A/B_1) + P(B_2)\cdot P(A/B_2) + \ldots + P(B_n)\cdot P(A/B_n)$

## *Baye's Rule*

Now from the total theorem just explained above, if a teacher is picked at random and found out to be an examiner, the probability that this teacher is a male is given by P(M/E)

Now $P(M/E) = \dfrac{P(M \cap E)}{P(E)}$

But $P(E) = P(M \cap E) + P(F \cap E)$

Hence $P(M/E) = \dfrac{P(M)\cdot P(E/M)}{P(M)\cdot P(E/M) + P(F)\cdot P(E/F)}$

This equation represents Baye's rule.

Generally, for *n* partitions, $B_1, B_2, B_3, \ldots B_n$ of sample space S, for any event A in the sample space S, we have;

$$P(B_i/A) = \frac{P(B_i \cap A)}{P(A)} \text{ where } i = 1, 2, 3, \ldots, n$$

***Examples***

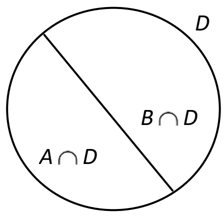
1. In a manufacturing plant, there are two different machines A and B. 20% and 80% of the items are produced by A and B respectively. It has been established that 5% of the items produced by A and 8% of the items produced by B are defective. If one item is selected at random from the lot produced, determine the probability that:
   (i) It is defective
   (ii) It is produced by A given that it is defective

***Solution***

Let A = item is produced by A
B = item is produced by B
D = item is defective

Now P(A) = 20% = 0.2
P(B) = 80% = 0.8
P(D/A) = 5% = 0.05
P(D/B) = 8% = 0.08

(i)




$P(D) = P(A \cap D) + P(B \cap D)$
$= P(A)\cdot P(D/A) + P(B)\cdot P(D/B)$
$= 0.2 \times 0.05 + 0.8 \times 0.08$
$= 0.01 + 0.064$
$= 0.074$

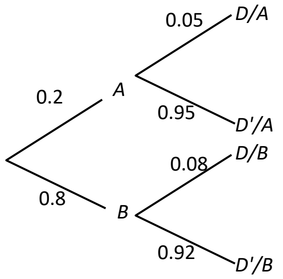
(ii) $P(A/D) = \dfrac{P(A \cap D)}{P(D)}$

$$= \frac{P(A)\cdot P(D/A)}{P(D)}$$

$$= \frac{0.2\times 0.05}{0.074} = \frac{0.01}{0.074} = 0.1351$$

**Alternatively**

(i) By using the probability factor tree diagram



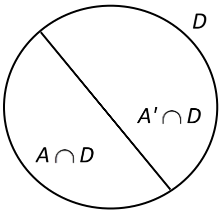

$$P(D) = P(A)\cdot P(D/A) + P(B)\cdot P(D/B)$$
$$= 0.2 \times 0.05 + 0.8 \times 0.08$$
$$= 0.01 + 0.064 = 0.074$$

**2.** A certain device has probability $\frac{1}{3}$ of detecting an 'AIDS' victim as having AIDS and probability of $\frac{1}{4}$ of detecting 'Non-AIDS' victim as having AIDS. The device is used on a population, 20% of which have AIDS. What is the probability that a person is detected as an AIDS victim by the device?

***Solution***

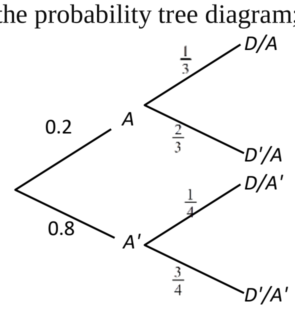
Let $D$ = detecting a person as having AIDS
$A$ = person has AIDS
$A'$ = person has no AIDS

**Note**: A person could be detected as having AIDS when he has AIDS or when he has no AIDS




$$P(D) = P(A\cap D) + P(A'\cap D)$$
$$= P(A)\cdot P(D/A) + P(A')\cdot P(D/A')$$
$$= 0.2\times\frac{1}{3}+0.8\times\frac{1}{4}$$
$$= \frac{0.2}{3}+\frac{0.8}{4}$$
$$= \frac{0.8+2.4}{12} = \frac{3.2}{12} = 0.2667$$

**OR** by using the probability tree diagram;

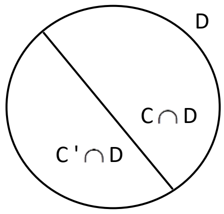



$$P(D) = P(A)\cdot P(D/A) + P(A')\cdot P(D/A')$$
$$= 0.2\times\frac{1}{3}+0.8\times\frac{1}{4}$$
$$= \frac{3.2}{12} = 0.2667$$

**3**. In a Kampala clinic, it is known that the probability of selecting a person with cancer is 0.02. If the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.7 and the probability of incorrectly diagnosing a person without cancer as having the disease is 0.05, what is the probability that a person is diagnosed as having cancer?

***Solution***

Let C = A person has cancer
D = A person is diagnosed as having cancer




$$P(D) = P(C\cap D) + P(C'\cap D)$$
$$= P(C)\cdot P(D/C) + P(C')\cdot P(D/C)$$
$$= 0.02 \times 0.7 + 0.98 \times 0.05$$
$$= 0.014 + 0.049 = 0.063$$

4. In a certain city, 30% of the people are conservatives, 50% are liberals and 20% are independents. Records indicate that in an election, 65% of the conservatives voted, 85% of the liberals voted and 50% of the independents voted. A person in the city is selected at random.

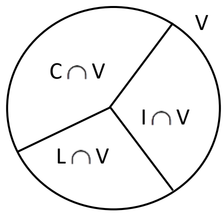
i) Determine the probability that he voted
ii) Given that he didn't vote, determine the probability that he is a conservative

***Solution***

Let C = A person is conservative
L = A person is liberal
I = A person is independent
V = A person voted

Now P(L) = 0.5, P(C) = 0.3, P(I) = 0.2,
P(V/L) = 0.85, P(V/C) = 0.65, P(V/I) = 0.5

(i)




$$P(V) = P(C\cap V) + P(L\cap V) + P(I\cap V)$$
$$= P(C)\cdot P(V/C) + P(L)\cdot P(V/L) + P(I)\cdot P(V/I)$$
$$=0.3 \times 0.65 + 0.5 \times 0.85 + 0.20 \times 0.5$$
$$= 0.195 + 0.425 + 0.1$$
$$= 0.72$$

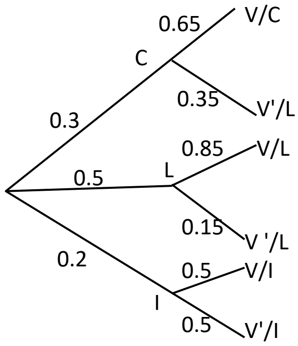
(ii) $P(C/V') = \dfrac{P(C\cap V')}{P(V')} = \dfrac{P(C)\cdot P(V'/C)}{1-P(V)}$

But P(V'/C) = 1 – P(V/C) = 1 – 0.65
= 0.35

$\Rightarrow$ P(V'/C) = $\dfrac{0.3\times 0.35}{1-0.72}$

$$= \frac{0.105}{0.28} = 0.375$$

**OR**; by using the probability tree diagram;




$$P(V) = P(C)\cdot P(V/C) + P(L)\cdot P(V/L) + P(I)\cdot P(V/I)$$

$= 0.3 \times 0.65 + 0.5 \times 0.85 + 0.20 \times 0.5$

$= 0.195 + 0.425 + 0.1 = 0.72$

**5.** In a certain university, 75% of the students are full time students, 45% of the students are female, and 40% of the students are male full time students. Find the probability that:

a) a student chosen at random from the university is a part time student

b) a student chosen at random from the university is a female and a part-time student.

c) a student chosen at random from the female students in the university is a part-time student.

***Solution***

Let $U$ = Event that student is a full-time student

Then $U'$ = Event that a student is a part time

$F$ = Female student

$M$ = Male student

$\Rightarrow$ $P(U) = 0.75$

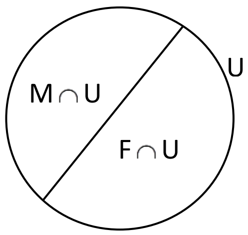
$P(F) = 0.45$

$P(M \cap U) = 0.4$

(a) $P(U') = 1 - P(U)$

$= 1 - 0.75 = 0.25$

b) We need to find $P(F \cap U')$ but a student is full time when he is male or she is a female.




$P(U) = P(M \cap U) + P(F \cap U)$

$0.75 = 04 + P(F \cap U)$

$P(F \cap U) = 0.35$

But $P(F) = P(F \cap U) + P(F \cap U')$

$0.45 = 0.35 + P(F \cap U')$

$P(F \cap U') = 0.1$

c) Here, we need to find $P(U'/F)$

But $P(U'/F) = \dfrac{P(U' \cap F)}{P(F)}$

$P(U'/F) = \dfrac{0.1}{0.45} = 0.2222$

# Examination Questions

**1**. (a) A box contains 3 red, 2 green and 5 blue crayons. Two crayons are randomly selected from the box without replacement, find the probability that:

i) the crayons are of the same colour

ii) at least one red crayon is selected

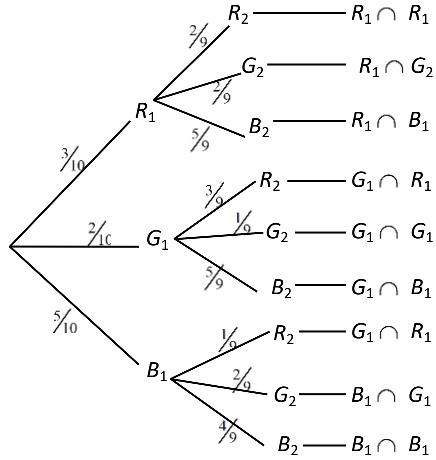
b) In an experiment, a box contains 2 green and 5 blue balls. A second box contains 5 green and 3 blue balls. One ball is drawn at random from the second box and placed into the first box. What is the probability that a ball now drawn at random from the first box is green? **(1990 No. 13)**

**Solution**

a)

| Red (R) | Green (G) | Blue (B) | Total |
|---|---|---|---|
| 3 | 2 | 5 | 10 |

$P(R) = \dfrac{3}{10}$, $P(G) = \dfrac{2}{10}$ and $P(B) = \dfrac{5}{10}$




(i) P(All are of the same colour)

$= P(R_1 \cap R_2) + P(G_1 \cap G_2) + P(B_1 \cap B_1)$

$= \dfrac{3}{10}\times\dfrac{2}{9} + \dfrac{2}{10}\times\dfrac{1}{9} + \dfrac{5}{10}\times\dfrac{4}{9}$

$= \dfrac{6}{90} + \dfrac{2}{90} + \dfrac{20}{90}$

$= \dfrac{28}{90} = \dfrac{14}{45}$

(ii) P(at least one is red) $= P(R_1 \cap R_2) + P(R_1 \cap G_2) + P(R_1 \cap G_1) + P(G_1 \cap R_2) + P(B_1 \cap R_2)$

$= \dfrac{3}{10}\times\dfrac{2}{9} + \dfrac{3}{10}\times\dfrac{2}{9} + \dfrac{3}{10}\times\dfrac{5}{9} + \dfrac{2}{10}\times\dfrac{3}{9} + \dfrac{5}{10}\times\dfrac{3}{9}$

$= \dfrac{6}{90} + \dfrac{6}{90} + \dfrac{15}{90} + \dfrac{6}{90} + \dfrac{15}{90}$

$= \dfrac{48}{90} = \dfrac{8}{15}$

(b)

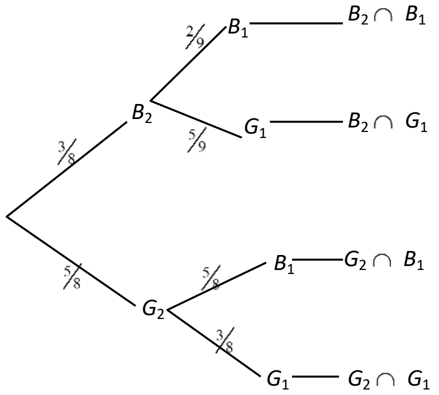
| | Green(G) | Blue(B) | Total |
|---|---|---|---|
| First box | 2 | 5 | 7 |
| Second box | 5 | 3 | 8 |

Let $B_1$ = Blue ball from first box

$G_1$ = green ball from first box

$B_2$ = Blue ball from second box

$G_2$ = Green ball from second box




$P(G_1) = P(B_2 \cap G_1) + P(G_2 \cap G_1)$

$$= \frac{3}{8} \times \frac{2}{8} + \frac{5}{8} \times \frac{3}{8} = \frac{6}{64} + \frac{15}{64} = \frac{21}{64}$$

Hence $P(G_1) = \frac{21}{64}$

**2** (a) Three balls are drawn at random one after the other without replacement from a bag containing 21 white, 9 blue, 40 red and 12 orange balls. Determine the probability that the first ball is blue, the second red or blue and the third is white.

b) Tom is to travel from Lira to Kampala for an interview. The probabilities that he will be in time for the interview when he travels by bus and taxi are 0.1 and 0.2 respectively. The probabilities that he will travel by bus and taxi are 0.6 and 0.4 respectively.

i) Find the probability that he will be on time

ii) Given that he is not on time, what is the probability that he travelled by taxi? **(1991 No 10)**

***Solution***

a)

| White | Blue | Red | Orange | Total |
|---|---|---|---|---|
| 21 | 9 | 40 | 12 | 82 |

Let $W$ = White ball

$B$ = Blue ball

$R$ = Red ball

P(1st is blue and second is red or blue and third is white)

$= P(B_1 \cap R_2 \cap W_3) + P(B_1 \cap B_2 \cap W_3)$

$$= \frac{9}{82} \times \frac{40}{81} \times \frac{21}{80} + \frac{9}{82} \times \frac{8}{81} \times \frac{21}{80}$$

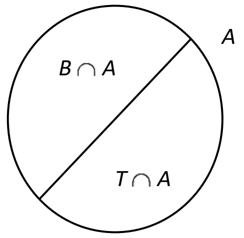
$$= \frac{7560}{531360} + \frac{1512}{531360} = \frac{9072}{531360} = 0.0171$$

b) Let B = Travelling by bus

T = Travelling by tax

A = arriving on time for the meeting

(i)




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

$= P(B) \cdot P(A/B) + P(T) \cdot P(A/T)$

$= 0.6 \times 0.1 + 0.08$

$= 0.14$

(ii) $P(T/A) = \dfrac{P(T \cap A')}{P(A')}$

$$= \frac{P(T \cap A')}{1 - 0.4} = \frac{P(T \cap A')}{0.86}$$

But $P(T) = P(T \cap A) + P(T \cap A')$

$0.4 = 0.8 + P(T \cap A')$

$P(T \cap A') = 0.32$

$\Rightarrow P(T/A') = \dfrac{0.32}{0.86} = 0.3721$

***OR*** $P(T/A') = \dfrac{P(T \cap A')}{P(A')}$

$$= \frac{P(T) \cdot P(A'/T)}{P(A')} = \frac{0.4 \times [1 - P(A/T)]}{0.86}$$

$$= \frac{0.4 \times (1 - 0.2)}{0.86} = \frac{0.4 \times 0.8}{0.86} = 0.3721$$

**3**. The probability of Mukasa waking up late is 0.3. When he wakes up late, the probability of being in time for school is 0.2. When he wakes up early, the probability of being punctual for school is 0.9. What is the probability of Mukasa:

i) waking up late and being punctual?

ii) waking up early and being punctual?

iii) If he is on time, what is the probability that he woke up early?

iv) What is the probability of being late for school?

***Solution***

Let $L$ = Waking up late

$T$ = being in time or punctual for school

Then $L'$ = Waking up early

Now $P(L) = 0.3$,

$P(T/L) = 0.2$

$P(T/L') = 0.9$

(i) $P(L \cap T) = P(L) \cdot P(T/L)$

$= 0.3 \times 0.2 = 0.06$

(ii) $P(L' \cap T) = P(L') \cdot P(T/L')$

$= 0.7 \times 0.9 = 0.63$

(iii) $P(L'/T) = \dfrac{P(L' \cap T')}{P(T)}$

$$= \frac{P(L' \cap T)}{P(L \cap T) + P(L' \cap T)}$$

$$= \frac{0.63}{0.06 + 0.063} = \frac{0.63}{0.69} = 0.913$$

(iv) $P(T') = 1 - P(T)$

$= 1 - 0.913 = 0.087$

**4**. (a) Two biased tetrahedrons have each their faces numbered 1 to 4. The chances of getting any one face showing uppermost is inversely proportional to the number on it. If the two tetrahedrons are thrown and the number on the uppermost face noted, determine the probability that the faces show the same number.

(b) If it is a fine day, the probability that Alex goes to play football is $\frac{9}{10}$ and the probability that Bob goes is

$\frac{3}{4}$. If it is not fine, Alex's probability is $\frac{1}{2}$ and Bob's is $\frac{1}{4}$. Their decisions are independent. In general, it is known that it is twice as likely to be fine as not fine.

(i) Determine the probability that both go to play

(ii) If they both go to play, what is the probability that is a fine day? **(1994 No 12)**

**Answer: (a) 0.328 (b)(i)** $\frac{59}{120}$ **(ii)** $\frac{54}{120}$

**5**. (**a**) In an examination, only two papers, namely mathematics and Physics were done. The failure rates were 45% and 40% respectively.

The number of candidates who sat for the examination was 2000. Find the probability that a candidate selected at random

(i) failed both mathematics and physics

(ii) passed both mathematics and physics

(iii) passed mathematics and failed physics

Determine the number of candidates who passed both papers in other grades given that 21.8% and 22.9% passed with distinction in mathematics and physics respectively

(**b**) When visiting a friend, john may go by road, air or rail. The probabilities of using road, air or rail are 0.3, 0.8 and 0.6 respectively. The corresponding probabilities of arriving on an agreed time are 0.2, 0.8 and 0.1 respectively. Find the probability of having used the road given that he arrived on time

**(1995 No. 1)**

**Answer** (a) (i) 0.18 (ii) 0.33 (iii) 0.22 (b) 0.0789

**6**. In a survey of newspaper reading of members of staff of a university, it is found that 80% read NEW VISION (N), 50 percent read MONITOR (*M*) and 30% read the EAST AFRICAN (*E*). Further, 20% read both *M* and *N*, 15% read both *N* and E and 10% read both *M* and E.

a) If a member of staff is chosen at random from the university, find the probabilities

i. that the member reads none of the three papers

ii. the member is one of those who read at least one of the three papers.

b) Estimate the number of members of staff who read at least two papers if the total number is 500

**c**) What is the probability that given that a member of staff reads two papers, he reads all the three?

**(1996 No. 14)**

**Answer:** (a) ***This question was incorrectly set!***

**7.** There are 3 black and 2 white balls in each of the two bags. A ball is taken from the first bag and put in the second, then a ball is taken from the second into the first, what is the probability that there are now the same number of black and white balls in each bag as there were to begin with? **(1998 March No 5)**

**Answer:** $\frac{3}{5}$

**8.** 64% of the students at A' level take science subjects and 36% do Arts subjects. The probability of them being successful is ¾ for Science students and 5/6 for Arts students. Find the probability that a student chosen at random will fail. **(1998 Mar. No. 8)**

**Answer: 0.22**

**9.** The probability that two independent events occur together is $\frac{2}{15}$. The probability that either or both events occur is $\frac{3}{5}$. Find the individual probabilities of the two events. **(1998 Nov/Dec No 1)**

**Answer:** $\frac{1}{3}$ and $\frac{2}{5}$.

**10**. Given that A and *B* are mutually exclusive events and $P(A) = \frac{2}{3}$ and $P(B) = \frac{1}{2}$, find:

i) $P(A \cup B)$, ii)$P(A \cap B^1)$ iii)$P(A^1 \cap B^1)$.

**(1999 No 1)**

**Answer:** (i) $\frac{9}{10}$ (ii) $\frac{2}{5}$ (iii) $\frac{1}{10}$

**11**. At a bus park, 60% of the buses are of Isuzu make, 25% are styer type and the rest are of Tata make. Of the Isuzu type, 50% have radios while for the Styer and Tata make types ,only 5% and 1% have radios, respectively. If a bus is selected at random from the park, determine the probability that:

i) it has a radio

ii) a styer type is selected given that it has a radio

**(1999 No. 7)**

**Answer**: (i) 0.0315 (ii) 0.0398

**12**. A family plans to have 3 children.

i) Write down the possible sample space and construct its probability distribution table.

ii) Given that X is the number of boys in the family, find the expected number of boys. **(2000 No. 1)**

**Answer:**

(i)

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |
| $xp(X=x)$ | 0 | $\frac{3}{8}$ | $\frac{6}{8}$ | $\frac{3}{8}$ |

(ii) 1.5

**13**. Two balls are randomly drawn without replacement from a bag containing 10 white and 6 red balls. Find the probability that the second ball drawn is

b) red given that the first one was white

*c)* white **(2000 No. 3)**

**Answer** (i) 0.4 (ii) 0.375

**14**. The events A and *B* are neither independent nor mutually exclusive. Given that $P(B) = 1/3$, $P(A)= ½$ and $P(A \cap B^1) = 1/3$, Find:

(i) $P(A^1 \cup B^1)$, (ii) $P(A^1/B^1)$ **(2001 No. 1)**

**Answer:** (i) $\frac{5}{6}$ (ii) $\frac{1}{2}$

**15**. (a) Bag A contains 3 green and 2 blue balls, while bag *B* contains 2 green and 3 blue balls. A bad is selected at random and 2 balls drawn from it without replacement. Find the probability that the balls are of different colours.

(b) A fair die is drawn 6 times. Calculate the probability that

i) A 2 or 4 appears on the first throw,

ii) Four 5s will appear in the 6 throws.

**(2001 No. 10)**

**Answer:** $(a)\ \dfrac{19}{30}\qquad (b)(i)\dfrac{1}{3}\quad (ii)\,0.008$

**16**. On a certain day, fresh fish from lakes: Kyoga, Victoria, Albert and George were supplied to one of the central markets of Kampala in the ratios 30%, 40%, 20% and 10% respectively. Each lake had an estimated ratio of poisoned fish of 2%, 3% and 1% respectively. If a health inspector picked a fish at random,

i. What is the probability that the fish was poisoned?
ii. Given that the fish was poisoned, what is the probability that it was from Lake Albert? **(2002 No. 1)**

**Answers:** (i) 0.025 (ii) 0.24

**17.** The chance that a person picked from a Kampala street is 30 in every 48. The probability that that a person is a university graduate given that he is employed is 0.6. Find the

i) probability that the person picked at random from the street is a university graduate and is employed
ii) number of people that are not university graduates and are employed from a group of 120 people. **2002 No. 4)**

**Answers**: (i) 0.0.375 (ii) 30

**18**. Events A and B are such that P(A) = ½,

P(B) = 3/8 and P(A/B) = 7/12, find

(i) P(A$\cap$B) (ii) $P(B/\overline{A})$ **(2003 No. 1)**

**Answer**: (i) $\dfrac{7}{32}$ (ii) $\dfrac{5}{16}$

**19.** The probability of two independent events *P* and *Q* occurring together is 1/8. The probability that either or both events occur is 5/8. Find

a) P(*P*)
b) P(*Q*) **(2004 No. 2)**

**Answer: (a)** $\left(\dfrac{1}{2},\dfrac{1}{4}\right)$ (b) $\left(\dfrac{1}{4},\dfrac{1}{2}\right)$

**20.** a) Abel, Bob and Charles applied for the same job in a certain company. The probability that Abel will take the job is ¾. The probability that *B*ob will take it is ½ while the probability that Charles won't take the job is 1/3. What is the probability that:

i) none of them will take the job?
ii) one of them will take the job?

b) Two events *A* and *B* are independent. Given that $P(A\cap B^1) = \frac{1}{4}$ and $P(A^1/B) = \frac{1}{6}$, use a Venn diagram to find the probabilities

i) $P(A)$
ii) $P(B)$
iii) $P(A\cap B)$
iv) $P(A\cup B)^1$. **(2004 No. 9)**

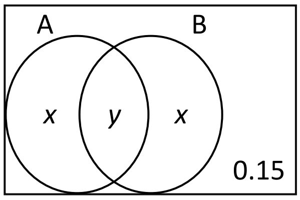
**Answer:** (a) (i) $\dfrac{1}{24}$ (ii) $\dfrac{1}{4}$

$(b)(i)\ \dfrac{5}{6},\quad (ii)\ \dfrac{7}{10}\quad (iii)\ \dfrac{7}{12}\quad (iv)\ \dfrac{1}{20}$

**21**. A good football striker is nursing his injury in the leg. The probability that his team will win the next match when he is playing is 4/5, otherwise it is 2/3. The probability that he will have recovered by the time of the match is ¼. Find the probability that his team will win the match. **(2005 No. 3)**

**Answer:** $\dfrac{7}{10}$.

**22**. a) A and *B* are intersecting sets as shown in the Venn diagram below.




Given that *P*(A) = 0.6, $P\left(A^1/B\right)=\dfrac{5}{7}$, and $P(A\cup B) = 0.85$, find

(i) the value of *x*, *y* and *z* (ii) $P(A/B)$

b) A bag contains 4 white balls and 1 black ball. A second bag contains 1 white ball and 4 black balls. A ball is drawn at random from the first bag and put into the second bag, then a ball is taken from the second bag and put into the first bag. Find the probability that a white ball will be picked when a ball is selected from the first bag. **(2005 No. 9)**

**Answer:** (a) (i) 0.5, 0.1 and *z* = 0.25 (ii) $\dfrac{2}{7}$ b) $\dfrac{7}{10}$

**23.** A and B are two independent events with A twice as likely to occur as B. If $P(A)=\dfrac{1}{2}$, find:

(ii) $P(A\cup B)$

(iii) $p\left[(A\cap B)/A\right]$ **(2006 No. 1)**

**Answer: (i)** $\frac{5}{8}$ (ii) $\frac{1}{8}$

**24**. The table below shows the likelihood of where A and B spend a Saturday evening.

| | A | B |
|---|---|---|
| Goes to dance | $\frac{1}{2}$ | $\frac{2}{3}$ |
| Visits a neighbour | $\frac{1}{3}$ | $\frac{1}{6}$ |
| Stays at home | $\frac{1}{6}$ | $\frac{1}{6}$ |

(i) Find the probability that they both go out.
(ii) If we know they both go out, what is the probability that they both went to dance?

**(2007 No. 5)**

**Answer:** (i) $\dfrac{25}{36}$ (ii) $\dfrac{12}{25}$

**25** (a) A box contains 7 red balls and 6 blue balls. Three balls are selected at random without replacement. Find the probability that:
(i) they are of the same colour.
(ii) at most two are blue.

(b) Two boxes P and Q contain white and brown cards. P contains 6 white cards and 4 brown cards. Q contains 2 white cards and 3 brown cards. A box is selected at random and a card is selected.
Find the probability that:
(i) a brown card is selected.
(ii) box Q is selected given that the card is white.

**(2007 No. 15)**

**Answer:**
(a) (i) 0.1923 (ii) 0.9301 (b) (i) 0.5 (ii) 0.4

**26.** The probability that Anne reads the *New Vision* is 0.75 and the probability that she reads the *New Vision* and not the *Daily Monitor* is 0.65. The probability that she reads neither of the papers is 0.15. Find the probability that she reads the *Daily Monitor*. **(2008 No. 1)**

**Answer:** 0.2

**27.** If A and B are independent events;
(i) show that events A and B' are also independent
(ii) find P(B) given that P(A) = 0.4 and $P(A \cup B) = 0.8$
**(2009 No. 9)**

**Answer:** (ii) 0.667

**28.** A box contains two types of balls, red and black. When a ball is picked from the box, the probability that it is red is $7/12$. Two balls are selected at random from the box without replacement.
Find the probability that
(i) the second ball is black
(ii) the first ball is red, given that the second one is black
b) An interview involves written, oral and practical tests. The probability that an interviewee passes a written test is 0.8, the oral test is 0.6 and the practical test is 0.7. What is the probability that the interviewee will pass
(i) the entire interview?
(ii) exactly two of the interview tests? **(2009 No. 13)**

**Answer:** (i) $\frac{5}{11}$ (ii) $\frac{7}{11}$ (b) (i) 0.336 (ii) 0.452

**29.** Two events *M* and *N* are such that $P(M) = 0.7$, $P(M \cap N) = 0.45$ and $P(M' \cap N') = 0.18$. Find:
(a) $P(N')$,
(b) P(*M* or *N* but not both *M* and *N*)
**(2010 No. 1)**

**Answers**: (a) 0.43 (b) 0.37

**30.** (a) The probabilities that three players A, B and C score in a netball game are $1/5$, $1/4$ and $1/3$ respectively. If they play together in a game, what is the probability that:
(i) only C scores,
(ii) at least one player scores,
(iii) two and only two players score.

(b) There are 100 students taking principal mathematics in a certain school. 56 of the students are boys and the remainder are girls. The probability that a student takes principal mathematics given that the student is a boy is 1/5. The probability that a student takes principal mathematics given that the student is a girl is 1/11. If a student is chosen at random from the school, find the probability that the student:
(i) is a boy given that the student takes principal mathematics,
(ii) does **not** take principal mathematics
**(2010 No. 12)**

**Answers:** (a) (i) 0.2 (ii) 0.6 (iii) 0.15
(b) (i) 0.7368 (ii) 0.848

**31**. Two events *A* and *B* are such that $P(A' \cap B) = 3x$, $P(A \cap B') = 2x$, $P(A' \cap B') = x$, and $P(B) = \frac{4}{7}$.
Using a Venn diagram, find the values of
(a) $x$,
(b) $P(A \cap B)$ **(2011 No. 4)**

**Answers:** (a) $\frac{1}{7}$ (b) $\frac{1}{7}$

**32.** Box *A* contains 4 red sweets and 3 green sweets. Box *B* contains 5 red sweets and 6 green sweets. Box *A* is twice as likely to be picked as box *B*. If a box is chosen at random and two sweets are removed from it, one at a time without replacement;
(a) find the probability that the two sweets removed are of the same colour.
(b) (i) construct a probability distribution table for the number of red sweets removed.
(ii) find the mean number of red sweets removed.
**(2011 No. 15)**

**Answers:** (a) 0.4372
(b) (i)

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X = x)$ | 0.1861 | 0.5628 | 0.2511 |

(ii) 1

**33.** Two events A and B are such that $P(A) = \frac{1}{5}$ and $P(B) = \frac{1}{2}$. Find $P(A \cup B)$ when A and B are:
(a) independent events,
(b) Mutually exclusive events. **(2012 No. 2)**

**Answers:** (a) 0.6 (b) 0.7

34. A box of oranges contains 20 good and 4 bad oranges. If 5 oranges are picked at random, determine the probability that 4 are good and the other is bad. **(2012 No .15)**

**Answer**: 0.456

**35.** Events A and B are such that
$P(A \cap B) = \frac{1}{2}$ and $P\left(A/B\right) = \frac{1}{3}$.
Find $P(B \cap A')$. **(2013 No. 4)**

**Answers:** $\frac{1}{6}$

**36.** (a) A bag contains 30 white (W), 20 blue (B) and 20 red (R) balls. Three balls are drawn at random one after the other without replacement. Determine the probability that the first ball is white and the third ball is also white.
(b) Events A and B are such that $P(A) = \frac{4}{7}$, $p(A \cap B') = \frac{1}{3}$ and $P(A/B) = \frac{5}{14}$.
Find (i) P(B) (ii) $P(A' \cap B')$. **(2014 No. 9)**

**Answers**: (a) 0.1801 (b) (i) $\frac{2}{3}$ (ii) 0

**37.** Events *A* and *B* are independent. $P(A) = x$, $P(B) = x + 0.2$ and $P(A \cup B) = 0.65$.
Find the value of $x$. **(2015 No.5)**

**Answer: 0.3**

16. A box *A* contains 4 white and 2 red balls. Another box *B* contains 3 white and 3 red balls. A box is selected at random and two balls are picked one after the other without replacement.
(a) Find the probability that the two balls picked are red.

(b) Given that two white balls are picked, what is the probability that they are from box B?

**(2015 No.16)**

**Answer:** (a) $\frac{2}{15}$ (b) 0.3 (c) $\frac{1}{3}$

## EXERCISE 3

1. (a) The letters of the word PROBABILITY are arranged at random. Find the probability that the two I'$_s$ is are separated.

(b) If the letter in the word ABSTEMIOUS are arranged at random, find the probability that the vowels and the consonants appear alternately

**2**. (a) From a group of 6 men and 8 women, 5 women are chosen at random. Find the probability that there are more men chosen than women

b) From a bag containing 6 white counters and 8 blue counters, 4 counters are chosen at random. Find the probability that 2 white counters and 2 blue counters are chosen

**3**. Two events C and D are such that P(C) = 0.7, $P(D \cup C) = 0.9$, $P(C \cap D) = 0.3$. Find:

(a) P(D) (b) $P(D' \cap C)$ (c) $P(D \cap C')$

(d) $P(D' \cap C')$

**4**. In a large group of people, it is known that 10% have a hot breakfast, 20% have lunch and 25% have a hot breakfast or a hot lunch. Find the probability that a person chosen at random from this group

(a) has a hot breakfast and a hot lunch

(b) has a hot lunch, given that the person chosen had a hot breakfast

**5.** Two children Akello (A) and Baale (B) play a game. An ordinary die is thrown and the first person to throw a four wins. Akello and Baale take it in turns to throw the die starting with Akello. Find the probability that Baale wins.

**6.** A stone is fired at the target and the probability that the target is hit is 0.7

a) Find how many stones should be thrown so that the probability that the target is hit at least once is greater than 0.995

b) Find how many stones should be thrown so that the probability that the target is not hit is less than 0.001

**7**. Each customer at a supermarket pays by one of the following; Cash, Cheque or Credit card. The probability of a randomly selected customer paying by cash is 0.54 and by cheque is 0.18.

a) Determine the probability of a randomly selected customer paying by credit card.

b) If three customers are selected at random, find the probability of:

(i) all three paying by cash

(ii) exactly one paying by cash

(iii) One paying by cash, one by cheque and one by credit card

**8.** A box contains 20 chocolates of which 15 have soft centres and 5 have hard centres. Two chocolates are taken at random, one after the other. Calculate the probability that

a) both chocolates have soft centres
b) one of each sort of chocolates is taken
c) both chocolates have hard centres given that the 2$^{nd}$ chocolate has a hard centre

**9.** In a group of 100 people, 40 own a cat, 25 own a dog and 15 own a cat and a dog. Find the probability that a person chosen at random:

a) Owns a dog or a cat
b) Owns a dog or a cat but not both
c) Owns a dog given that he owns a cat
d) Does not own a cat, given that he owns a dog

**10**. A bag contains seven black and 3 white marbles. Three marbles are chosen at random and in succession, each marble being replaced after it has been taken out of the bag. Calculate the probability of choosing:

**a)** three black marbles

**b)** a white marble, a black marble and a black marble in that order

**c)** two white marbles and black marble in that order

**d)** at least one black marble.

**11.** Otim's chances of passing physics are 0.60, of chemistry 0.75 and of mathematics 0.80.

i) Determine the chance that he passes only one subject

ii) If it is known that he passed at least two subjects, what is the probability that he failed chemistry?

**12.** I travel to work by route A or route B. The probability that I use route A is $\frac{1}{4}$. Probability that I am late on work if I go via route A is $\frac{2}{3}$ and the corresponding probability if I go via route B is $\frac{1}{3}$.

(a) What is the probability that I am late for work on Monday?

(b) Given that I am late for work, what is the probability that I went via route B?

**13.** In a group of 12 international referees there are three from Africa, four from Asia and five from Europe. To officiate at a tournament, three referees are chosen at random from the group. Calculate the probability that

(a) a referee is chosen from each container,

(b) exactly two referees are chosen from Asia,

(c) the three referees are chosen from the same continent.

# Chapter Three

# *PROBABILITY DENSITY FUNCTIONS (P.D.F)*

A function is said to be a P.d.f if its random variable, say *X* takes on either specific values or values within a given range.

## Types of pdf

There are two types of P.d.fs

- Discrete P.d.f
- Continuous P.d.f

## DISCRETE *PROBABILITY DENSITY FUNCTIONS*

If a random variable takes on specific values of a P.d.f, then such a variable is known as a discrete random variable

**Properties of a discrete P.d.f**

i) $\sum_{all\ X} f(x) = \sum_{all\ X} P(X = x) = 1$

ii) $P(X = x) \geq 0$

***Note***: We use the first property to find the unknown constants

***Example***

1. A random variable of discrete p.d.f is defined as

$$f(x) = \begin{cases} kx, & x = 1, 2, 3, 4 \\ 0 & \text{otherwise/Elsewhere} \end{cases}$$

Find the value of *k*

***Solution***

This means $f(x) = kx$, for only $X = 1, 2, 3, 4$

Otherwise $f(x) = 0$, for example $f(5) = f(6) = 0$

Now $\sum_{X=1}^{X=4} f(x) = 1$

$\Rightarrow f(1) + f(2) + f(3) + f(4) = 1$

$k + 2k + 3k + 4k = 1$

$10k = 1$

$k = \frac{1}{10}$

2. A random variable *X* of a discrete p.d.f is defined as: $P(X = 0) = 0.1$, $P(X = 1) = P(X = 2) = a$, $P(X = 3) = 0.2$ and $P(X = 4) = 0.3$

Find the value of *a*

***Solution***

$\sum_{X=0}^{X=4} P(X = x) = 1$

$\Rightarrow P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 1$

$0.1 + a + a + 0.2 + 0.3 = 1$

$2a + 0.6 = 1$

$2a = 0.4$

$a = 0.2$

## Finding Probabilities

This involves substituting directly the values of the random variable into the function for a given domain

Note: If *a* is a value in the domain, then:

i) $P(X = a) = f(a)$
ii) $P(X \geq a) = f(a) + . .... + f(\text{upper limit})$
iii) $P(X > a) = f(a + 1) + .... + f(\text{upper limit})$

**Examples**

1. A random variable *X* of a discrete p.d.f is given as

$$f(x) = \begin{cases} kx, & x = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases}$$

Find:

i) the value of *K*
ii) $P(X = 2)$
iii) $P(X > 2)$
iv) $P(X \leq 3)$
v) $P(1 < X \leq 4$
vi) $P(X > 1/X \leq 4)$

***Solution***

(i) $\sum_{X=1}^{X=5} f(x) = 1$

$\Rightarrow f(1) + f(2) + f(3) + f(4) + f(5) = 1$

$k + 2k + 3k + 4k + 5k = 1$

$15k = 1$

$k = \frac{1}{15}$

Hence $f(x) = \begin{cases} \frac{1}{15}x, & x = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases}$

(ii) $P(X = 2) = f(2) = \frac{2}{15}$

(iii) $P(X > 2) = P(X = 3, 4, 5)$

$= f(3) + f(4) + f(5)$

$= \frac{3}{15} + \frac{4}{15} + \frac{5}{15} = \frac{12}{15} = \frac{4}{5}$

(iv) $P(X \leq 3) = P(X = 1, 2, 3)$

$= f(1) + f(2) + f(3)$

$= \frac{1}{15} + \frac{2}{15} + \frac{3}{15} = \frac{6}{15} = \frac{2}{5}$

(v) $P(1 < X \leq 4) = P(X = 2, 3, 4)$

$= f(2) + f(3) + f(4)$

$= \frac{2}{15} + \frac{3}{15} + \frac{4}{15} = \frac{9}{15} = \frac{3}{5}$

(vi) $P(X > 1/X \leq 4) = \frac{P(X > 1 \cap X \leq 4)}{P(X \leq 4)}$

$= \frac{P(X = 2, 3, 4)}{P(X = 1, 2, 3, 4)}$

$= \frac{\frac{2}{15} + \frac{3}{15} + \frac{4}{15}}{\frac{1}{15} + \frac{2}{15} + \frac{3}{15} + \frac{4}{15}} = \frac{9}{15} \times \frac{15}{10} = \frac{9}{10}$

2. Two tetrahedral dice are each with faces labelled 1, 2, 3 and 4 are thrown and the sum of the scores on the upper shown faces noted. Find the probability distribution function of X, where X is a random variable representing the sum of the scores.

***Solution***

Probability space for possible outcomes

| | | first die | | | |
|---|---|---|---|---|---|
| second die | | 1 | 2 | 3 | 4 |
| | 1 | 2 | 3 | 4 | 5 |
| | 2 | 3 | 4 | 5 | 6 |
| | 3 | 4 | 5 | 6 | 7 |
| | 4 | 5 | 6 | 7 | 8 |

Probability distribution function

| $X$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| P($X = x$) | $\frac{1}{16}$ | $\frac{2}{16}$ | $\frac{3}{16}$ | $\frac{4}{16}$ | $\frac{3}{16}$ | $\frac{2}{16}$ | $\frac{1}{16}$ |

However, if the question required us to find the probability density function of $X$, then the following would be noted: The pattern for the probabilities relating to $X$ from 2 to 5:

$f(x) = P(X = x) = \dfrac{X-1}{16}$, for $X$ = 2, 3, 4, 5

For 2 to 6, the pattern changes to:

$f(x) = P(X = x) = \dfrac{9-x}{16}$, for $X$ = 6, 7, 8

Hence the p.d.f becomes;

$$f(x) = \begin{cases} \frac{x-1}{16}, & x = 2,3,4,5 \\ \frac{9-x}{16}, & x = 6,7,8 \\ 0 & \text{otherwise} \end{cases}$$

3. The probability density function of a random variable $Y$ is given by

P($Y = y$) = $cy^2$; for $y$ = 0, 1, 2, 3, 4

Find: (i) value of the constant, $c$

(ii) P($Y \geq 2$)

***Solution***

(i) P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) + P(Y = 4) = 1

$c + 4c + 9c + 16c = 1$

$30c = 1$

$c = \dfrac{1}{30}$

$P(Y \geq 2) = P(Y = 2) + P(Y = 3) + P(Y = 4)$

$= \dfrac{4}{30} + \dfrac{9}{30} + \dfrac{16}{30} = \dfrac{29}{30}$

## Cumulative Mass (Distribution) Function

The cumulative distribution function of f(x) denoted by F($X$) is defined as: F($x$) = $P(X \leq x)$

**Properties of *F*(*x*)**

i) F(+∞) = 1 where +∞ is the upper limit of the distribution

ii)The probabilities are non-decreasing

**Examples**

1. Given the following probability distribution function;

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| P($X = x$) | $\frac{1}{6}$ | $\frac{2}{6}$ | $\frac{2}{6}$ | $\frac{1}{6}$ |

Determine the cumulative mass function

***Solution***

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| P($X = x$) | $\frac{1}{6}$ | $\frac{2}{6}$ | $\frac{2}{6}$ | $\frac{1}{6}$ |
| F($X$) | $\frac{1}{6}$ | $\frac{3}{6}$ | $\frac{5}{6}$ | $\frac{6}{6}$ |

Note: $F(1) = P(X \leq 1) = \dfrac{1}{6}$

$F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = \dfrac{3}{6}$

$F(3) = P(X \leq 3) = \dfrac{5}{6}$

$F(4) = P(X \leq 4) = \dfrac{6}{6} = 1$

**2.** The discrete random variable $X$ has cumulative mass function $F(x) = \dfrac{x}{6}$, for $X$ = 1, 2, 3, …., 6

Find: (i) P(X ≤ 3)

(ii) the probability distribution of $X$

***Solution***

(i) $P(X \leq 3) = F(3) = \dfrac{3}{6} = \dfrac{1}{2}$

(ii) $P(X = 1) = F(1) = \dfrac{1}{6}$

$P(X = 2) = F(2) - F(1)$

$= \dfrac{2}{6} - \dfrac{1}{6} = \dfrac{1}{6}$

$P(X = 3) = F(3) - F(2) = \dfrac{3}{6} - \dfrac{2}{6} = \dfrac{1}{6}$

$P(X = 4) = F(4) - F(3) = \dfrac{4}{6} - \dfrac{3}{6} = \dfrac{1}{6}$

$P(X = 5) = F(5) - F(4) = \dfrac{5}{6} - \dfrac{4}{6} = \dfrac{1}{6}$

$P(X = 6) = F(6) - F(5) = \dfrac{6}{6} - \dfrac{5}{6} = \dfrac{1}{6}$

Hence the distribution of $X$ is:

| $X$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| P($X = x$) | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

## Finding Parameters

### a) Mean or Expected Value

The expected value of $X$ denoted by E($X$) *or* $\mu$ is defined as: $E(X) = \sum_{all\ x} xP(X = x)$

### b) Variance

The variance of $X$ denoted by var($X$) is defined as:

$$\text{Var}(X) = E(X^2) - [E(X)]^2$$

Where $E(X^2) = \sum_{all\ X} x^2 P(X = x)$

***Properties of the mean***

(i) E($a$) = $a$; where $a$ is a constant

(ii) E($aX$) = $a$E($X$)
(iii) E($ax + b$) = E($ax$) + E($b$)
$= aE(X) + b$

Properties of variance
(i) Var($a$) = 0
(ii) Var($aX$) = $a^2$var($X$)
(iii) Var($aX + b$) = var($aX$) + var($b$)
$= a^2\text{var}(X) + 0$
$= a^2\text{var}(X)$

## Mode

This is the value of the function associated with the highest probability

## Median

This is the smallest value of $X$ for which $F(X) \geq 0.5$

If $M$ is the median, then M must satisfy the following conditions
(i) $F(M) \geq 0.5$
(ii) $1 - F(M - 1) \geq 0.5$

***Note***: If there exists $X_1$, such that $F(X_1) = 0.5$, then there will be two values of $X$ i.e. $X_1$ and $X_2$ will satisfy the two conditions and hence median = $\frac{X_1 + X_2}{2}$

## Quartiles

The lower quartile is the value of $X$ which satisfies the following conditions:
(i) $F(q_1) \geq 0.25$
(ii) $1 - F(q_1 - 1) \geq 0.75$

The upper quartile is the value of $X$ which satisfies the following conditions
(i) $F(q_3) \geq 0.75$
(ii) $1 - F(q_3 - 1) \geq 0.25$

Interquartile range = Upper quartile – lower quartile
$= q_3 - q_1$

***Examples***

1. Given the probability distribution of $X$ below:

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| P($X = x$) | 0.1 | 0.4 | 0.3 | 0.2 |

Find: (i) E($X$)
(ii) var($X$)
(iii) mode
(iv) median
(v) interquartile range

***Solution***

| $X$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| P($X = x$) | 0.1 | 0.4 | 0.3 | 0.2 |
| $xP(X = x)$ | 0.1 | 0.8 | 0.9 | 0.8 |
| $x^2P(X = x)$ | 0.1 | 1.6 | 2.7 | 3.2 |
| $F(X)$ | 0.1 | 0.5 | 0.8 | 1.0 |

(i) $E(X) = \sum x\text{P}(X = x)$
$= 0.1 + 0.8 + 0.9 + 0.8 = 2.6$

(ii) Var($X$) = E($X^2$) – [E($X$)]$^2$
E($X^2$) = 0.1 + 1.6 + 2.7 + 3.2 = 7.6

Var($X$) = 7.6 – (2.6)$^2$
= 7.6 – 6.76= 0.84

(iii) Mode = 2

(iv) Taking median = 2
F(2) = 0.5
1 – F(1) = 1 – 0.1
= 0.9 > 0.5

Taking median = 3;
$F$(3) = 0.8 > 0.5
1 – $F$(2) = 1 – 0.5= 0.5

Hence median = $\frac{2+3}{2}$ = 2.5

(v) Taking lower quartile = 2;
$F$(2) = 0.5
1 – F(1) = 1 – 0 = 0.9 > 0.75

Hence $q_1 = 2$

Taking upper quartile = 3;
$F$(3) = 0.8 > 0.75
1 – $F$(2) = 1 – 0.5
= 0.5 > 0.25

Hence $q_3 = 3$

Interquartile range = 3 – 2 = 1

2. Given the following probability distribution;

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| P($X = x$) | 0.1 | 0.2 | 0.3 | 0.2 | 0.2 |

Find (i) E($X$) and E(6$X$ + 2)
(ii) Var($X$) and Var(6$X$ + 2)
(iii) mode
(iv) median
(v) semi-interquartile range

***Solution***

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| P($X = x$) | 0.1 | 0.2 | 0.3 | 0.2 | 0.2 |
| $x$P($X = x$) | 0 | 0.2 | 0.6 | 0.6 | 0.8 |
| $x^2$P($X = x$) | 0 | 0.2 | 1.2 | 1.8 | 3.2 |
| F($X$) | 0.1 | 0.3 | 0.6 | 0.8 | 1.0 |

(i) E($X$) = $\sum x\text{P}(X = x)$
= 0.2 + 0.6 + 0.6 + 0.8 = 2.2

E(6$X$ + 2) = E(6$X$) + E(2)
= 6E($X$) + 2
= 6 × 2.2 + 2 = 15.2

(ii) Var($X$) = E($X^2$) – [E($X$)]$^2$
E($X^2$) = 0.2 + 1.2 + 1.8 + 3.2 = 6.4
Var($X$) = 6.4 – (2.2)$^2$ = 1.56
Var(6$X$ + 2) = Var(6$X$) + Var(2)
= 36var($X$) + 0
= 36 ×1.56 = 56.16

(iii) Mode = 2

(iv) Taking median = 2
F(2) = 0.6 > 0.5
1 – F(1) = 1 – 0.3 = 0.7 > 0.5

Taking median = 3;

F(3) = 0.8 > 0.5
1 – F(2) = 1 – 0.6 = 0.4 < 0.5
Hence median = 2
(Since it satisfies the two conditions above)

(v) Taking lower = quartile, $q_1$ = 1;
F(1) = 0.3 > 0.25

$1 - F(0) = 1 - 0.1$
$= 0.9 > 0.75$

Hence lower quartile = 1

Taking upper quartile, $q_3 = 3$;
$F(3) = 0.8 > 0.75$;
$1 - F(2) = 1 - 0.6 = 0.4 > 0.25$

Hence upper quartile = 3

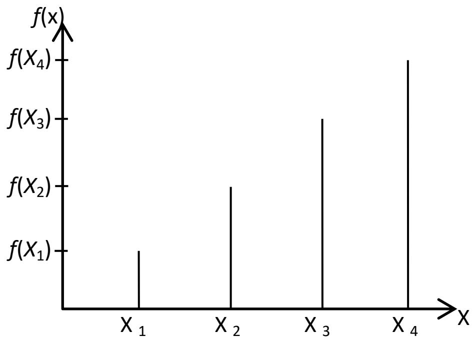
Interquartile range = 3 – 1 = 2

Semi-interquartile range = $\frac{2}{2} = 1$

**<u>Graphs of *f*(*x*) and *F*(x)</u>**

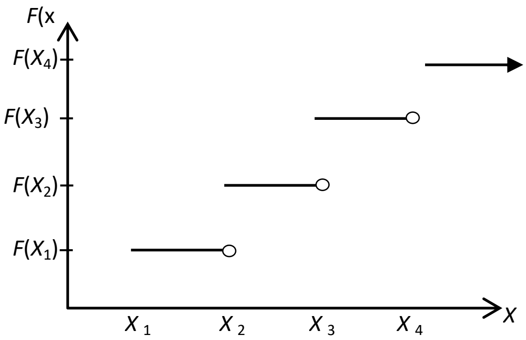
**Graphs of *f*(*x*)**

This comprises of vertical lines drawn from specific values of *x* corresponding to their respective probabilities




**Graph of *F*(*x*)**

This comprises of horizontal lines drawn from specific values of *X* towards the next values corresponding to their respective cumulated probabilities




***Examples***

1. Given probability distribution function of *X*;
$P(X = 1) = 0.1$, $P(X = 2) = P(X = 3) = 0.2$, $P(X = 4) = 0.3$ and $P(X = 5) = 0.2$.

a) Show that the distribution given is a discrete .d.f

b) Sketch the graph of *f*(*X*) and F(*X*)

***Solution***

a) Conditions for a discrete p.d.f:

(i) $\sum P(X = x) = 1$

Now $P(X = 1)+P(X = 2)+P(X = 3)+P(X = 4)+P(X = 5)$

$= 0.1 + 0.2 + 0.2 + 0.3 + 0.2 = 1.0$

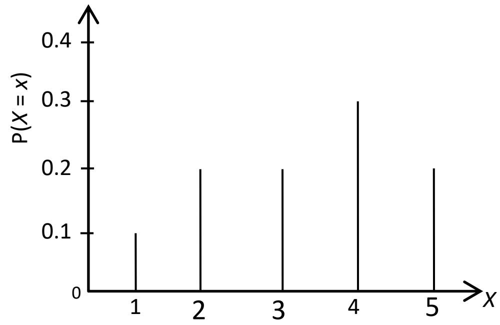
(ii) $P(X = x) \geq 0$ for all values of *X*,
Now for $x = 1, 2, 3, 4, 5$, $P(X = x) > 0$

Hence the distribution is a discrete p.d.f

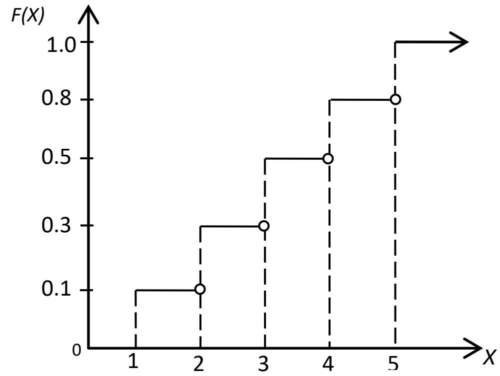
b)

| *x* | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P(*X*=*x*) | 0.1 | 0.2 | 0.2 | 0.3 | 0.2 |
| F(*x*) | 0.1 | 0.3 | 0.5 | 0.8 | 1.0 |

**Graph of *f*(*x*)**




Graph of F(*x*)



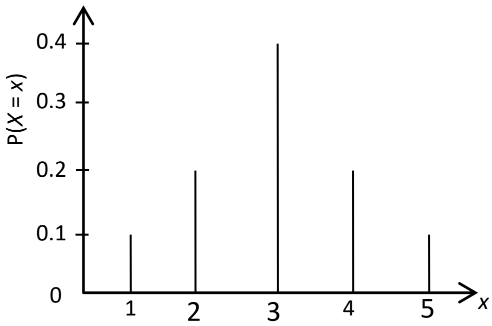

***Note***: If a graph of a discrete p.d.f is symmetrical about a specific value of *X*, then the expectation of *X*, *E*(*X*) is equal to the value of *X*

**2**. A random variable *X* of a probability distribution function is given as:

| *x* | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P(*X* = *x*) | 0.1 | 0.2 | 0.4 | 0.2 | 0.1 |

Draw the graph of the distribution and hence find E(*X*)

***Solution***




The graph is symmetrical about $x = 3$,
Hence $E(X) = 3$.

**GENERAL EXAMPLES**

**1.** A news agent stocks 12 copies of a magazine each week. He has regular orders for nine copies, and the number of additional copies sold varies from week to week. The news agent uses previous sales data to estimate the probability for each possible total number of copies solid, as follows:

| Number of copies | 9 | 10 | 11 | 12 |
|---|---|---|---|---|
| Probability | 0.20 | 0.35 | 0.30 | 0.15 |

a) Calculate the expected number of copies that he sells in a week

b) The news agent buys the magazines at 1500/- each and sells them at 2000/- each. Any copies left unsold are destroyed.

i) Find the profit on these magazines in a week when he sells 10 copies

ii) Construct a probability distribution table for the news agent's weekly profit from the sale of these magazines. Hence or otherwise, calculate the expected weekly profit

***Solution***

a) Let $X$ = number of copies the news agent sells per week

| $x$ | 9 | 10 | 11 | 12 |
|---|---|---|---|---|
| $P(X = x)$ | 0.20 | 0.35 | 0.30 | 0.15 |
| $xP(X = x)$ | 1.80 | 3.50 | 3.30 | 1.80 |

$E(X) = \sum xP(X = x)$

$= 1.80 + 3.50 + 3.30 + 1.80 = 10.4$

Hence the expected number of copies sold per week is 10.4

b) (i) Profits for 10 copies = total sales – cost of sales

$= 10 \times 2000 - 12 \times 1500$

$= 20{,}000 - 18{,}000$

$= 2000/-$

Hence the profit on 10 copies is 5000/-

(ii) Profits for 9 copies $= 9 \times 2000 - 12 \times 1500$

$= 18000 - 18000 = 0/-$

Profits for 11 copies $= 11 \times 2000 - 12 \times 1500$

$= 22{,}000 - 18{,}000$

$= 4000/-$

Profits for 12 copies $= 12 \times 2000 - 12 \times 1500$

$= 24{,}000 - 18{,}000$

$= 6000$

Let $Y$ = weekly profit,

Then the probability distribution of $Y$ will be:

| $y$ | 0 | 2000 | 4000 | 6000 |
|---|---|---|---|---|
| $P(Y = y)$ | 0.20 | 0.35 | 0.30 | 0.15 |
| $y(Y = y)$ | 0 | 700 | 1200 | 900 |

$E(Y) = 0 + 700 + 1200 + 900$

$= 2{,}800/-$

Hence the expected weekly profit is 2800/-

**2.** The discrete random variable $X$ can take on values 0, 1, 2 and 3 only. Given $P(X \le 2) = 0.9$, $P(X \le 1) = 0.5$ and $E(X) = 1.4$, find:

a) $P(X = 0)$ b) $P(X = 1)$

***Solution***

(a) ***Note***: $P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2)$

$\Rightarrow 0.9 = P(X = 0) + P(X = 1) + P(X = 2)$.............. (i)

Also $P(X \le 1) = P(X = 0) + P(X = 1)$

$\Rightarrow 0.5 = P(X = 0) + P(X = 1)$................................ (ii)

Substituting Eqn (ii) into Eqn (i);

$0.9 = 0.5 + P(X = 2)$

$P(X = 2) = 0.4$

But $\sum_{x=0}^{x=3} P(X = x) = 1$

$\Rightarrow P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1$

$0.9 + P(X = 3) = 1$

$P(X = 3) = 0.1$

Let $P(X = 0) = a$, then from Eqn (i);

$P(X = 1) = 0.5 - a$

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $a$ | $0.5 - a$ | 0.4 | 0.1 |
| $xP(X = x)$ | 0 | $0.5 - a$ | 0.8 | 0.3 |

$E(X) = \sum xP(X = x)$

$1.4 = 0.5 - a + 0.8 + 0.3$

$1.4 = 1.6 - a$

$a = 0.2$

Hence $P(X = 0) = 0.2$

(b) $P(X = 1) = 0.5 - 0.2 = 0.3$

3. The following table shows the probability distribution for a random variable $X$

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $c$ | $c^2$ | $c^2 + c$ | $3c^2 + 2c$ |

Calculate (a) the value of $c$

(b) $E(X)$

***Solution***

a) $\sum_{all\ X} P(X = x) = 1$

$\Rightarrow c + c^2 + c^2 + c + 3c^2 + 2c = 1$

$5c^2 + 4c - 1 = 0$

$5c^2 + 5c - c - 1 = 0$

$5c(c + 1) - 1(c + 1) = 0$

$(5c - 1)(c + 1) = 0$

$5c - 1 = 0$

$c = \frac{1}{5} = 0.2$

| $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{5}$ | $\frac{1}{25}$ | $\frac{1}{25} + \frac{1}{5} = \frac{6}{25}$ | $\frac{3}{25} + \frac{2}{5} = \frac{13}{25}$ |
| $xP(X = x)$ | 0 | $\frac{1}{25}$ | $\frac{12}{25}$ | $\frac{39}{25}$ |

$E(X) = \sum xP(X = x)$

$= \frac{1}{25} + \frac{12}{25} + \frac{39}{25} = \frac{52}{25} = 2.08$

**4**. The discrete random variable $X$ has a probability distribution function $P(X = x) = k|x|$, where $x$ takes the values -3, -2, -1, 0, 1, 2, 3. Find the:

a) value of the constant $k$

b) $E(X)$

c) standard deviation of $X$

***Solution***

| $x$ | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|
| $P(X=x)$ | $3k$ | $2k$ | $k$ | 0 | $k$ | $2k$ | $3k$ |

$\sum_{all\ X} P(X = x) = 1$

$3k + 2k + k + k + 2k + 3k = 1$

$12k = 1$

$k = \frac{1}{12}$

b) $E(X)$

| $x$ | -3 | -2 | -1 | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|
| $P(X=x)$ | $\frac{3}{12}$ | $\frac{2}{12}$ | $\frac{1}{12}$ | 0 | $\frac{1}{12}$ | $\frac{2}{12}$ | $\frac{3}{12}$ |
| $xP(X=x)$ | $\frac{-9}{12}$ | $\frac{-4}{12}$ | $\frac{-1}{12}$ | 0 | $\frac{1}{12}$ | $\frac{4}{12}$ | $\frac{9}{12}$ |
| $x^2P(X=x)$ | $\frac{27}{12}$ | $\frac{8}{12}$ | $\frac{1}{12}$ | 0 | $\frac{1}{12}$ | $\frac{8}{12}$ | $\frac{27}{12}$ |

$E(X) = \sum xP(X = x)$

$$= \frac{-9}{12} + \frac{-4}{12} + \frac{-1}{12} + \frac{1}{12} + \frac{4}{12} + \frac{9}{12} = 0$$

c) Standard deviation = $\sqrt{\text{var}(X)}$

$$= \sqrt{E(X^2) - [E(X)]^2}$$

$E(X^2) = \sum x^2P(X = x)$

$$= \frac{27}{12} + \frac{8}{12} + \frac{1}{12} + \frac{8}{12} + \frac{27}{12} = \frac{72}{12} = 6$$

Standard deviation = $\sqrt{6-0}$

$$= \sqrt{6} = 2.449$$

**5.** A random variable $X$ of a pdf is given by:

$$f(x) = \begin{cases} kx, & x = 1, 2, 3, \ldots n \\ 0, & \text{else where} \end{cases}$$

Given that $E(X) = 3$,

a) find the value of constants $K$ and $n$

b) sketch the graph of $f(X)$ and $F(X)$

***Solution***

a) $\sum_{all\ X} P(X = x) = 1$

$\Rightarrow f(1) + f(2) + f(3) + \ldots\ldots + f(n) = 1$

$k[1 + 2 + 3 + \ldots + n] = 1$

$k\left[\frac{n}{2}(n+1)\right] = 1$

$kn(n + 1) = 2$ ..................................... (i)

$E(X) = \sum_{all\ X} xf(x)$

$= k\sum_{all\ X} x^2$

$= k(\mathbf{1^2 + 2^2 + 3^2 + \ldots\ldots + n^2})$

**But E($X$) =** 3

$\Rightarrow k(1^2 + 2^2 + 3^2 + \ldots\ldots + n^2) = 3$

$k\left[\frac{n}{6}(n+1)(2n+1)\right] = 3$

$Kn(n + 1)(2n + 1) = 18$ ...................... (ii)

Eqn (ii) ÷ Eqn (i);

$$\frac{kn(n+1)(2n+1)}{kn(n+1)} = \frac{18}{2}$$

$2n + 1 = 9$

$n = 4$

Substituting for $n$ into Eqn (i);

$4k(4 + 1) = 2$

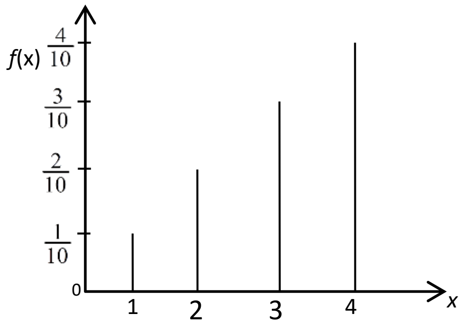
$10k = 1$

$$k = \frac{1}{10}$$

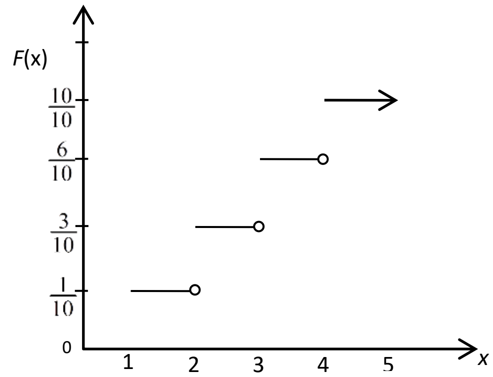
b)

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $f(X)$ | $\frac{1}{10}$ | $\frac{2}{10}$ | $\frac{3}{10}$ | $\frac{4}{10}$ |
| $F(X)$ | $\frac{1}{10}$ | $\frac{3}{10}$ | $\frac{6}{10}$ | $\frac{10}{10}$ |

Graph of $f(x)$




Graph of $F(x)$




**6**. A random variable X takes on integral values within the given range of the function:

$$f(x) = \begin{cases} a2^x, & 0 \le x \le 6 \\ 0, & \textit{else where} \end{cases}$$

Find: (i) the value of $a$ (ii) $E(X)$

(iii) $P(X < 4/X \ge 2)$

***Solution***

***Note***: When $X$ takes on integral values, this means that the distribution is a discrete one.

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $f(x) = a2^x$ | $a$ | $2a$ | $4a$ | $8a$ | $16a$ | $32a$ | $64a$ |

$\sum_{all\ X} f(x) = 1$

$\Rightarrow a + 2a + 4a + 8a + 16a + 32a + 64a = 1$

$127a = 1$

$$a = \frac{1}{127}$$

ii) $E(X) = \sum_{all\ X} xf(x)$

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $f(x)$ | $\frac{1}{127}$ | $\frac{2}{127}$ | $\frac{4}{127}$ | $\frac{8}{127}$ | $\frac{16}{127}$ | $\frac{32}{127}$ | $\frac{64}{127}$ |
| $xf(x)$ | 0 | $\frac{2}{127}$ | $\frac{8}{127}$ | $\frac{24}{127}$ | $\frac{64}{127}$ | $\frac{160}{127}$ | $\frac{384}{127}$ |

$$E(X) = \frac{2}{127} + \frac{8}{127} + \frac{24}{127} + \frac{64}{127} + \frac{160}{127} + \frac{384}{127}$$

$$= \frac{642}{127} = 5.055$$

(iii) $P(X < 4/X \ge 2) = \dfrac{P(X < 4 \cap X \ge 2)}{P(X \ge 2)}$

$$= \frac{P(X=2,3)}{P(X=2,3,4,5,6)}$$

$$= \frac{\frac{4}{127}+\frac{8}{127}}{\frac{4}{127}+\frac{8}{127}+\frac{16}{127}+\frac{32}{127}+\frac{64}{127}}$$

$$= \frac{12}{127}\times\frac{127}{124} = \frac{12}{124} = \frac{3}{31}$$

**7.** A bag contains 6 red sweets and 4 yellow sweets. Two sweets are to be picked at random in succession from the bag without replacement.

(a) (i) Write down the possible sets of colours of sweets that will be picked from the bag

(ii) Determine the probability of picking each set of the sweets in (i) above.

(b) If $X$ denotes the number of red sweets picked in (i) above, construct a probability distribution function of $X$.

(c) (i) Determine the cumulative distribution function $F(x)$, of $X$

(ii) Plot the graph of $F(x)$.

**Solution**

(a) (i)

Red = 6

Yellow = 4

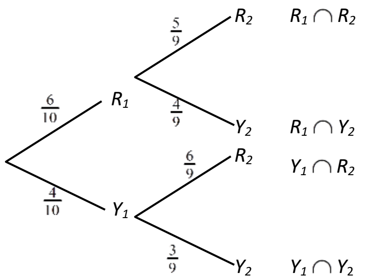
Total = 10

Let $R_1$ = first sweet drawn is red

$R_2$ = second sweet drawn is red

$W_1$ = first sweet drawn is yellow

$W_2$ = second sweet drawn is yellow




$$S = \{R_1R_2, R_1Y_2, Y_1R_2, Y_1Y_2\}$$

(ii) $P(R_1R_2) = \frac{6}{10}\cdot\frac{5}{9} = \frac{30}{90} = \frac{5}{15}$

$P(R_1Y_2) = \frac{6}{10}\cdot\frac{4}{9} = \frac{24}{90} = \frac{4}{15}$

$P(Y_1R_2) = \frac{4}{10}\cdot\frac{6}{9} = \frac{24}{90} = \frac{4}{15}$

$P(Y_1Y_2) = \frac{4}{10}\cdot\frac{3}{9} = \frac{12}{90} = \frac{2}{15}$

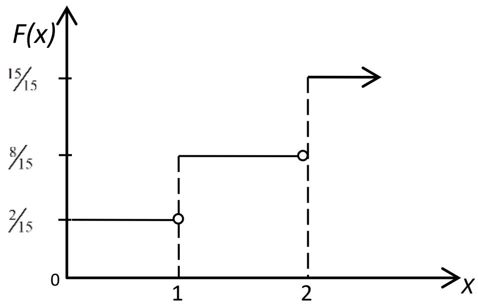
(b)

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| P($X$ = $x$) | $\frac{2}{5}$ | $\frac{8}{15}$ | $\frac{5}{15}$ |

(c) (i)

| $x$ | 0 | 1 | 2 |
|---|---|---|---|
| P($X$ = $x$) | $\frac{2}{5}$ | $\frac{8}{15}$ | $\frac{5}{15}$ |
| $F(x)$= P($X \le x$) | $\frac{2}{5}$ | $\frac{10}{15}$ | $\frac{15}{15}$ |

(ii)




7. The random variable $X$ has the distribution shown in the table below

| $x$ | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|
| P($X$=$x$) | $\frac{1}{20}$ | $\frac{1}{20}$ | $m$ | $\frac{3}{20}$ | $\frac{6}{20}$ | $\frac{2}{20}$ | $\frac{3}{20}$ | $\frac{1}{20}$ |

Find the:,

(i) value of $m$

(ii) probability distribution of $Y$, where $Y = 2X - 3$,

(iii) expectation of $Y$ .

**Solution**

(i) $\sum_{all x} P(X=x) = 1$

$\frac{1}{20}+\frac{1}{20}+m+\frac{3}{20}+\frac{6}{20}+\frac{2}{20}+\frac{3}{20}+\frac{1}{20}=1$

$\frac{17}{20}+m=1$

$m = \frac{20}{20}-\frac{17}{20}$

$m = \frac{3}{20}$

(ii) $Y = 2x - 3$

| $y$ | -5 | -3 | -1 | 1 | 3 | 5 | 7 | 9 |
|---|---|---|---|---|---|---|---|---|
| P($Y$= $y$) | $\frac{1}{20}$ | $\frac{1}{20}$ | $\frac{3}{20}$ | $\frac{2}{20}$ | $\frac{6}{20}$ | $\frac{2}{20}$ | $\frac{3}{20}$ | $\frac{1}{20}$ |

(iii)

| $y$ | -5 | -3 | -1 | 1 | 3 | 5 | 7 | 9 |
|---|---|---|---|---|---|---|---|---|
| P($Y$= $y$) | $\frac{1}{20}$ | $\frac{1}{20}$ | $\frac{3}{20}$ | $\frac{2}{20}$ | $\frac{6}{20}$ | $\frac{2}{20}$ | $\frac{3}{20}$ | $\frac{1}{20}$ |
| $y$P($Y$= $y$) | $-\frac{5}{20}$ | $-\frac{3}{20}$ | $-\frac{3}{20}$ | $\frac{3}{20}$ | $\frac{18}{20}$ | $\frac{10}{20}$ | $\frac{21}{20}$ | $\frac{9}{20}$ |

$$E(Y) = \sum yP(Y=y)$$

$$= -\tfrac{5}{20} + -\tfrac{3}{20} + -\tfrac{3}{20} + \tfrac{3}{20} + \tfrac{18}{20} + \tfrac{10}{20} + \tfrac{21}{20} + \tfrac{9}{20}$$

$$= -\tfrac{11}{20} + \tfrac{61}{20}$$

$$= \tfrac{50}{20} = 2.5$$

**OR**

$$E(Y) = E(2X-3)$$

$$= E(2X) - E(3)$$

$$= 2E(X) - 3$$

| $x$ | -1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|
| P($X$= $x$) | $\frac{1}{20}$ | $\frac{1}{20}$ | $\frac{3}{20}$ | $\frac{2}{20}$ | $\frac{6}{20}$ | $\frac{2}{20}$ | $\frac{3}{20}$ | $\frac{1}{20}$ |
| $x$P($X$= $x$) | $-\frac{1}{20}$ | 0 | $\frac{3}{20}$ | $\frac{6}{20}$ | $\frac{18}{20}$ | $\frac{8}{20}$ | $\frac{15}{20}$ | $\frac{6}{20}$ |

$$E(X) = -\tfrac{1}{20} + \tfrac{3}{20} + \tfrac{6}{20} + \tfrac{18}{20} + \tfrac{8}{20} + \tfrac{15}{20} + \tfrac{6}{20}$$

$$= \tfrac{55}{20}$$

$E(Y) = \frac{55}{20}\times 2 - 3$

$= \frac{110}{20} - 3$

$= 5.5 - 3$

$= 2.5$

8. The table below shows a random variable $X$ with the following probability distribution

| $x$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P($X$ = $x$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

(i) Construct tables for the $W$ and $Z$, such that $W = 3x$ and $Z = 2x + 4$.

(ii) Find the expectation of $W$ and $Z$.

(iii) Calculate the variance of $Z$.

***Solution***

(i) W = 3$x$

| $w$ | 3 | 6 | 9 | 12 | 15 |
|---|---|---|---|---|---|
| P($W$ = $w$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

Z = 2$x$ + 4

| $z$ | 6 | 8 | 10 | 12 | 14 |
|---|---|---|---|---|---|
| P($Z$ = $z$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

(ii)

| **$w$** | **3** | **6** | **9** | **12** | **15** |
|---|---|---|---|---|---|
| P($W$ = $w$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $w$P($W$ = $w$) | $\frac{3}{4}$ | $\frac{6}{8}$ | $\frac{9}{8}$ | $\frac{12}{4}$ | $\frac{15}{4}$ |
| **$z$** | **6** | **8** | **10** | **12** | **14** |
| P($Z$ = $z$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $z$P($Z$ = $z$) | $\frac{6}{4}$ | $\frac{8}{8}$ | $\frac{10}{8}$ | $\frac{12}{4}$ | $\frac{14}{4}$ |
| $z^2$P($Z$ = $z$) | $\frac{36}{4}$ | $\frac{64}{8}$ | $\frac{100}{8}$ | $\frac{144}{4}$ | $\frac{196}{4}$ |

$E(W) = 3/4 + 6/8 + 9/8 + 12/4 + 15/4$

$= 75/8 = 9.375$

$E(Z) = 6/4 + 8/8 + 10/8 + 12/4 + 14/4$

$= 82/8 = 10.25$

**Alternatively:**

| $x$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P($X$ = $x$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $x$P($X$ = $x$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $x^2$P($X$ = $x$) | $\frac{1}{4}$ | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

$E(X) = 1/4 + 2/8 + 3/8 + 4/4 + 5/4$

$= 25/8 = 3.125$

$E(W) = 3E(x)$

$= 3\times 3.125$

$= 9.375$

$E(Z) = E(2x+4)$

$= 2E(x) + E(4)$

$= 2E(x) + 4$

$= 2\times 3.125 + 4$

$= 10.25$

(iii) $Var(Z) = E(Z^2) - [E(Z)]^2$

$E(Z^2) = \frac{36}{4} + \frac{64}{8} + \frac{100}{8} + \frac{144}{8} + \frac{196}{8}$

$= \frac{916}{8} = 114.5$

$Var(Z) = 114.5 - (10.25)^2$

$= 9.4375$

**OR**

$Var(Z) = Var(2x+4)$

$= 4Var(x)$

$But\ Var(x) = E(x^2) - [E(x)]^2$

$E(x^2) = \frac{1}{4} + \frac{4}{8} + \frac{9}{8} + \frac{16}{4} + \frac{25}{4}$

$= \frac{97}{8} = 12.125$

$Var(x) = 12.125 - (3.125)^2$

$= 2.359375$

$\Rightarrow Var(Z) = 4\times 2.359375$

$= 9.4375$

# Exercise 3.1

1. The discrete random variable $X$ has the given probability distribution

| $x$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P($X$ = $x$) | 0.2 | 0.25 | 0.4 | $a$ | 0.05 |

a) Find the value of $a$

b) Find: (i) P($1 \le X \le 3$) (ii) P($X > 2$) (iii) P($2 < X < 5$) (iv) mode

2. A game consists of throwing tennis balls into a bucket from a given distance. The probability that Peter will get the tennis ball in the bucket is 0.4. A turn consists of three attempts.

a) Construct the probability distribution for $X$, the number of tennis balls that land in the bucket in a turn.

b) Peter wins a prize if at the end of his turn, there are two or more tennis balls in the bucket. What is the probability that Peter doesn't win a prize?

3. A bag contains five black counters and six red counters. Two counters are drawn one at a time and not replaced. By letting $X$ be the number of red counters drawn. Find $E(X)$.

4. The discrete random variable $X$ can take on values 0, 1, 2 and 3 only. Given that P(X ≤ 2) = 0.9, P($X$ ≤ 1) = 0.5 and E($X$) = 1.4. Find

a) P($X$ = 1)

b) P($X$ = 0)

5. A random variable $X$ of a p.d.f takes on only integral values whose function is given below

$$f\ \ x = \begin{cases} \frac{1}{2}^{\,x}, & 1 \le x \le 5 \\ c, & x = 6 \\ 0, & \text{otherwise} \end{cases}$$

where $c$ is a constant.

Determine the value of $c$ and hence the mode and mean of $X$

6. (a) The discrete random variable $X$ has *p.d.f* P($X$ = $x$) = $k|x|$, where $x$ takes the values -3, -2, -1, 0, 1, 2, 3.

Find: (i) the value of the constant $k$

(ii) $E(X)$

(iii) the standard deviation of $X$

(b) The discrete random variable $X$ has distribution function $F(X)$ where

$F(X) = 1 - (1 - \frac{1}{4}x)^x$ for $x = 1, 2, 3, 4$

(i) Show that $F(3) = \frac{63}{64}$ and $F(2) = \frac{3}{4}$

(ii) Obtain the probability distribution of $X$

(iii) Find E($X$) and Var($X$)

(iv) Find P($X$ > E($X$))

7. A discrete random variable $X$ has a probability function

$$P(X = x) = \begin{cases} x/k, & x = 1, 2, 3........, n \\ o & otherwise \end{cases}$$

Where $k$ and $n$ are real numbers.
Given that the expectation of $X$ is 3, find

(a) the values of $n$ and $k$

(b) the median and variance of $X$

(c) $P(X = 2 / X \geq 2)$

8. The discrete random variable X has p.d.f
$P(X = x) = k$ for $x$ =1, 2, 3, 4, 5, 6.
Find (a) E($X$), (b) E($X^2$),
(c) E($3X + 4$), (d) Var($X$).

9. The discrete random variable X has probability function given by

$$P(X = x) = \begin{cases} \left(\frac{1}{2}\right)x & x = 1, 2, 3, 4, 5, \\ c & x = 6, \\ o & otherwise \end{cases}$$

where $c$ is a constant.
Determine the value of $c$ and hence the mode and mean of $X$.

10. A game consists of tossing four unbiased coins simultaneously. The total score is calculated by giving three points for each head and one point for each tail. The random variable $X$ represents the total score.

(a) Show that $P(X = 8) = \frac{3}{8}$.

(b) Copy and complete the table, given below, for the symmetrical probability distribution of $X$.

| $x$ | 4 | 6 | 8 | 10 | 12 |
|---|---|---|---|---|---|
| P($X = x$) | | | $\frac{3}{8}$ | | |

(c) Calculate the variance of $X$.

# *CONTINUOUS PROBABILITY DENSITY FUNCTIONS*

A function is said to be a continuous probability density function if its domain is continuous. For example $f(x) = a \leq X \leq b$ where $X$ takes on any value within the given range

***Note***: For a continuous function,

$P(a \leq X \leq b) = P(a < X < b)$

***Properties of a continuous p.d.f***

(i) $f(x) \geq 0$ for all values of $X$

(ii) $\int_{x=a}^{x=b} f(x)dx = 1$ or $\int_{-\infty}^{+\infty} f(x)dx = 1$

Where $a$ = lower limit (+$\infty$) and
$b$ = upper limit of the domain (-$\infty$)

We use the second property to find the constants

***Examples***

1. A random variable $X$ of a continuous *p.d.f* is given by:

$$f(x) = \begin{cases} kx, & 0 \leq x \leq 6 \\ 0, & \text{else where} \end{cases}$$

Find the value of $k$

***Solution***

$$\int_0^6 kx\,dx = 1$$

$$k\int_0^6 x\,dx = 1$$

$$k\left[\frac{x^2}{2}\right]_0^6 = 1$$

$$k\left[\frac{36}{2} - 0\right] = 1$$

$$18k = 1$$

$$k = \frac{1}{18}$$

**2.** A random variable $X$ of a continuous *p.d.f* is given by

$$f(x) = \begin{cases} k(x+1), & 0 \leq x \leq 4 \\ kx, & 4 < x \leq 6 \\ 0, & else\ where \end{cases}$$

Find the value $k$

***Solution***

$$k\int_0^4 (x+1)\,dx + k\int_4^6 x\,dx = 1$$

$$k\left[\frac{x^2}{2} + x\right]_0^4 + k\left[\frac{x^2}{2}\right]_4^6 = 1$$

$$k\big[(8+4) - (0)\big] + k\big[18 - 8\big] = 1$$

$$12k + 10k = 1$$

$$22k = 1$$

$$k = \frac{1}{22}$$

## *Finding probabilities*

This involves integrating the function given for given limits of the probability

**Examples**

1. Given a continuous *p.d.f* as

$$f(x) = \begin{cases} \frac{1}{18}x, & 0 \le x \le 6 \\ 0, & \text{else where} \end{cases}$$

Find: (i) P($X > 2$)
(ii) P($X < 3$)
(iii) P($1 < X < 3$)
(iv) P($X > 2/X \le 4$)

***Solution***

(i) P($X > 2$) = $\frac{1}{18}\int_2^6 x\,dx$

$$= \frac{1}{8}\left[\frac{x^2}{2}\right]_2^6 = \frac{1}{36}(36-4) = \frac{8}{9}$$

(ii) P($X < 3$) = $\frac{1}{18}\int_0^3 x\,dx$

$$= \frac{1}{8}\left[\frac{x^2}{2}\right]_0^3 = \frac{1}{36}(9-0) = \frac{1}{4}$$

(iii) P($1 < X < 3$) $= \frac{1}{18}\int_1^3 x\,dx$

$$= \frac{1}{8}\left[\frac{x^2}{2}\right]_1^3 = \frac{1}{36}[9-1] = \frac{8}{36} = \frac{2}{9}$$

(iv) P(X > 2/X ≤ 4) = $\frac{P(X > 2 \cap X \le 4)}{P(X \le 4)}$

$$= \frac{\frac{1}{18}\int_2^4 x\,dx}{\frac{1}{18}\int_0^4 x\,dx} = \frac{\frac{1}{36}\left[x^2\right]_2^4}{\frac{1}{36}\left[x^2\right]_0^4} = \frac{16-4}{16-0}$$

$$= \frac{12}{16} = \frac{3}{4}$$

**2.** A continuous random variable has a *p.d.f.* given by

$$f(x) = \begin{cases} \frac{1}{4}x & 0 \le x \le 2 \\ \frac{1}{4}(4-x) & 2 \le x \le 4 \\ 0 & \text{Otherwise} \end{cases}$$

(a) Show that the distribution is a *p.d.f*
(b) Find:
i) P(X < 1)
ii)P(X > 3)
iii) P($1 \le X \le 3$)
iv) P(X ≥ 1/ X ≤ 3)

***Solution***

(a) For a continuous function to be a *p.d.f*, then:

(i) $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

$$\int_0^4 (x)\,dx = \frac{1}{4}\int_0^2 x\,dx + \frac{1}{4}\int_2^4 (4-x)\,dx$$

$$= \frac{1}{4}\left[\frac{x^2}{2}\right]_0^2 + \frac{1}{4}\left[4x - \frac{x^2}{2}\right]$$

$$= \frac{1}{8}[4-0] + \frac{1}{4}[8-6]$$

$$= \frac{4}{8} + \frac{2}{4} = \frac{8}{8} = 1$$

(ii) $f(x) \ge 0$,
Now for all values of $x$ from $x = 0$ to $x = 4$,
$f(x_i) \ge 0$ where $i = 1, 2, 3, 4$

(b) (i) $P$(X < 1) = $\frac{1}{4}\int_0^1 x\,dx$

$$= \left[\frac{x^2}{8}\right]_0^1 = \frac{1}{8}$$

(ii) P(X > 3) = $\frac{1}{4}\int_3^4 (4-x)\,dx$

$$= \frac{1}{4}\left[4x - \frac{x^2}{2}\right]_3^4$$

$$= \frac{1}{8}\left[8x - x^2\right]_3^4$$

$$= \frac{1}{8}(16-15) = \frac{1}{8}$$

(iii) P($1 \le X \le 3$) = $\frac{1}{4}\int_1^2 x\,dx + \frac{1}{4}\int_2^3 (4-x)\,dx$

$$= \frac{1}{4}\left[\frac{x^2}{2}\right]_1^2 + \frac{1}{4}\left[4x - \frac{x^2}{2}\right]_2^3$$

$$= \frac{1}{2}\left[2 - \frac{1}{2}\right] + \frac{1}{4}[7.5-6]$$

$$= \frac{1}{2}\times 1.5 + \frac{1}{4}\times 1.5 = 0.75$$

(iv) $P(X \ge 1/X \le 3) = \frac{P(X \ge 1 \cap X \le 3)}{P(X \le 3)}$

$$= \frac{P(1 \le X \le 3)}{1 - P(X > 3)} = \frac{0.75}{1 - \frac{1}{8}}$$

$$= \frac{0.75}{7/8} = \frac{0.75 \times 8}{7}$$

$$= \frac{6}{7} = 0.857$$

**FINDING PARAMETERS**

### a) **Mean or Expected value**

The expected value of $X$ denoted by E($X$) or $\mu$ is defined as:

$$E(X) = \int_{-\infty}^{+\infty} xf(x)\,dx$$

Where $-\infty$ = lower limit and $+\infty$ upper limit

### b) **Variance**

The variance of X denoted by Var(X) is defined as

$Var(X) = E(X^2) - [E(X)]^2$

Where $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx$

***Note***: Standard deviation = $\sqrt{Var(X)}$

c) **Mode**

This is the value of $X$ for which $f(X)$ is greatest in the given range of $X$.
There are basically two approaches used for finding the mode

(i) Using differential approach: With this approach, we obtain the mode by differentiating the function given and equate it to zero i.e. mode is the value of $X$ for which $f'(x) = \frac{d}{dx} f(x) = 0$
If there are more than one value for which $f'(x) = 0$, then a maximum value is confirmed for $f''(x) < 0$

(ii) Using the graphical approach. With this approach, mode is the value of $X$ corresponding to the highest point of the curve/line (See graphical sketching in the next section)

**Note**: if a function $f(x) = ax$ for a given range, it is quite difficult to obtain the mode by differentiating it as $f'(x) = a$(constant), so in such a situation, the graphical approach is more appropriate.

## Median

This is the value of M for which either

$$\int_{-\infty}^{M} f(x)\,dx = \frac{1}{2} \text{ or } \int_{M}^{+\infty} f(x)\,dx = \frac{1}{2}$$

## Quartiles

The lower quartile is the value of $q_1$ for which either

$$\int_{-\infty}^{q_1} f(x)\,dx = \frac{1}{4} \text{ or } \int_{q_1}^{+\infty} f(x)\,dx = \frac{3}{4}$$

The upper quartile is the value of $q_3$ for which either

$$\int_{-\infty}^{q_3} f(x)\,dx = \frac{3}{4} \text{ or } \int_{q_3}^{+\infty} f(x)\,dx = \frac{1}{4}$$

The interquartile range = upper quartile – lower quartile

***Examples***

1. The continuous random variable $X$ has p.d.f $f(x)$ where

$$f(x) = \begin{cases} \frac{1}{8}x, & 0 \le x \le 4 \\ 0, & \text{else where} \end{cases}$$

Find: (i) E(X)
(ii) Var(X)
(iii) mode
(iv) median
(v) interquartile range

***Solution***

(i) $E(X) = \int_{all\ X} x f(x)dx$

$$= \int_0^4 \tfrac{1}{8}x^2\,dx$$

$$= \frac{1}{8}\left[\frac{x^3}{3}\right]_0^4 = \frac{1}{24}[64-0] = \frac{64}{24} = \frac{8}{3}$$

$= 2.6667$ (4 d.p.)

(ii) $var(X) = E(X^2) - [E(X)]^2$

$$E(X^2) = \int_{all\ X} x^2 f(x)\,dx$$

$$= \int_0^4 \tfrac{1}{8}x^3\,dx = \frac{1}{8}\left[\frac{x^4}{4}\right]_0^4 = \frac{64}{8} = 8$$

$$\text{Var}(X) = 8 - \left(\frac{8}{3}\right)^2 = 0.88889$$

(iii) Mode is the value of X for which $f'(x) = 0$

Given $f(x) = \frac{1}{8}x$;

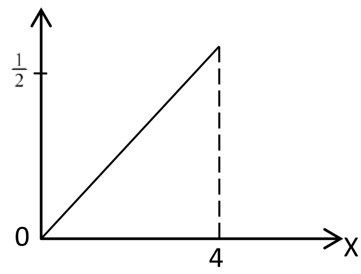
$f'(x) = \frac{1}{8}$

Since $f'(x) = \frac{1}{8}$ is a constant, therefore we cannot obtain the mode using this approach. So we use the graphical approach.

$f(x) = \frac{1}{8}x$

If $x = 0$, $f(x) = 0$

If $x = 4$, $f(x) = \frac{4}{8} = \frac{1}{2}$




For the range of $0 \le x \le 4$, the graph is maximum at $x = 4$, hence mode = 4

(**iv**) Let median = $m$

$$\int_0^m \tfrac{1}{8}x\,dx = \frac{1}{2}$$

$$\frac{1}{8}\left[\frac{x^2}{2}\right]_0^m = \frac{1}{2}$$

$$\frac{m^2}{2} - 0 = \frac{1}{2}$$

$$m^2 = 8$$

$m = 2.828$ (3 d.p.)

(v) Interquartile range = $q_3 - q_1$

For lower quartile $q_1$, we have:

$$\int_0^{q_1} \tfrac{1}{8} x\,dx = \frac{1}{4}$$

$$\frac{1}{8}\left[\frac{x^2}{2}\right]_0^{q_1} = \frac{1}{4}$$

$$\frac{{q_1}^2}{2} = 2 \Rightarrow q_1 = 2$$

For upper quartile $q_3$, we have

$$\int_0^{q_3} \tfrac{1}{8} x\,dx = \frac{3}{4}$$

$$\frac{1}{8}\left[\frac{x^2}{2}\right]_0^{q_3} = \frac{3}{4}$$

$$\frac{{q_3}^2}{2} = 6$$

$${q_3}^2 = 12$$

$$q_3 = 3.464$$

Interquartile range = 3.464 – 2 = 1.464

**2.** A random variable $X$ has probability density function given by:

$$f(x) = \begin{cases} \tfrac{2}{3}(x+1), & -1 \le x \le 0 \\ \tfrac{1}{3}(2-x), & 0 < x \le 2 \\ 0, & \text{otherwise} \end{cases}$$

Find: (i) E($X$) (ii) var($X$)
(iii) mode (iv) median
(v) semi-interquartile range

***Solution***

$$E(X) = \int_{all\,x} xf(x)dx$$

$$= \frac{2}{3}\int_{-1}^{0} x(x+1)\,dx + \frac{1}{3}\int_0^2 x(2-x)\,dx$$

$$= \frac{2}{3}\int_{-1}^{0} (x^2+x)\,dx + \frac{1}{3}\int_0^2 (2x-x^2)\,dx$$

$$= \frac{2}{3}\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_1^0 + \frac{1}{3}\left[x^2 - \frac{x^2}{3}\right]_0^2$$

$$= \frac{2}{3}\left[(0) - \left(\frac{-1}{3} + \frac{1}{2}\right)\right] + \frac{1}{3}\left[\left(4 - \frac{8}{3}\right) - 0\right]$$

$$= \frac{2}{3} \times \frac{-1}{6} + \frac{1}{3} \times \frac{4}{3}$$

$$= \frac{-1}{9} + \frac{4}{9} = \frac{3}{9} = \frac{1}{3}$$

(ii) Var($X$) = E ($X^2$) –[E($X$)]$^2$

$$E(X^2) = \int_{all\,X} x^2 f(X)dx$$

$$= \frac{2}{3}\int_{-1}^{0} x^2(x+1)\,dx + \frac{1}{3}\int_0^2 x^2(2-x)\,dx$$

$$= \frac{2}{3}\int_{-1}^{0} (x^3+x^2)\,dx + \frac{1}{3}\int_0^2 (2x^2-x^3)\,dx$$

$$= \frac{2}{3}\left[\frac{x^4}{4} + \frac{x^3}{3}\right]_{-1}^0 + \frac{1}{3}\left[\frac{2x^3}{3} - \frac{x^4}{4}\right]_0^2$$

$$= \frac{2}{3}\left[(0) - \left(\frac{1}{4} - \frac{1}{3}\right)\right] + \frac{1}{3}\left[\left(\frac{16}{3} - \frac{16}{4}\right) - 0\right]$$

$$= \frac{2}{3} \times \frac{-1}{12} + \frac{1}{3} \times \frac{16}{12}$$

$$= \frac{-2}{36} + \frac{16}{36} = \frac{14}{36} = \frac{7}{18}$$

$$\text{Var}(X) = \frac{7}{18} - \frac{1}{9} = \frac{5}{18}$$

(**iii**) By sketching the graph of $f(x)$

For $-1 \le x \le 0$, $f(x) = \frac{2}{3}(x+1)$, which is a line

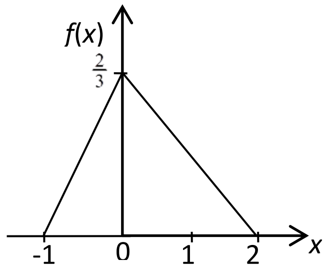
If $x = -1$, $f(x) = 0$

If $x = 0$, $f(x) = \frac{2}{3}$

For $0 \le x \le 2$, $f(x) = \frac{1}{3}(2-x)$ which is also a line

If $x = 0$, $f(x) = \frac{2}{3}$

If $x = 2$, $f(x) = 0$




The graph is maximum at $x = 0$. Hence mode = 0

(iv) ***Note***: If there is more than one interval, we must test for the appropriate interval where the median lies

Now if the median lies in the first interval of

($-1 \le x \le 0$), then $\int_{-1}^{0} \frac{2}{3}(x+1)\,dx \ge 0.5$

$$\frac{2}{3}\int_{-1}^{0}(x+1)\,dx \ge 0.5 = \frac{2}{3}\left[\frac{x^2}{2} + x\right]_{-1}^0$$

$$= \frac{2}{3}\left[(0) - \left(\frac{1}{2} - 1\right)\right]$$

$$= \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$$

But $\frac{1}{3} < \frac{1}{2}$

Hence the median lies in the next interval of $0 < x \le 2$

Let $m$ = median

$$\Rightarrow \int_{-1}^{0} f(x)\,dx + \int_{0}^{m} \frac{1}{3}(2-x)\,dx = \frac{1}{2}$$

$$\frac{1}{3} + \frac{1}{3}\left[2x - \frac{x^2}{2}\right]_0^m = \frac{1}{2}$$

$$\frac{1}{3}\left[2m - \frac{m^2}{2}\right] = \frac{1}{6}$$

Multiplying through by 6;

$4m - m^2 = 1$

$m^2 - 4m + 1 = 0$

$$m = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$m = \frac{4 \pm \sqrt{16-4}}{2}$$

$$m = \frac{4 \pm \sqrt{12}}{2}$$

Either $m = \dfrac{4+\sqrt{12}}{2} = 3.732$

Or $m = \dfrac{4-\sqrt{12}}{2} = 0.2679$

Since median lies in the interval $0 < x \leq 2$,
then median = 0.2679

(iv) Since $\frac{1}{3} > \frac{1}{4}$, hence lower quartile lies in the first interval

Let $q_1$ = lower quartile

$$\frac{2}{3}\int_{-1}^{q} (x+1)dx = \frac{1}{4}$$

$$\frac{2}{3}\left[\frac{x^2}{2} + x\right]_{-1}^{q} = \frac{1}{4}$$

$$\frac{2}{3}\left[\left(\frac{q_1^{\ 2}}{2} + q_1\right) - \left(\frac{1}{2} - 1\right)\right] = \frac{1}{4}$$

$$\frac{2}{3}\left[\frac{q_1^{\ 2}}{2} + q_1 + \frac{1}{2}\right] = \frac{1}{4}$$

$$q_1^{\ 2} + q_1 + \frac{1}{2} = \frac{3}{8}$$

$$q_1^{\ 2} + q_1 + \frac{1}{8} = 0$$

$$8q_1^{\ 2} + 8q_1 + 1 = 0$$

$$q_1 = \frac{-8 \pm \sqrt{64-32}}{16}$$

$$q_1 = \frac{-8+\sqrt{32}}{16} = -0.1464$$

Let $q_3$ = upper quartile

$$\frac{1}{3} + \frac{1}{3}\int_{0}^{q_3} (2-x)dx = \frac{3}{4}$$

$$\frac{1}{3}\left[2x - \frac{x^2}{2}\right]_0^{q_3} = \frac{3}{4}$$

$$\frac{1}{3}\left[2q_3 - \frac{q_3^{\ 2}}{2}\right] = \frac{3}{4} - \frac{1}{3}$$

$$\frac{1}{3}\left(2q_3 - \frac{q_3^{\ 2}}{2}\right) = \frac{5}{12}$$

$$2q_3 - \frac{q_3^{\ 2}}{2} = \frac{5}{4}$$

Multiplying through by 4;

$$2q_3^{\ 2} - 8q_3 + 5 = 0$$

$$q_3 = \frac{8 \pm \sqrt{64-40}}{4}$$

$$q_3 = \frac{8 \pm \sqrt{24}}{4}$$

Either $q_3 = \dfrac{8+\sqrt{24}}{4} = 3.2247$

Or $q_3 = \dfrac{8-\sqrt{24}}{4} = 0.7753$

Since 0.7753 lies in the interval $0 < x < 2$,
hence $q_3 = 0.7753$

semi-interquartile range $= \dfrac{0.7753 - {}^{-}0.1464}{2} = 0.46085$

**3**. A random variable $X$ of a pdf is given by

$$f(x) = \begin{cases} \frac{3}{80}(2+x)(4-x) & 0 \leq x \leq 4 \\ 0, & \text{otherwise} \end{cases}$$

Find the mode

***Solution***

$$f(x) = \frac{3}{80}(2+x)(4-x) = \frac{3}{80}(8 + 2x - x^2)$$

Mode is the value of $x$ for which $f'(x) = 0$

Now $f'(x) = \dfrac{3}{80}(2-2x)$

For $f'(x) = 0$

$$\Rightarrow \frac{3}{80}(2-2x) = 0$$

$2 - 2x = 0$

$x = 1$

Hence mode = 1

**Note**: To confirm that the mode exists, we find $f''(x)$ and see whether its value is less than zero

$$f'(x) = \frac{3}{80}(2-2x)$$

$$f''(x) = -\frac{6}{80}$$

Since $f''(x) < 0$ for all values of $x$, this indicates that there is a maximum point at $x = 1$

**4.** A random variable $X$ of a p.d.f is given by

$$f(x) = \begin{cases} Ax(6-x)^2, & 0 \leq x \leq 6 \\ 0, & \text{else where} \end{cases}$$

a) Find the value of $A$
b) Calculate the mean
c) Calculate the mode

***Solution***

a) $\int_{all\ X} f(x)dx = 1$

$$\int_0^6 Ax(6-x)^2 dx = 1$$

$$A\int_0^6 36x - 12x^2 + x^3 \ dx = 1$$

$$A\left[18x^2 - 4x^3 + \frac{1}{4}x^4\right]_0^4 = 1$$

$$108A = 1$$

$$A = \frac{1}{108}$$

b) E($X$) = $\int_{all\ X} xf(x)dx$

$$= \frac{1}{108}\int_0^6 x^2(6-x)^2 dx$$

$$= \frac{1}{108}\int_0^6 (36x^2 - 12x^3 + x^4)dx$$

$$= \frac{1}{108}\left[12x^3 - 3x^4 + \frac{x^5}{5}\right]_0^6$$

$$= \frac{259.2}{108} = 2.4$$

c) Mode is the value of $x$ for which $f'(x) = 0$

$$f(x) = \frac{1}{108}\ 36x - 24x^2 + x^3$$

$$f'(x) = \frac{1}{108}\ 36 - 48x + 3x^2$$

For $f'(x) = 0$;

$$\Rightarrow \frac{1}{108}\ 36 - 48x + 3x^2 \ = 0$$

$$36 - 48x + 3x^2 = 0$$

$$x^2 - 8x + 12 = 0$$

$$(x-6)(x-2) = 0$$

Either $x - 6 = 0$

$x = 6$

Or $x - 2 = 0$

$x = 2$

Since these two values of $x$ lie in the required range, we find $f''(x)$ and substitute for these two values.

Now $f'(x) = \frac{3}{108}\ x^2 - 8x + 12$

$$f''(x) = \frac{3}{108}\ 2x - 8$$

Substituting for $x = 6$

$f''(6) = \frac{3}{108} \times 4 = \frac{12}{108} > 0$, but for mode to be, $f''(x) < 0$,

hence $x = 6$ is not the mode

Substituting for $x = 2$;

$$f''(2) = \frac{3}{108}\ 4 - 8 \ = \frac{-12}{108} < 0$$

Hence mode = 2

5. The mass $X$ kg of a particular product produced per hour in a certain factory is modelled by a continuous random variable with probability density function given by:

$$f(x) = \begin{cases} \frac{3}{32}x^2, & 0 \le x \le 2 \\ \frac{3}{32}(6-x), & 2 < x \le 6 \\ 0, & \text{else where} \end{cases}$$

a) If the product is sold at Ush. 2000 per kg and the running costs amount to Ush. 800/- per hour, taking $Y$ as the profit made in each hour, express $Y$ in terms of $X$

b) Find the expected value of $Y$

***Solution***

a) Total sales per hour = 2000$X$

Running costs per hour = 800

Profits = 2000$X$ – 800

Hence $Y$ = 2000$X$ – 800

b) E($Y$) = E(2000$X$ – 800)

= E(2000$X$) – E(800) = 2000E($X$) –800

Now $E(X) = \int_0^2 \frac{3}{32}x^3 dx + \int_2^6 \frac{3}{32}(6x - x^2)dx$

$$= \frac{3}{32}\left[\frac{x^4}{4}\right]_0^2 + \frac{3}{32}\left[3x^2 - \frac{x^3}{3}\right]_2^6 = \frac{23}{8}$$

$$E(Y) = 2000 \times \frac{23}{8} - 800$$

= 5750 – 800 = 4950

Hence the expected profit is USh 4950

## The cumulative distribution function, F($x$)

As for discrete probability density function, the cumulative distribution function of a continuous function is denoted by F(x)

## Properties of F($x$)

(i) None decreasing function

(ii) $F(-\infty) = 0$

(iii) $F(+\infty) = 1$

## Steps taken for finding $F(x)$

(i) Consider interval by interval

(ii) For every interval, integrate the function given from lower limit to some specified value $x$ with respect to a dummy variable

(iii) Substitute for the upper limit to be carried forward to the next interval

(iv) Continue this process until you find F(+$\infty$)

**Examples**

1. A random variable $X$ of a continuous pdf is given by

$$f\ \ x\ = \begin{cases} \frac{1}{8}x, & 0 \le\ x \le 4 \\ 0, & \text{elsewhere} \end{cases}$$

Find $F(x)$

***Solution***

For $x < 0$, F($x$) = 0

For $0 \le x \le 4$, $F(x) = \int_0^x \frac{1}{8}t \; dt$

$$\left[\frac{t^2}{16}\right]_0^x = \frac{x^2}{16}$$

$F(4) = \dfrac{16}{16} = 1$

For $x > 4$, $F(x) = 1$

Hence $F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{16}, & 0 \le x \le 4 \\ 1, & x > 4 \end{cases}$

**Note**: when $F(x)$ is given, we obtain $f(x)$ with respect to $x$ by differentiation. I.e. $f(x) = \dfrac{d}{dx}F(x)$

For example given;

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{16}, & 0 \le x \le 4 \\ 1, & x > 4 \end{cases}$$

Find $f(x)$

**Solution**

For $x < 0$, $f(x) = \dfrac{d}{dx}(0) = 0$

For $0 \le x \le 4$, $f(x) = \dfrac{d}{dx}\left(\dfrac{x^2}{16}\right) = \dfrac{2x}{16} = \dfrac{x}{8}$

For $x > 4$, $f(x) = \dfrac{d}{dx}(1) = 0$

Hence $f(x) = \begin{cases} \frac{x}{8}, & 0 \le x \le 4 \\ 0, & \text{elsewhere} \end{cases}$

**2.** $X$ is a continuous random variable with a pdf $f(x)$ given by:

$$f(x) = \begin{cases} \frac{x}{3}, & 0 \le x \le 2 \\ \frac{-2x}{3} + 2, & 2 < x \le 3 \\ 0, & \text{otherwise} \end{cases}$$

Find $F(x)$

***Solution***

For $x < 0$, $F(0) = 0$

For $0 \le x \le 2$, $F(x) = \int_2^x \frac{t}{3}dt = \left[\frac{t^2}{6}\right]_0^x = \frac{x^2}{6}$

$F(2) = \dfrac{4}{6} = \dfrac{2}{3}$

For $2 < x \le 3$, $F(x) = \dfrac{2}{3} + \int_2^x \left(\frac{-2t}{3} + 2\right)dt$

$$= \frac{2}{3} + \left[\frac{-t^2}{3} + 2t\right]_2^x$$

$$= \frac{2}{3} + \left[\left(\frac{-x^2}{3} + 2x\right) - \left(\frac{-4}{3} + 4\right)\right]$$

$$= \frac{2}{3} + \left[\frac{-x}{3} + 2x - \frac{8}{3}\right]$$

$$= \frac{-x^2}{3} + 2x - 2$$

$F(3) = -3 + 6 - 2 = 1$

For $x > 3$, $F(x) = 1$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{6} & 0 \le x \le 2 \\ \frac{-x^2}{3} + 2x - 3, & 2 < x \le 3 \\ 1, & x > 3 \end{cases}$$

Finding $f(x)$ given $F(x)$ above,

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

For $x < 0$, $f(x) = 0$

For $0 \le x \le 2$, $f(x) = \dfrac{d}{dx}\left(\dfrac{x^2}{6}\right) = \dfrac{2x}{6} = \dfrac{x}{3}$

For $2 < x \le 3$, $f(x) = \dfrac{d}{dx}\left(\dfrac{-x^2}{3} + 2x - 3\right) = \dfrac{-2x}{3} + 2$

For $x > 3$, $f(x) = \dfrac{d}{dx}(1) = 0$

Hence $f(x) = \begin{cases} \frac{x}{3}, & 0 \le x \le 2 \\ \frac{-2x}{3} + 2, & 2 < x \le 3 \\ 0, & \text{otherwise} \end{cases}$

**3.** A random variable $X$ of a continuous p.d.f is given by

$$f(x) = \begin{cases} \frac{2}{12}(x+1), & 0 \le x \le 1 \\ \frac{2}{12}, & 2 < x \le 3 \\ \frac{2x}{12}, & 3 < x \le 4 \\ 0, & \text{otherwise} \end{cases}$$

Find $F(x)$

***Solution***

For $x < 0$, $F(x) = 0$

For $0 \le x \le 1$, $F(x) = \dfrac{2}{12}\int_0^x (t+1)dt$

$$= \frac{2}{12}\left[\frac{t^2}{2} + t\right]_0^x = \frac{2}{12}\left(\frac{t^2}{2} + x\right)$$

$F(1) = \dfrac{2}{12}\left(\dfrac{1}{2} + 1\right) = \dfrac{3}{12}$

For $1 < x < 2$, $F(x) = \dfrac{3}{12} + \int_1^x 0\,dx = \dfrac{3}{12}$

For $2 \le x \le 3$, $F(x) = \dfrac{3}{12} + \dfrac{2}{12}\int_2^x dt$

$$= \frac{3}{12} + \frac{2}{12}\, t\,\big|_2^x = \frac{3}{12} + \frac{2}{12}\; x-2$$

$$= \frac{2}{12}x - \frac{1}{12}$$

$F(3) = \frac{6}{12} - \frac{1}{12} = \frac{5}{12}$

For $3 < x \leq 4$, $F(x) = \frac{5}{19} + \frac{2}{19}\int_3^x ddt$

$$F(x) = \frac{5}{12} + \frac{2}{12}\int_3^x ddt$$

$$= \frac{5}{12} + \frac{2}{12}\left[\frac{t^2}{12}\right]_3^x = \frac{5}{12} + \frac{2}{12}\left[\frac{x^2}{2} - \frac{9}{2}\right]$$

$$= \frac{x^2}{12} - \frac{4}{12}$$

$$F(4) = \frac{14-4}{12} = 1$$

For $x > 4$, $F(x) = 1$

$$F\;x = \begin{cases} 0 & x < 0 \\ \frac{1}{12}(x^2 + 2x), & 0 \leq x \leq 4 \\ \frac{1}{4}, & 1 < x < 2 \\ \frac{1}{12}(x^2 - 4), & 2 \leq x \leq 3 \\ \frac{1}{12}(2x - 1), & 3 < x \leq 4 \\ 1, & x > 4 \end{cases}$$

## Finding probabilities, median and quartiles given *F*(*x*)

**a) Finding probabilities**

(i) $P(x \leq x_1) = F(x_1)$

(ii) $P(x_1 < x < x_2) = F(x_2) - F(x_1)$

b) **Finding the median**

After carrying out a test for the interval where the median lies, then the median, $m$ is obtained from

$$F(m) = \frac{1}{2}$$

**Note:** We substitute $m$ directly into the function whose interval gives a value which is at least 0.5

c) **Finding quartiles**

Like for median, the interval over which the lower and upper quartiles lie must be established first and then the lower quartiles $q_1$ is obtained from;

$$E(q_1) = \frac{1}{4}$$

And the upper quartile, $q_3$ is obtained from

$$F(q_3) = \frac{3}{4}$$

***Examples***

**1.** Given the following distribution

$$F\;x = \begin{cases} 0 & x < 0 \\ \frac{x^2}{16}, & 0 \leq x \leq 4 \\ 1, & x > 4 \end{cases}$$

Find: (a) $P(X \leq 2)$

(b) $P(1 \leq X \leq 3)$

(c) median

(d) interquartile range

(e) the value of $a$ such that $P(X > a) = 0.4$

***Solution***

(a) $P(X \leq 2) = P(0 \leq X \leq 2)$

$$= F(2) - F(0)$$

$$= \frac{2^2}{16} = \frac{4}{16} = \frac{1}{4}$$

(b) $P(1 \leq X \leq 3) = F(3) - F(1)$

$$= \frac{3^2}{16} - \frac{1^2}{16} = \frac{9}{16} - \frac{1}{16} = \frac{1}{2}$$

(c) Let $m$ = median

$$\frac{m^2}{16} = \frac{1}{2} \Rightarrow m^2 = 8$$

$$m = 2.828$$

(d) Let $q_1$ and $q_3$ be the lower and upper quartile respectively

$$\Rightarrow \frac{{q_1}^2}{16} = \frac{1}{4}$$

$$q_1 = 2$$

$$\Rightarrow \frac{{q_3}^2}{16} = \frac{3}{4}$$

$${q_3}^2 = 12$$

$$q_3 = \sqrt{12} = 3.464$$

Interquartile range = 3.464 – 2 = 1.464

(e) $P(X > a) = P(a < X \leq 4)$

$F(4) - F(a)$

$$\Rightarrow F(4) - F(a) = 0.4$$

$$\frac{a^2}{16} - \frac{a^2}{16} = 0.4$$

$$1 - \frac{a^2}{16} = 0.4$$

$$\frac{a^2}{16} = 0.6$$

$$a = 3.098$$

**2**. The random variable $X$ has cumulative distribution function

$$F\;x = \begin{cases} 0 & x < 0 \\ \frac{x^2}{6}, & 0 \leq x \leq 2 \\ \frac{-x^2}{3} + 2x - 2, & 2 < x \leq 3 \\ 1 & x > 3 \end{cases}$$

Determine (i) $P(1 \leq X \leq 3)$

(ii) $P(X > 1/X < 2.5)$

(iii) median

(iv) interquartile range

***Solution***

(i) $P(1 \leq X \leq 3) = F(3) - F(1)$

$$= \left(-\frac{9}{3} + 6 - 2\right) - \left(\frac{1}{6}\right) = 1 - \frac{1}{6} = \frac{5}{6}$$

(ii) $P(X > 1/X < 2.5) = \dfrac{P(X > 1 \cap X < 2.5)}{P(X < 2.5)}$

$$= \frac{P(1 < X < 2.5)}{P(X < 2.5)}$$

$$= \frac{F(2.5) - F(1)}{F(2.5)}$$

$$= \frac{\frac{-6.25}{3} + 5 - 2\ - (\frac{1}{6})}{\frac{-6.25}{3} + 5 - 2} = \frac{45}{55} = \frac{9}{11}$$

(iii) Testing for the interval where median lies:

For $0 \le X \le 2,\ F(2) = \dfrac{2^2}{6} = \dfrac{4}{6} = 0.6667 > 0.5$

Hence median lies in the interval $0 \le X \le 2$

Let $m$ = median

$$\frac{m^2}{6} = \frac{1}{2}$$

$$m = 1.732$$

(iv) Testing for the interval where the lower quartile lies

Since 0.6667 > 0.25, hence the lower quartile lies in the interval $0 \le X \le 2$

Let $q_1$ = lower quartile

$$\frac{q_1^{\,2}}{6} = \frac{1}{4}$$

$$q_1 = 1.2247$$

Testing for the interval where the upper quartile lies:

Since 0.6667 < 0.75, hence the upper quartile lies in the interval $2 < X \le 3$

Let $q_3$ = upper quartile

$$\frac{-q_1^{\,2}}{3} + 2q_3 - 2 = \frac{3}{4}$$

Multiplying through by 12;

$$q_3 = \frac{24 \pm \sqrt{576 - 528}}{8} = \frac{24 \pm \sqrt{48}}{8}$$

Either $q_3 = \dfrac{24 + \sqrt{48}}{8} = 3.866$

Or $q_3 = \dfrac{24 - \sqrt{48}}{8} = 2.134$

Since $q_3$ lies in the interval $2 \le X \le 3$;

Hence $q_3 = 2.134$

Interquartile range = 2.134 – 1.2247 = 0.9093

## Graphs of $f(x)$ and $F(x)$

Unlike for discrete pdf, the graph of $f(x)$ and $F(x)$ for continuous pdf are sketched in the same way.

**Steps taken:**

i) If for a given interval, $f(x)$ or $F(x)$ is a line, then:
(a) Find the starting and ending points of the line
(b) Join the points with a straight line

ii) If for a given interval, $f(x)$ or $F(x)$ is a curve, then:
(a) Find the starting and ending points of the curve
(b) Join the points with a curve but not a line

***Note*****:** Steps used in sketching curves in *Pure Mathematics* are not very necessary in Paper II mathematics

### Importance of graphs of *f*(*x*)

1. We use the curve drawn to find the probability between limits; say $P(X_1 \le X \le X_2)$ by finding the area under the curve between $X_1$ and $X_2$
2. If the graph drawn is symmetrical, say about $X = X_1$, then the mean is $X_1$
3. We use these graphs to find the mode of the distribution at the highest(maximum) point of the graph
4. We use graphs to find the constants by equating the area under the graph to one

***Examples***

**1.** A random variable $X$ of a continuous *p.d.f* is given by:

$$f\ x = \begin{cases} \dfrac{x}{8}, & 0 \le x \le 4 \\ 0, & \text{otherwise} \end{cases}$$

Sketch the graph of $f(x)$ and use it to find:

(i) $P(X \le 1)$

(ii) $P(1 \le X \le 2)$

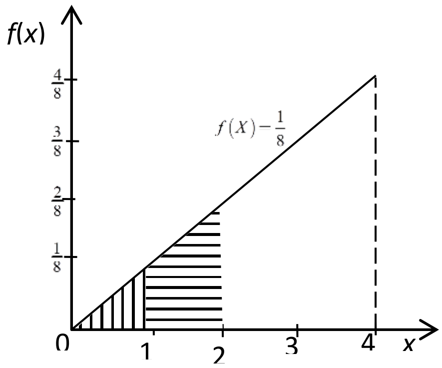
***Solution***

$f(x) = \frac{1}{8}x$

If $x = 0$, $f(x) = 0$

If $x = 4$, $f(x) = \frac{4}{8} = \frac{1}{2}$

**Graph of *f*(*x*)**




(i) $P(X \le 1)$ = Area under the graph between $x = 0$ and $x = 2$

$$= \frac{1}{2} \times \frac{1}{8} \times 1 = \frac{1}{16}$$

(ii) $P(1 \le X \le 2)$ = Area under the graph between $x = 1$ and $x = 2$

$$= \frac{1}{2} \times 1 \times \left(\frac{1}{8} + \frac{2}{8}\right) = \frac{1}{2} \times \frac{3}{8} = \frac{3}{16}$$

**2**. A random variable X of a continuous p.d.f is given by:

$$f(x) = \begin{cases} kx, & 0 \le x \le 2 \\ k(4 - x), & 2 < x \le 4 \\ 0, & \text{otherwise} \end{cases}$$

Sketch the graph of the $f$(x) and use it to determine the:

(i) value of the constant $k$

(ii) $P(1 \le x \le 3)$

(iii) $E(X)$

(iv) mode

***Solution***

For $0 \le x \le 2$, $f(x) = kx$;

If $x = 0$, $f(x) = 0$

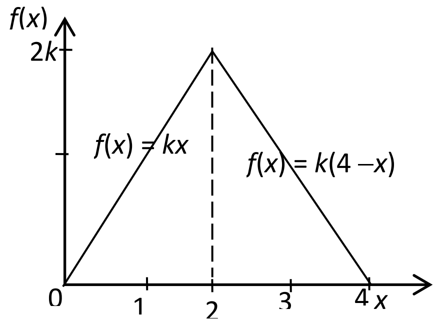
If $x = 2$, $f(x) = 2k$

For $2 \le X \le 4$, $f(x) = k(4 - x)$

If $x = 2$, $f(x) = 2k$

If $x = 4$, $f(x) = 0$

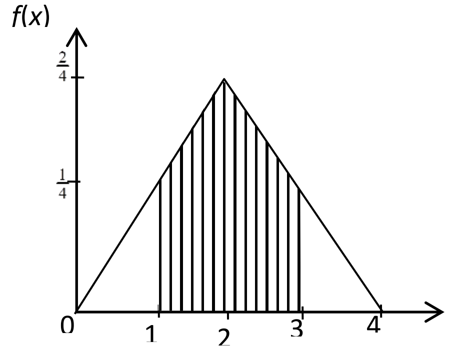
**Graph of $f(x)$**




(i) Area under the graph = 1

$\frac{1}{2}\times 4\times 2k = 1 \Rightarrow k = \frac{1}{4}$

(ii)




$P(1 \le X \le 3)$ = Area under the graph between $x = 1$ and $x = 3$

$$= \left[\frac{1}{2}\times 1\times\left(\frac{1}{4}+\frac{2}{4}\right)\right]\times 2 = \frac{1}{4}+\frac{2}{4} = \frac{3}{4} = 0.75$$

(iii) Since the graph of f($x$) is symmetrical about the line $x = 2$, hence $E(X) = 2$

(iv) The graph has got the highest point at $x = 2$. Hence mode = 2

**3.** A random variable $X$ of a continuous p.d.f is given as

$$f(x) = \begin{cases} \frac{x^2}{27}, & 0 \le x \le 3 \\ \frac{1}{3}, & 3 \le x \le 5 \\ 0, & \text{otherwise} \end{cases}$$

Sketch the graph of f($x$)

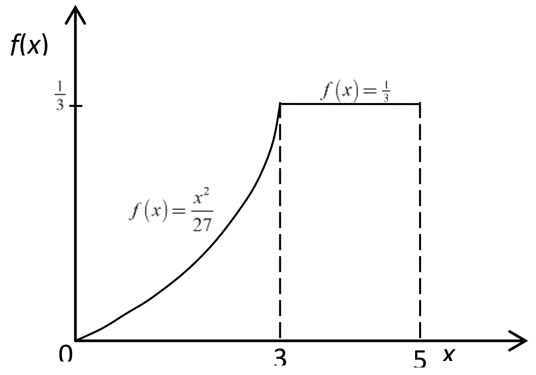
***Solution***

For $0 \le x \le 3$, $f(x) = \frac{x^2}{27}$, which is a curve;

When $x = 0$, $f(x) = 0$

If $x = 3$, $f(x) = \frac{9}{27} = \frac{1}{3}$

For $0 \le x \le 5$, $f(x) = \frac{1}{3}$ which is a horizontal line parallel to the $x-$axis




## Exercise 3.2

**1.** The continuous random variable $X$ has p.d.f $f(x)$ where $f(x) = k(4 - x)$, $1 \le x \le 3$

(a) Find the value of the constant $k$

(b) Sketch $y = f(x)$

(c) Find $P(1.2 \le X \le 2.4)$

**2.** A continuous random variable $X$ has p.d.f $f(x)$ where:

$$f(x) = \begin{cases} kx, & 0 \le x \le 1 \\ k(2-x), & 1 \le x \le 2 \\ 0, & \text{Otherwise} \end{cases}$$

Find (a) the value of the constant $k$

(b) $E(X)$ (iii) var($X$)

(c) $P(\frac{3}{4} \le X \le 1½)$

(d) the mode

**3.** (a) The continuous random variable $X$ has continuous p.d.f $f(x)$, where

$$f(x) = \begin{cases} \frac{1+x}{6}, & 1 \le x \le 3 \\ 0, & \text{otherwise} \end{cases}$$

(i) Sketch the graph of $f(x)$

(ii) Calculate the $E(X)$

(iii) Obtain the cumulative distribution function $F(x)$

(iv) Find $m$ such that $P(X \le m) = \frac{1}{2}$

**4.** (a) The continuous random variable $X$ has cumulative distribution function $F(x)$ where

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{2x}{3}, & 0 \le x \le 1 \\ \frac{x}{3}+k, & 1 \le x \le 2 \\ 1, & x \ge 2 \end{cases}$$

Find the: (a) value of $k$

(b) p.d.f $f(x)$ and sketch it

(c) mean

(d) standard deviation

**5.** A continuous random variable X takes values in the interval 0 to 3.

It is given that $P(X > x) = a + bx^3$, $0 \le x \le 3$

(a) Find the values of the constants $a$ and $b$

(b) Find the cumulative distribution function $F(x)$

(c) Find the probability density function $f(x)$

(d) Show that $E(X) = 2.25$

6. The continuous random variable $X$ has *p.d.f.* given by *f(x)* where

$$f(x)=\begin{cases} kx^2 & 0\le x<3 \\ k & 3\le x\le 5 \\ 0 & \text{otherwise}\end{cases}$$

Find
(a) the value of $k$.
(b) E($X$)
(c) E($X^2$)
(d) the standard deviation of $X$.

7 . The random variable $X$ has probability density function

$$f(x)=\begin{cases} \frac{1}{4} & 0\le x<2 \\ \frac{1}{4}(2x-3) & 2\le x\le 3\end{cases}$$

Find the:
(a) cumulative distribution function, *F(x)*,
(b) median.

8. The continuous random variable $X$ has probability density function $f$ given by

$$f(x)=\begin{cases} k(x+3) & -3\le x\le 3 \\ 0 & \text{otherwise}\end{cases}$$

where $k$ is a constant.
(a) Show that $k=\frac{1}{18}$
(b) Find E($X$) and Var($X$)
(c) Find the lower quartile of $X$,
(d) Let $Y = aX + b$, where $a$ and $b$ are constants with $a > 0$. Find the values of $a$ and $b$ for which E($Y$) = 0 and Var($Y$) = 1.

9. The continuous random variable, $X$, has probability density function defined by

$$f(x)=\begin{cases} kx & 0\le x<8 \\ 8k & 8\le x\le 9 \\ 0 & \text{otherwise}\end{cases}$$

where $k$ is a constant.
(a) Sketch the graph of *f(x)*.
(b) Show that $k = 0.025$.
(c) Determine, for all $x$, the distribution function *F(x)*.
(d) Calculate the probability that an observed value of $X$ exceeds 6.

10. The cumulative distribution function of $X$ is given by

$$F(x)=\begin{cases} 0 & x<0 \\ x^3 & 0\le x\le 1 \\ 1 & x\ge\end{cases}$$

Find the:
(a) value of $k$,
(b) probability density function *f(x)*,
(c) median of $X$,
(d) variance of $X$.

11. The continuous random variable $X$ has (cumulative) distribution function given by

$$F(x)=\begin{cases} 0 & x<-1 \\ \frac{1+x}{8} & -1\le x<0 \\ \frac{1+3x}{8} & 0\le x<2 \\ \frac{5+x}{8} & 2\le x\le 3 \\ 1 & x>3\end{cases}$$

(a) Sketch the graph of the probability density function $f(x)$.
(b) Determine the expectation of X and the variance of X.
(c) Determine P($3 \le 2X \le 5$).

12. The continuous random variable X has probability density function given by

$$f(x)=\begin{cases} \frac{k}{x} & 1\le x\le 9 \\ 0 & \text{otherwise}\end{cases}$$

where $k$ is a constant. Giving your answers correct to three significant figures where, appropriate, find the:
(a) value of $k$, and also the median value of $X$,
(b) mean and variance of $X$,
(c) cumulative distribution function, $F$, of $X$, and sketch the graph of $y = F(x)$.

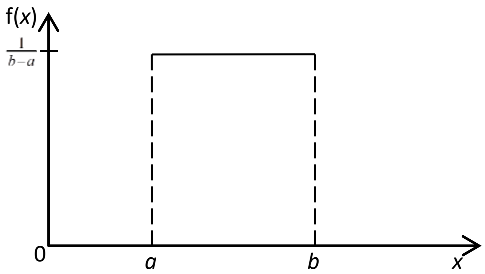
## *Rectangular (Uniform) Distribution*

If a random variable $X$ is said to be uniformly distributed over the interval [$a$, $b$], then its p.d.f is given by:

$$f(x)=\begin{cases} \frac{1}{b-a}, & a\le x\le b \\ 0, & \text{otherwise}\end{cases}$$

**Graphical representation of $f(x)$**

Since the function is a constant over the given interval, $f(x)$ is a horizontal line parallel to the $x$-axis




**Expectation of $X$**

Since the graph is symmetrical about the mid point of a and b, hence the expectation of $X$ is $E(X) = \frac{a+b}{2}$

OR: By calculation, we have:

$$E(X)=\int_{\text{all } x} xf(x)\,dx$$
$$=\int_a^b \frac{x}{b-a}\,dx$$
$$=\left[\frac{x^2}{2(b-a)}\right]_a^b$$
$$=\frac{b^2-a^2}{2(b-a)}=\frac{(b+a)(b-a)}{2(b-a)}$$
$$=\frac{b+a}{2}$$

**Variance of $X$**

The variance of $X$ denoted by var($X$) is given by

$$\text{var}(X) = E(X^2) - [E(X)]^2$$

where $E(X^2) = \int_{all\ x} x^2 f(x)\ dx$

$$= \int_a^b \frac{x^2}{b-a}\ dx$$

$$= \left[\frac{x^3}{3(b-a)}\right]_a^b$$

$$= \frac{b^3 - a^3}{3(b-a)} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)}$$

$$= \frac{b^2+ab+a^2}{3}$$

$$\text{var}(X) = \frac{b^2+ab+a^2}{3} - \left(\frac{b+a}{2}\right)^2$$

$$= \frac{b^2+ab+a^2}{3} - \left(\frac{b^2+2ab+a^2}{4}\right)$$

$$= \frac{b^2-2ab+a^2}{12}$$

$$= \frac{(b-a)(b-a)}{12} = \frac{(b-a)^2}{12}$$

***Examples***

1. A random variable $X$ is distributed uniformly over the interval $^-5 \le x \le ^-2$

Find:

(a) P(-4.3 < $X$ < -28*)*

(b) $E(X)$

(c) the standard deviation of $X$

***Solution***

(a) We know that for a uniform distribution, for any interval a ≤ $x$ ≤ b;

$$f(x) = \frac{1}{b-a}$$

For -5 ≤ $x$ ≤ -2, $f(x) = \frac{1}{-2-^-5} = \frac{1}{3}$

P(-4.3 < $X$ < -2.8 = $\int_{-4.3}^{-2.8} \frac{1}{3}\ dx$

$$= \left[\frac{1}{3}x\right]_{-4.3}^{-2.8}$$

$$= \frac{-2.8+4.3}{3} = \frac{1.5}{3} = 0.5$$

b) $E(X) = \frac{a+b}{2} = \frac{-5+^-2}{2} = -3.5$

c) $S.D. = \sqrt{\text{var}(X)} = \sqrt{\frac{(b-a)^2}{12}}$

$$= \sqrt{\frac{(-2+5)^2}{12}}$$

$$= \frac{3}{\sqrt{12}} = 0.866$$

2. The continuous random variable $Y$ has a rectangular distribution

$$f(y) = \begin{cases} \frac{1}{\pi}, & -\frac{\pi}{2} \le y \le \frac{\pi}{2} \\ 0, & \text{Otherwise} \end{cases}$$

a) Find the mean of $Y$

b) Find the variance of $Y$

***Solution***

a) We know that for a ≤ $x$ ≤ b,

$$E(X) = \frac{a+b}{2}$$

So $E(Y) = \frac{-\frac{\pi}{2}+\frac{\pi}{2}}{2} = 0$

b) $\text{var}(Y) = \frac{(\frac{\pi}{2}+\frac{\pi}{2})^2}{12} = \frac{\pi^2}{12}$

3. The continuous random distribution $X$ is uniformly distributed in the interval $a < x < b$

The lower quartile is 5 and the upper quartile is 9

Find: (a) the value of $a$ and $b$

b) P(6 < $X$ < 7)

c) the cumulative function $F(x)$

***Solution***

a) $f(x) = \frac{1}{b-a}$

For the lower quartile;

$$\int_a^{q_1} f(x)\ dx = \frac{1}{4}$$

$$\int_a^{q_1} \frac{1}{b-a}\ dx = \frac{1}{4}$$

$$\left[\frac{x}{b-a}\right]_a^5 = \frac{1}{4}$$

$$\frac{5-a}{b-a} = \frac{1}{4} \quad \text{.............................. (i)}$$

For upper quartile,

$$\int_a^{q_3} f(x)\ dx = \frac{3}{4}$$

$$\int_a^{q_3} \frac{1}{b-a}\ dx = \frac{3}{4}$$

$$\left[\frac{x}{b-a}\right]_a^9 = \frac{3}{4}$$

$$\frac{9-a}{b-a} = \frac{3}{4} \quad \text{..........................(ii)}$$

Eqn (ii) ÷ Eqn (i)

$$\frac{9-a}{5-a} = 3$$

9 – $a$ = 3(5 – $a$)

9 – $a$ = 15 – 3$a$

2$a$ = 6

$a$ = 3

Substituting for $a$ into Eqn (i)

$$\frac{2}{b-3} = \frac{1}{4}$$

$b$ = 11

b) $f(x) = \frac{1}{11-3} = \frac{1}{8}$

P(6 < $X$ < 7) = $\int_6^7 \frac{1}{8}\ dx = \left[\frac{1}{8}x\right]_6^7 = \frac{7-6}{8} = \frac{1}{8}$

c) $f(x) = \begin{cases} \frac{1}{8}, & 3 \le x \le 11 \\ 0, & \text{Otherwise} \end{cases}$

For $X$ < 3, $F(x) = 0$

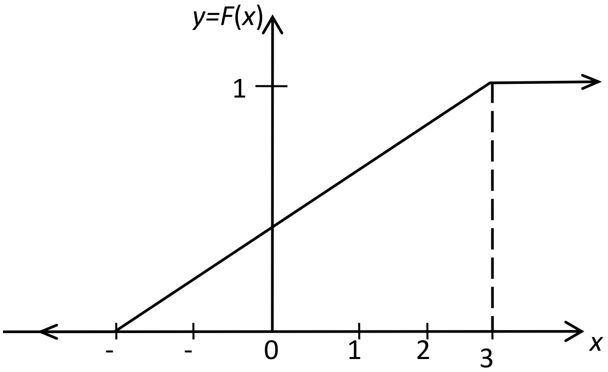
For $3 \le X \le 11$, $F(x) = \int_3^x \frac{1}{8}\, dt = \left[\frac{1}{8}t\right]_3^x = \frac{x-3}{8}$

$$F(11) = \frac{11-3}{8} = 1$$

For $x > 11$, $F(x) = 1$

$$F(x) = \begin{cases} 0, & x < 3 \\ \frac{x-3}{8}, & 3 \le x \le 11 \\ 1, & x > 11 \end{cases}$$

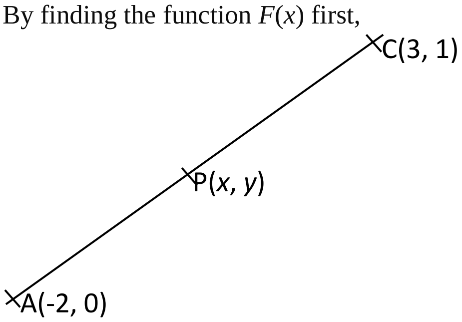
**4**. A random variable $X$ has cumulative distribution *F(x)* illustrated as follows




a) Find the probability density function f($x$)
b) Find the standard deviation of $X$
c) Find the interquartile range
d) Find the 20$^{th}$ percentile

***Solution***

a) By finding the function $F(x)$ first,




Gradient of $\overline{AP}$ = Gradient of $\overline{AC}$

$$\frac{y-0}{x+2} = \frac{1-0}{3+2}$$

$$\frac{y}{x+2} = \frac{1}{5}$$

$$y = \frac{1}{5}(x+2)$$

Hence for $-2 \le x \le 3$, F($x$) = $\frac{1}{5}(x+2)$

For $x < -2$, $F(x) = 0$
For $x > 3$, $F(x) = 1$

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

For x < -2 and x > 3, $f(x) = 0$

For $-2 \le x \le 3$, $f(x) = \frac{1}{5}\frac{d}{dx}(x+2) = \frac{1}{5}$

$$f(X) = \begin{cases} \frac{1}{5}, & -2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

b) $\text{S.D.} = \sqrt{\text{var}(X)}$

but $\text{var}(X) = \frac{(b-a)^2}{12} = \frac{(3+2)^2}{12} = \frac{25}{12}$

$$S.D. = \sqrt{\frac{25}{12}} = 1.443$$

c) Interquartile range = $q_3 - q_1$

For lower quartile, we have;

$$\int_{-2}^{q_1} \frac{1}{5}\, dx = \frac{1}{4}$$

$$\left[\frac{x}{5}\right]_{-2}^{q_1} = \frac{1}{4}$$

$$\frac{q_1+2}{5} = \frac{1}{4}$$

$$q_1 + 2 = \frac{5}{4}$$

$$q_1 = -0.75$$

For the upper quartile;

$$\int_{-2}^{q_3} \frac{1}{5}\, dx = \frac{1}{4}$$

$$\left[\frac{x}{5}\right]_{-2}^{q_3} = \frac{3}{4}$$

$$\frac{q_3+2}{5} = \frac{3}{4}$$

$$q_3 + 2 = \frac{15}{4}$$

$$q_3 = 1.75$$

Interquartile range = 1.75 – -0.75
= 2.5

d) Let $p$ = 20$^{th}$ percentile

$$\int_{-2}^{p} \frac{1}{5}\, dx = \frac{20}{100}$$

$$\left[\frac{x}{5}\right]_{-2}^{p} = \frac{1}{5}$$

$$\frac{p+2}{5} = \frac{1}{5}$$

$p + 2 = 1$
$p = -1$

**5**. The error in percentage made by an examiner may be modelled by the random variable $X$, with probability density function:

$$f(x) = \begin{cases} 0.1, & -3 \le x \le 7 \\ 0, & \text{otherwise} \end{cases}$$

Find the probability that the:
(a) error is positive
(b) magnitude of an error exceeds 2%
(c) magnitude of an error is less than 4%

***Solution***

(a) For an error to be positive, $X > 0$

so $P(X > 0) = \int_0^7 \frac{1}{10}\, dx = \left[\frac{1}{10}x\right]_0^7$

$$\frac{7}{10} - 0 = 0.7$$

(b) $P|X| > 2 = 1 - P(|X| < 2)$

$= 1 - P(-2 < X < 2) = 1 - \int_2^2 \frac{1}{10}\, dx$

$= 1 - \left[\frac{1}{10}x\right]_{-2}^2$

$= 1 - \frac{1}{10}(2+2)$

$= 1 - \frac{4}{10} = \frac{6}{10} = 0.6$

(b) P|$X$| < 4 = P(-4 < $X$ < 4)

**Note**: Care must be taken with this part because for $X < -3$, f($x$) = 0

So P(-4 < $x$ < 4) = P(-3 < $x$ < 4)

$$= \int_{-3}^{4} \tfrac{1}{10}\, dx = \tfrac{1}{10}[x]_{-3}^{4}$$

$$= \frac{4 - ^{-}3}{10} = \frac{7}{10} = 0.7$$

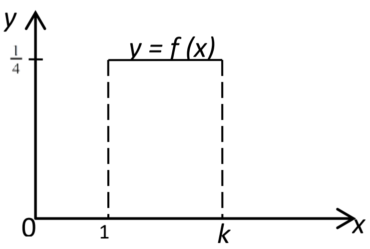
## Exercise 3.3

**1.** $X$ follows a uniform distribution with probability density function: $f(x) = k$, $3 \le x \le 6$

Find (a) the vale of $k$ (b) E($X$)

(c) var($X$) (d) P($X$ > 5)

**2.** The continuous random variable $X$ has a p.d.f $f(x)$ as shown in the diagram




Find:

a) the value of $k$
b) P(2.1 < $X$ < 3.4)
c) E($X$)
d) Var($X$)

**3.** The random variable $Y$ has probability density function given by:

$$f(y) = \begin{cases} 0.2, & 32 \le y \le 37 \\ 0, & \text{otherwise} \end{cases}$$

Find the probability that $Y$ lies within one standard deviation of the mean

**4.** $X$ has a cumulative distribution function

$$F(x) = \frac{x-2}{5},\ 2 \le x \le 7$$

Find (a) $E(X)$ (b) Var($X$)

**5.** The random variable $X$ has probability density function $f(x) = k$, $1 \le x \le 6$

Find the:

(a) value of $k$
(b) cumulative distribution function $F(x)$
(c) 20th percentile
(d) interquartile range

**6.** The length of an off-cut of a wooden planking is a random variable which can take any value up to 0.5m. It is known that the probability of the length being not more than $x$ metres ($0 \le x \le 0.5$) is equal to k$x$.

Determine the:

(a) value of $k$
(b) probability density function of $X$
(c) expected value of $X$
(d) standard deviation of $X$

# Examination Questions

1. (a) A random variable $X$ has the following distribution: P($X$ = 0) = P($X$ = 1), P($X$ = 2) = 0.2, P($X$ = 3) = P($X$ = 4) = 0.3. Find the mean and variance of $X$

(b) A continuous random variable $X$ has the distribution function:

$$F(x) = \begin{cases} 3kx(1 - \frac{x^2}{3}), & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

Determine: (i) the value of $k$

(ii) the probability density function of $X$

(iii) the mean of $X$

(iv) P(X > 0.5/0.25 ≤ X ≤ 1) (**1988 No. 11)**

***Solution***

**(a)**

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| P($X$ =$x$) | 0.1 | 0.1 | 0.2 | 0.3 | 0.3 |
| $x$P($X$ = $x$) | 0 | 0.1 | 0.4 | 0.9 | 1.2 |
| $x^2$P($X$=$x$) | 0 | 0.1 | 0.8 | 2.7 | 4.8 |

$$E(X) = \sum_{all\ X} xP(X = x)$$

$$= 0.1 + 0.4 + 0.9 + 1.2 = 2.6$$

$$\text{var}(X) = E(X^2) - [E(X)]^2$$

$$E(X^2) = \sum x^2 P(X = x)$$

$$= 0.1 + 0.8 + 2.7 + 4.8 = 8.4$$

var(X) = 8.4 – $(2.6)^2$

= 8.4 – 6.76 = 1.64

(b) (i) Since this is already a cumulated probability,

` F(1) = 1

$$\Rightarrow 3k\left(1 - \frac{1}{3}\right) = 1$$

$$2k = 1$$

$$k = \frac{1}{2}$$

(ii) For $0 \le X \le 1$, $f(x) = \frac{d}{dx}\left(\frac{3}{2}x - \frac{x^3}{2}\right)$

$$= \frac{3}{2} - \frac{3x^2}{2} = \frac{3}{2}(1 - x^2)$$

For $X > 1$, $f(x) = \frac{d}{dx}(1^2) = 0$

Hence $f(x) = \begin{cases} \frac{3}{2}(1 - x^2), & 0 \le x \le 1 \\ 0, & \text{Otherwise} \end{cases}$

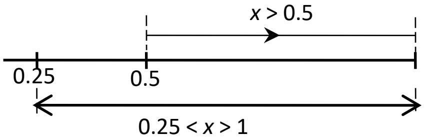
(iii) $E(x) = \int_{-\infty}^{\infty} xf(x)dx = \frac{3}{2}\int_0^1 x(1 - x^2)dx$

$$= \frac{3}{2}\int_0^1 (x - x^3)\, dx$$

$$= \frac{3}{2}\left[\frac{x^2}{2} - \frac{x^4}{4}\right]_0^1$$

$$= \frac{3}{2}\left[\left(\frac{1}{2} - \frac{1}{4}\right) - 0\right] = \frac{3}{2}\cdot\frac{1}{4} = \frac{3}{8}$$

(iv) $P(X > 0.5/0.25 \le X \le 1) = \dfrac{P(X > 0.5 \cap 0.25 \le X \le 1)}{P(0.25 \le X \le 1)}$




$P(X > 0.5 \cap 0.25 \le X \le 1) = P(X > 0.5)$
$= P(0.5 < X < 1)$

$$P(X > 0.5/0.25 \le X \le 1) = \frac{P(0.5 < X < 1)}{P(0.25 \le X \le 1)}$$

$$= \frac{F(1) - F(0.5)}{F(1) - F(0.25)}$$

$$F(x) = \frac{3}{2}x\left(1 - \frac{x^2}{3}\right)$$

Now $F(1) = 1$

$$F(0.5) = \frac{3}{2} \times 0.5\left(1 - \frac{0.25}{3}\right) = 0.6875$$

$$F(0.25) = \frac{3}{2} \times 0.25\left(1 - \frac{0.0625}{3}\right) = 0.3671875$$

$$P(X > 0.5/0.25 \le X \le 1) = \frac{1 - 0.6875}{1 - 0.3671875}$$

$$= \frac{0.3125}{0.6328125} = 0.4938$$

**2.** The outputs of 9 machines in a factory are independent variables each with probability density function given by

$$f(x) = \begin{cases} ax, & 0 \le x \le 10 \\ a(20 - x), & 10 \le x \le 20 \\ 0, & \text{otherwise} \end{cases}$$

Find the:

(i) value of $x$

(ii) expected value and variance of the output of each machine. Hence or otherwise find the expected value and variance of the total output from all machines

***Solution***

$$\int_{all\ x} f(x)dx = 1$$

$$\Rightarrow a\int_0^{10} x\ dx + a\int_{10}^{20} (20 - x)^2\ dx = 1$$

$$a\left[\frac{x^2}{2}\right]_0^{10} + a\left[20x - \frac{x^2}{2}\right]_{10}^{20} = 1$$

$$50a + a(200 - 150) = 1$$

$$100a = 1$$

$$a = \frac{1}{100}$$

***Alternatively,*** this could be solved by sketching the graph of $f(x)$

For $0 \le x \le 10$, $f(x) = ax$

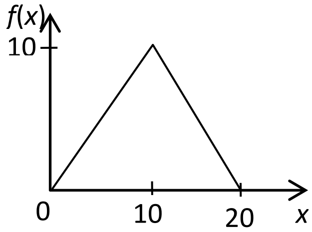
If $x = 0$, $f(x) = 0$

If $x = 10$, $f(x) = 10a$

For $10 \le x \le 10$, $f(x) = a(20 - x)$

If $x = 0$, $f(x) = 10a$

If $x = 10$, $f(x) = 0$




Area under graph = 1

$$\frac{1}{2} \times 20 \times 10a = 1$$

$$100a = 1$$

$$a = \frac{1}{100}$$

(ii) Since the graph is symmetrical about $x = 10$, hence

$E(X) = 10$

$var(X) = E(X^2) - [E(X)]^2$

$$E(X^2) = \int_{all\ x} xf(x)\ dx$$

$$= \frac{1}{100}\int_0^{10} x^3\ dx + \frac{1}{100}\int_{10}^{20} (20x^3 - x^2)\ dx$$

$$= \frac{1}{100}\left[\frac{x^4}{4}\right]_0^{10} + \frac{1}{100}\left[\frac{20x^3}{3} - \frac{x^4}{4}\right]_{10}^{20}$$

$$= \frac{2500}{100} + \frac{1}{100}\left[\left(\frac{160}{3} - 40{,}000\right) - \left(\frac{20{,}000}{3} - 2500\right)\right]$$

$$= 25 + \frac{1}{100}\left(\frac{140{,}000}{3} - 37500\right)$$

$$= 116.6667$$

$Var(X) = 116.667 - (10)^2 = 16.6667$

Expected value of total output from 9 machines is denoted by $E(9X)$

But $E(9X) = 9E(X)$

$= 9 \times 10 = 90$

Variance of total output $= var(9X)$

$= 9^2 var(X)$

$= 81 \times 16.6667 = 1350$

**3. (a)** The number of cars crossing the Owen falls dam daily is uniformly distributed between 1026 to 3025 cars

**(i)** Find the probability that at least 1625 cars cross the bridge

**(ii)** What is the expected number of cars that will cross the bridge on any given day?

(b) The probability density function of a random variable $X$ is

$$f(x) = \begin{cases} K\sin x & \text{for } 0 \le x \le \pi \\ 0, & \text{elsewhere} \end{cases}$$

Determine: (i) the value of $K$

(ii) $P(X > \frac{\pi}{3})$

(iii) the median value of $X$ **(1990 No. 11)**

***Solution***

(a) $f(x) = \frac{1}{3025-1026} = \frac{1}{1999}$

$P(X \geq 1625) = \int_{1625}^{3025} \frac{1}{1999} dx = \left[\frac{x}{1999}\right]_{1625}^{3025}$

$= \frac{3025-1625}{1999} = 0.7$

(b) E($X$) = $\int xf(x)\ dx$

$= \int_{1026}^{3025} \frac{x}{1999} dx = \left[\frac{x^2}{3998}\right]_{1026}^{3025}$

$= \frac{3025^2 - 1025^2}{3998}$

$= \frac{(3025+1026)(3025-1026}{3998}$

$= 2025.5 = 2026 \text{cars}$

(b) (i) $k\int_0^{\pi} \sin x dx = 1$

$k\left[-\cos x\right]_0^{\pi} = 1$

$k\left[-\cos\pi + \cos 0\right] = 1$

$2k = 1$

$k = \frac{1}{2}$

(ii) P(X > $\frac{\pi}{3}$) = $\frac{1}{2}\int_{\pi/3}^{\pi} \sin x dx = 1$

$= -\frac{1}{2}\left[\cos x\right]_{\frac{\pi}{3}}^{\pi} = \frac{-1}{2}\left[\cos\pi - \cos\frac{\pi}{3}\right]$

$= \frac{-1}{2}\left[-1-\frac{1}{2}\right] = \frac{3}{4} = 0.75$

(iii) Let $m$ = median

$\frac{1}{2}\int_0^{m} \sin x\ dx = \frac{1}{2}$

$\frac{-1}{2}\left[\cos x\right]_0^{m} = \frac{1}{2}$

$\cos(m) - \cos 0 = -1$

$\cos(m) - 1 = -1$

$\cos(m) = 0$

$m = \cos^{-1}(0)$

$m = \frac{\pi}{2}$

Hence median = $\frac{\pi}{2}$.

**4.** A random variable $X$ hold the probability function

$$f(x) = \begin{cases} k2^x, & x = 0,1,2,..6 \\ 0, & \text{elsewhere} \end{cases}$$

Determine:

i) the value of $k$
ii) E(X)
iii) P(X < 4/X > 1 **(1991 No. 11)**

***Solution***

(i) $\sum_{all\ x} f(x) = 1$

$k(2^0 + 2^1 + 2^2 + 2^3 + 2^4 + 2^5 + 2^6) = 1$

$k(1 + 2 + 3 + 4 + 5 + 6) = 0$

$127k = 1$

$k = \frac{1}{127}$

(ii)

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $F(x)$ | $\frac{1}{127}$ | $\frac{2}{127}$ | $\frac{4}{127}$ | $\frac{8}{127}$ | $\frac{16}{127}$ | $\frac{32}{127}$ | $\frac{64}{127}$ |
| $xf(x)$ | 0 | $\frac{2}{127}$ | $\frac{8}{127}$ | $\frac{24}{127}$ | $\frac{64}{127}$ | $\frac{160}{127}$ | $\frac{384}{127}$ |

$E(X) = \frac{2}{127} + \frac{8}{127} + \frac{24}{127} + \frac{64}{127} + \frac{160}{127} + \frac{384}{127}$

$= \frac{642}{127} = 5.055$

(iii) P($X$< 4/$X$ > 1) $= \frac{P(X < 4 \cap X > 1)}{P(X > 1)}$

$= \frac{P(x = 2,3)}{P(x = 2,3,4,5,6)}$

$= \frac{P(x=2) + P(x=3)}{1 - [P(x=0) + P(x=1)]}$

$= \frac{4/127 + 8/127}{1 - (\frac{1}{127} + \frac{2}{127})}$

$= \frac{12}{127} = \frac{3}{31} = 0.098$

**5.** A discrete random variable $X$ has the following probability distribution

| $x$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P($X$ = $x$) | $k$ | $2k$ | $3k$ | $4k$ | $5k$ |

i) Determine the value of $k$
ii) Evaluate P(2 < $x \leq$ 4)
iii) Calculate the mean and standard deviation of $X$

**(1992 No. 10)**

***Solution***

(i) $\sum_{all\ x} P(X = x) = 1$

$k + 2k + 3k + 4k + 5k = 1$

$15k = 1$

$k = \frac{1}{15}$

(ii) P(2 < $X \leq$ 4) = P(X = 3) + P(X = 4)

$= \frac{3}{15} + \frac{4}{15} = \frac{7}{15}$

(iii)

| $x$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P($X$ = $x$) | $\frac{1}{15}$ | $\frac{2}{15}$ | $\frac{3}{15}$ | $\frac{4}{15}$ | $\frac{5}{15}$ |
| $x$P($X$= $x$) | $\frac{1}{15}$ | $\frac{4}{15}$ | $\frac{9}{15}$ | $\frac{16}{15}$ | $\frac{25}{15}$ |
| $x^2$P($X$=$x$) | $\frac{1}{15}$ | $\frac{8}{15}$ | $\frac{27}{15}$ | $\frac{64}{15}$ | $\frac{125}{15}$ |

$E(X) = \frac{1}{15} + \frac{4}{15} + \frac{9}{15} + \frac{16}{15} + \frac{25}{15}$

$= \frac{225}{15} = 15$

$\text{var}(X) = 15 - \left(\frac{11}{3}\right)^2 = \frac{14}{9}$

$$S.D = \sqrt{\frac{14}{9}} = 1.247$$

**6**. (a) Which of the functions $f(x)$ and $g(x)$ below in which $c_1$ and $c_2$ are positive constants, can be used as probability density functions (pdfs) over the specified ranges?

(i) $f(x) = c_1 \sin x$ $\quad 0 \le x \le 3\pi$

(ii) $g(x) = c_2 e^{-2x}$; $\quad 0 < x < \ln 10$

Determine the cumulative probability function for the case where a pdf exists

(b) The random variable $X$ has the probability density function

$$f(x) = \begin{cases} cx, & 0 \le x \le 1 \\ c(2-x), & 1 \le x \le 2 \\ 0, & \text{Elsewhere} \end{cases}$$

where c is a constant. Find the median and mode of the probability distribution of $X$ **(1992 No. 12)**

***Solution***

For a continuous pdf, the following two conditions must hold

(i) $\int_{all\ x} f(x)dx = \mathbf{1}$ (ii) $f(x) \ge 0$ for all values of $x$

Considering $f(x)$;

(i) $c_1 \int_0^{3\pi} \sin x \, dx = 1$

$$c_1 \left[-\cos x\right]_0^{3\pi} = 1$$

$$-c_1 \left[\cos 3\pi - \cos 0\right] = 1$$

$$-c_1 \left[-1-1\right] = 1$$

$$2c_1 = 1$$

$$c_1 = \frac{1}{2}$$

Since $c_1$ is positive, hence the condition holds

(ii) For $0 < x < \pi$. $\sin x > 0$ e.g. $\sin\left(\frac{\pi}{2}\right) = 1$

For $\pi < x < 2\pi$, $\sin x < 0$ e.g. $\sin\left(\frac{3\pi}{2}\right) = -1$

For $2\pi < x < 3\pi$, $\sin x > 0$ e.g. $\sin\left(\frac{5\pi}{2}\right) = 1$

Since $f(x) \not\ge 0$ for all values of $x$, therefore this condition does not hold and therefore $f(x)$ is not a p.d.f

Considering $g(x)$;

(i) $c_2 \int_0^{\ln 10} e^{-2x} \, dx = 1$

$$c_2 \left[\frac{-1}{2} e^{-2x}\right]_0^{\ln 10} = 1$$

$$\frac{-c_2}{2}\left[e^{-2\ln 10} - 1\right] = 1$$

$$\frac{-c_2}{2}\left[\frac{1}{100} - 1\right] = 1$$

$$\frac{99c_2}{200} = 1$$

$$c_2 = \frac{200}{99}$$

(ii) For $0 \le x \le \ln 10$, $\quad ce^{-2x} > 0$

(By substitution)

Hence the condition holds, therefore $g(x)$ is a p.d.f

Finding $G(X)$

For $x < 0$, $\quad G(x) = 0$

For $0 \le x \le \ln 10$, $G(x) = \frac{200}{99}\int_0^x e^{-2t} dt$

$$= \frac{200}{99}\left[\frac{-1}{2} e^{-2t}\right]_0^x$$

$$= \frac{-100}{99}\left[e^{-2t} - 1\right]$$

$$= \frac{100}{99}\left(1 - e^{-2x}\right)$$

$$G(\ln 10) = \frac{100}{99}\left(1 - e^{-210}\right)$$

$$= \frac{100}{99}\left(1 - 0.01\right) = \frac{99}{99} = 1$$

For $x > \ln 10$, $G(x) = 1$

$$\text{Hence } G(X) = \begin{cases} 0, & x < 0 \\ \frac{100}{99}(1 - e^{-2x}), & 0 \le x \le x\ln 10 \\ 1, & x > \ln 10 \end{cases}$$

b) Sketching the graph of $f(x)$:

For $0 \le x \le 1$ $\quad f(x) = cx$

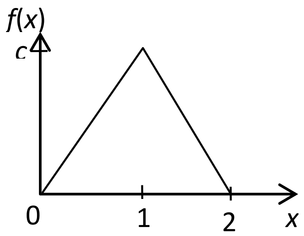
if $x = 0$, $\quad f(x) = 0$

if $x = 1$, $\quad f(x) = c$

For $1 \le x \le 2$, $f(x) = 2(2 - x)$

if $x = 1$, $\quad f(x) = c$

if $x = 2$, $\quad f(x) = 0$




Area under the graph = 1

$$\frac{1}{2} \times 2 \times c = 1$$

$$c = 1$$

Since the graph is symmetrical about $x = 1$, therefore median = 1.

Also since the graph has got the highest point at $x = 1$, therefore mode = 1

**7.** (a) A random variable $X$ has the probability density function

$$f(x) = \begin{cases} k(1 - x^2), & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

where $k$ is a constant. Find

(i) the value of the constant $k$

(ii) the mean of $X$

(iii) the variance of $X$

(b) The number of times a machine breaks down every month is a discrete random variable $X$ with the probability distribution

$$f(x) = \begin{cases} k(\frac{1}{4})^x, & x = 0, 1, 2, 3, \ldots \\ 0, & \text{elsewhere} \end{cases}$$

where $k$ is a constant

Determine the probability that the machine will break down not more than two times a month **(1993 No. 10)**

(i) For a continuous pdf,

$$\int_{all\ x} f(x)dx = 1$$

$$\Rightarrow k\int_0^1 x(1-x^2)dx = 1$$

$$k\left[x-\frac{x^3}{3}\right] = 1$$

$$k\left[\left(1-\frac{1}{3}\right)-0\right] = 1$$

$$\frac{2}{3}k = 1 \Rightarrow k = \frac{3}{2}$$

(ii) Mean of $X$, E($X$) = $\int_{all\ x} xf(x)dx$

$$= \frac{3}{2}\int_0^1 x(1-x^2)dx$$

$$= \frac{3}{2}\left[\frac{x^2}{2}-\frac{x^4}{4}\right]_0^1 = \frac{3}{2}\left[\left(\frac{1}{2}-\frac{1}{4}\right)-0\right]$$

$$= \frac{3}{2}\times\frac{1}{4} = \frac{3}{8}$$

(iii) Variance of $X$, var($X$) = E($X^2$) – [E($X$)]$^2$

E($X^2$) = $\int_{all\ x} x^2 f(x)dx$

$$= \frac{3}{2}\int_0^1 x^2(1-x^2)dx = \frac{3}{2}\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_0^1$$

$$= \frac{3}{2}\left[\left(\frac{1}{3}-\frac{1}{5}\right)-0\right] = \frac{3}{2}\times\frac{2}{15} = \frac{1}{5}$$

$$\text{var}(X) = \frac{1}{5}-\left(\frac{3}{8}\right)^2 = \frac{19}{320}$$

(b) For a discrete probability density function,

$$\sum_{\text{all } x} P(X = x) = 1$$

$\Rightarrow$ P($X$ = 0) + P($X$ = 1) + P($X$ = 2) + …. = 1

$$k+\frac{1}{4}k+\frac{1}{4^2}k+\frac{1}{4^3}k+... = 1$$

$$k\left(1+\frac{1}{4}+\frac{1}{4^2}+\frac{1}{4^3}+...\right) = 1$$

$$k\left[\frac{a}{1-r}\right] = 1$$

$$k\left[\frac{1}{1-\frac{1}{4}}\right] = 1$$

$$\frac{4k}{3} = 1$$

$$k = \frac{3}{4}$$

P($X \le 2$) = P($X$ = 0) + P($X$ = 1) + P($X$ = 2)

$$= \frac{3}{4}\left(1+\frac{1}{4}+\frac{1}{4^2}\right) = \frac{3}{4}\times\frac{21}{16} = \frac{63}{64}$$

**8.** A continuous random variable $X$ has the probability density function

$f(x) = kx(3-x)$

$f(x) = k(4-x)$ for $2 \le x \le 4$

$f(x) = 0$ else where

Find: (i) the value of $k$

(ii) the mean

(iii) $F(x)$, the cumulative distribution function

(iv) P($1.5 \le X \le 3$) **(1994 No 11)**

**Answer** (i) $\frac{3}{16}$ (ii) 1.75

(iii) $F(x) = \begin{cases} 0, & x \le 0 \\ \frac{3}{16}\left(\frac{3}{2}x^2-\frac{x^3}{3}\right), & 0 \le x \le 2 \\ \frac{3}{4}x-\frac{3}{32}x^2-\frac{1}{2}, & 2 \le x \le 4 \\ 1, & x \ge 4 \end{cases}$

(iv) $\frac{11}{16}$

**9.** A random variable $X$ has the probability density function

$$f(x) = \begin{cases} \frac{2}{3a}(x+a), & -a < x \le 0 \\ \frac{1}{3}(2a-x), & 0 < x \le 2a \\ 0, & \text{else where} \end{cases}$$

where $a$ is a constant

Determine: (i) the value of $a$

(ii) the median of $X$

(iii) P($X \le 1.5/X > 0$)

(iv) the cumulative distribution function $F(x)$. Sketch the graph of $F(x)$ **(1995 No 12)**

**Answers**: (i) 1 (ii) 0.2679 (iii) 0.9375

(iv) $f(x) = \begin{cases} 0, & x < -1 \\ \frac{1}{3}(x^2+2x+1), & -1 < x < 0 \\ \frac{1}{6}(2+4x-x^2), & 0 < x < 2 \\ 1 & x \ge 2 \end{cases}$

**10.** The probability density function of a random variable $X$ is given by:

$$f(x) = \begin{cases} k(x+2), & -1 < x \le 0 \\ 2k(1-x), & 0 < x \le 1 \\ 0, & \text{Elsewhere} \end{cases}$$

(i) Sketch the function

(ii) Find $k$ and the mean of $X$

(iii) Find the probability P($0 < X < \frac{1}{2}/X > 0$)

**(1997 No 10)**

**11.** A random variable $X$ has a distribution probability function given by

$$f(x) = \begin{cases} kx, & 0 \le x \le 1 \\ k(4-x^2), & 1 \le x \le 2 \\ 0, & \text{Elsewhere} \end{cases}$$

(i) Find the constant $k$

(ii) Determine $E(X)$ and var($X$)

(iii) Find the cumulative distribution function F($x$) and sketch it **(1998 No 13)**

**Answers:** (i) $\frac{6}{13}$ (ii) 1.1923, 0.1399

(iii) $F(x) = \begin{cases} 0, & x < 0 \\ \frac{3}{13}x^2, & 0 \le x \le 1 \\ \frac{24x-2x^3-19}{13}, & 1 \le x \le 2 \\ 1, & x > 2 \end{cases}$

**12.** A probability density function is given as:

$$f(x)=\begin{cases} kx(4-x^2), & 0\le x<2 \\ 0, & \text{elsewhere}\end{cases}$$

Find the: (i) value of $k$

(ii) median

(iii) mean

(iv) standard deviation (**1998 No 11)**

**Answers: (i)** $\frac{1}{4}$ (ii) 2.6131 (iii) 1.0667 (iv) 0.4422

**13.**(a) A man buys 10 tickets from a total of 200 tickets in a lottery. There is only one prize ticket of Shs. 10,000.

(i) Find the probability that one of the tickets is a prize tickets

(ii) If the prize of each ticket is Shs. 100 and assuming that all tickets are sold, find the expected loss

b) A man leaves at a point which is 20 minutes walk from the taxi stage. Taxis arrive at the stage punctually. If the probability density function for getting a taxi is given by

$$f(x)=\begin{cases} \frac{1}{20}, & 0\le x\le 20 \\ 0, & \text{elsewhere}\end{cases}$$

Determine the:

i) expected time it takes to wait for a taxi

ii) variance of the time it takes to wait for the taxi

(**1999 No. 10)**

**Answers:** a) (i) 0.0478 (ii) 522

b) (i) 10 min (ii) $\frac{100}{3}$ min

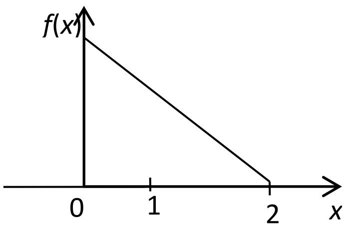
**14**. A random variable $X$ takes on the values of the interval $0 < X < 2$ and the probability density function given by

$$f(x)=\begin{cases} a, & 0<x\le 1\frac{1}{2} \\ \frac{a}{2}(2-x), & 1\frac{1}{2}\le x\le 2 \\ 0, & \text{Elsewhere}\end{cases}$$

Find (i) the value of $a$

(ii) $P(X < 1.6)$

**b)** The probability density function $f(x)$ of the random variable $X$ takes on the form shown in the diagram below




Determine the expression for $f(x)$. Hence obtain the:

i) expression for the cumulative probability density function of $X$

ii) mean and the variance of $X$ (**2000 No 14)**

**Answers:** a) (i) $\frac{16}{25}$ (ii) 0.9744

b) $f(x)=\begin{cases} 1-\frac{1}{2}x, & 0<x<2 \\ 0, & \text{Elsewhere}\end{cases}$

(i) $f(x)=\begin{cases} 0, & x<0 \\ x-\frac{x^2}{5}, & 0<x\le 2 \\ 1, & x>2\end{cases}$ (ii) $E(X)=\frac{2}{3}$, var($X$) = $\frac{2}{9}$

**15.** A continuous random variable $X$ *is* defined by the probability density function:

$$f(x)=\begin{cases} k(x-\frac{1}{a}), & 1\le x\le 3 \\ 0, & \text{elsewhere}\end{cases}$$

Given that P(X > 1) = 0.8, find:

(i) values of a and $k$

(ii) probability that $X$ lies between 0.5 and 2.5

**(iii)** mean of $X$ (**2001 No 15)**

**Answers:** $a = -1$, $k=\frac{2}{15}$ (iii) 0.6667 (iv) 1.8

**16.** (a) A random variable $X$ has the probability density function

$$f(x)=\begin{cases} \frac{1}{b-a}, & a\le x<b \\ 0, & \text{elsewhere}\end{cases}$$

Show that the variance of $X$ is $\frac{(b-a)^2}{12}$

**(b)** During rush hours, it was observed that the number of vehicles departing for Entebbe from Kampala old taxi park take on a random variable $X$ with a uniform distribution over the interval ($X_1$, $X_2$). If in one hour, the expected number of vehicles leaving the stage is 12 with variance of 3, calculate the:

i) value of $X_1$ and $X_2$

ii)Probability that at least 11 vehicles leave the stage

(**2002 No 11)**

**17**. Given the cumulative distribution function:

$$F(x)=\begin{cases} \frac{x^2-1-x}{2}, & 1\le x<2 \\ 3x-\frac{x^2}{2}, & 2\le x\le 3 \\ 1, & x>3\end{cases}$$

(a) Find: (i) the pdf

(ii) $P(1.2 < X < 2.4)$

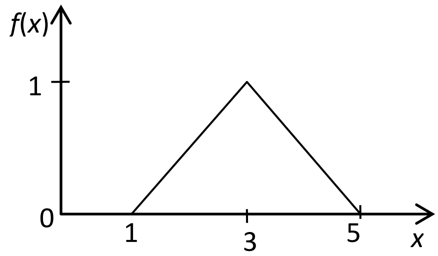
(iii) the mean of $x$

(b) Sketch $f(x)$ (**2003 No 10)**

**Answers**: (i) $f(x)=\begin{cases} x-1, & 1\le x<2 \\ 3-x, & 2\le x\le 3 \\ 0, & \text{elsewhere}\end{cases}$

(ii) **0**.8 (iii) 2

(b)




**18.** The probability density function of a random variable $X$ is given by:

$$f(x)=\begin{cases} k(x+2), & -1<x<0 \\ 2k, & 0\le x\le 1 \\ \frac{K}{2}(5-x), & 1<x\le 3 \\ 0, & \text{elsewhere}\end{cases}$$

b) Sketch the function $f(x)$

c) Find the:

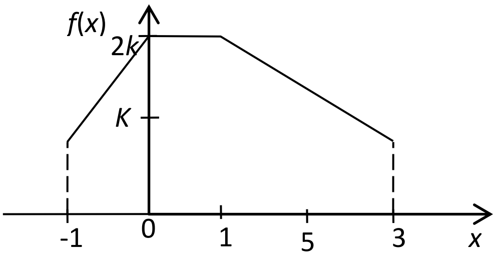
i) value of $k$

ii) mean of $X$

iii) $P(0 < X < \frac{1}{X} > 0)$ (**2004 No 11)**

**Answers**:

(a)




(b) (i) $\frac{2}{13}$ (ii) $\frac{12}{13}$ (iii) $\frac{7}{13}$

**19.** The probability density function of a random variable $X$ is given by:

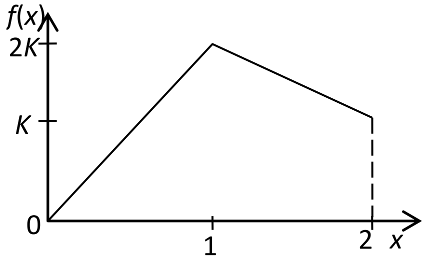
$$f(x)=\begin{cases}2kx, & 0\le x\le 1\\ k(3-x), & 1\le x\le 2\\ 0, & \text{otherwise}\end{cases}$$

a) Sketch the function $f(x)$

b) Find the (i) value of $k$
(ii) mean of $X$
(iii) $P\left(1<\frac{2}{x}>0\right)$ **2005 No 11**

**Answers:**

(a)




b) (i) $\frac{2}{5}$ (ii) $\frac{17}{15}$ (iii) No definite solution

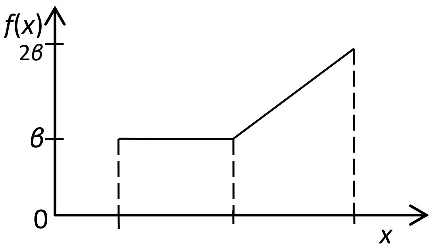
**20**. A continuous random variable $X$ has a probability density function given by:

$$f(x)=\begin{cases}\beta, & 2<x<3\\ \beta(x-2), & 3<x\le 4\\ 0, & \text{otherwise}\end{cases}$$

(i) Sketch $f(x)$
(ii) Find the value of β, hence $f(x)$
(iii) Median, $m$
(iv) P(2.5 < $X$ < 3.5) **2006 No 15**

**Answers:**

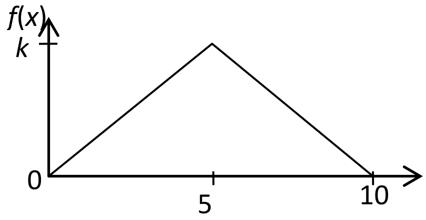
(i)




(ii) $\frac{2}{5}$ $f(x)=\begin{cases}\frac{2}{5}, & 2<x<3\\ \frac{2}{5}(x-2), & 3<x\le 4\\ 0, & \text{otherwise}\end{cases}$

(iii) 3.22475 (iv) 0.65

**21**. The departure, T of pupils from a certain day primary school can be modelled as in the diagram below, where $t$ is the time in minutes after the final bell at 5:00pm




Determine the:
(i) value of $k$
(ii) equation of the p.d.f
(iii) E($T$)
(iv) Probability that a pupil leaves between 4 and 7 minutes after the bell. (**2007 No. 9)**

**Answers: (i)** $\frac{1}{5}$ (ii) $f(t)=\begin{cases}\frac{1}{25}t, & 0\le t\le 5\\ \frac{1}{25}(10-t), & 5<t\le 10\\ 0, & \text{otherwise}\end{cases}$

(iii) 5 (iv) 0.5

**22**. A continuous random variable $X$ has the probability density

$$f(x)=\begin{cases}\lambda(1-\cos x); & 0\le x\le \frac{\pi}{2},\\ \lambda\sin x; & \frac{\pi}{2}<x\le \pi\\ 0 & \text{elsewhere.}\end{cases}$$

a) Find:
(i) the value of λ
(ii) $P\left(\pi/3<X<3\pi/4\right)$

b) Show that the mean μ of the distribution is $1+\frac{\pi}{4}$

**(2008 No 12)**

**Answers**: (a) (i) $\frac{2}{\pi}$ (ii) 0.6982

23. The random variable $X$ has a probability

function $f(x)=\begin{cases}k2^2; x=0,1,2,3\\ 0\ ;\ \text{elsewhere}\end{cases}$

Find:
(a) the value of the constant $k$
(b) $E(X)$ **(2009 No.4)**

**Answers**: (a) $\frac{1}{15}$ (b) $\frac{34}{15}$

24. The probability mass function of a discrete random variable $X$ is given by

$$f(x)=\begin{cases}\frac{16}{15}\times 2^{-x};\ x=1,2,3,4\\ 0 \quad ;\ \text{Otherwise}\end{cases}$$

Find the:
(a) mean of $X$,
(b) variance of $X$.

**Answers:** (a) 1.733 (b) 0.862

25. The continuous random variable $X$ has the probability density function (p.d.f.) given by

$$f(x)=\begin{cases}k_1x, & 1\le x\le 3,\\ k_2(4-x) & 3<x\le 4,\\ 0 & \text{otherwise}\end{cases}$$

where $k_1$ and $k_2$ are constants.

(a) Show that $k_2 = 3k_1$
(b) Find:

(i) the value of $k_1$ and $k_2$

(ii) $E(X)$, the expectation of $X$

**Answers**: (b) (i) $\frac{2}{11}$, $\frac{6}{11}$ (ii) 2.485

26. A random variable $X$ has the following probability distribution:

$P(X=0)=\frac{1}{8}$, $P(X=1)=P(X=2)=\frac{3}{8}$ and $P(X=3)=\frac{1}{8}$. Find the :

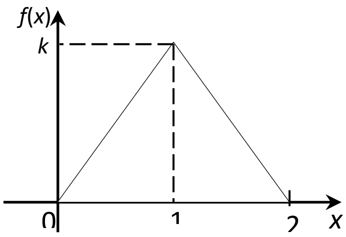
(a) mean value of $X$,

(b) variance of $X$. **(2012 No.8)**

**Answers:** (a) 1.5 (b) 0.75

**27.** A continuous random variable $X$ has a probability density function ($p.d.f$)

$f(x)$ as shown in the graph below.




(a) Find the :

(i) value of $k$,

(ii) expression for the probability density function ($p.d.f$) of $X$.

(b) Calculate the:

(i) mean of $X$,

(ii) $P(X<1.5/X>0.5)$. **(2012 No.12)**

**Answers:** (a) (i) 1 (b)(i) 1 (ii) 0.8751

28. The probability density function (p.d.f) of a continuous random variable X is given by

$$f(x)=\begin{cases} kx(16-x^2), & 0\le x\le 4 \\ 0 \quad , & \text{elsewhere;} \end{cases}$$

where $k$ is a constant.

Find the;

(a) value of $k$.

(b) mode of $X$

(c) mean of $X$ **(2013 No.11)**

**Answers**: (a) $\frac{1}{64}$ (b) 2.31 (c) 2.13

29. The daily number of patients visiting a certain hospital is uniformly distributed between 150 and 210.

(a) Write down the probability distribution function (p.d.f) of the number of patients.

(b) Find the probability that between 170 and 194 patients visit the hospital on a particular day.

**(2014 No.1)**

**Answers**: (a) $f(x)=\begin{cases} \frac{1}{60}, & 150\le x\le 210 \\ 0, & \text{otherwise} \end{cases}$

(b) 0.4

30. The probability density function ($p.d.f.$) of a random variable Y is given by

$$f(x)=\begin{cases} \frac{(y+1)}{4}, & 0\le y\le k \\ 0, & \text{elsewhere} \end{cases}$$

Find

(a) the value of $k$

(b) the expectation of Y.

(c) $P(1\le Y\le 1.5)$ **(2015 No.9)**

**Answers:**

(a) 2 (b) $\frac{7}{6}$ (c) 0.28125