Probability & Statistics

Probability is a concept of Mathematics which measures the degree of certainty or uncertainty of the occurence of events. If any event can happen in m ways and fails in n ways, and each of the $(m + n)$ ways are equally likely to occur, then probability of happening of the events is defined as the ratio, $\frac{m}{m+n}$ and that of its failing as $\frac{n}{m+n}$. If the probability of the happening is denoted by P and not happening by q, then $P + q = 1$, If the event is certain to happen, its probability is unity. If the happening is impossible, then the probability is zero.

Definition : If there are n exhaustive, mutually exclusive and equally likely cases and m of them are favourable to the happening of an event E, then

$$P(E) = \frac{\text{Favourable number of cases}}{\text{Exhaustive number of cases}} = \frac{m}{n}$$

Addition Law of Probability or Theroem of Total Probability

If the probability of an event A happening a result of a trial is P(A) and the probability of a mutually exclusive event B happening is P(B). then the probability of either of the events happening as a result of the trial is

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

Thus for a number of mutually exclusive events $A_1, A_2 \ldots\ldots A_n$ we have

$$\begin{aligned} P(A_1 + A_2 + \ldots\ldots A_n) &= P(A_1 \cup A_2 \cup A_3 \ldots\ldots \cup A_n) \\ &= P(A_1) + P(A_2) + \ldots\ldots + P(A_n) \end{aligned}$$

Conditional Probability—The probability of the happening of an event A when another event B is known to have already happened is called 'Conditional Probability' and is denoted by P(A/B).

Mutually Independent Events—An event A is said to be independent of an event B, if

$$P(A/B) = P(A)$$

i.e., if the probability of happening of A is independent of the happening of B.

Multiplicative Law of Probability—The probability of simultaneous occurence of two events is equal to the probability of the events multiplied by conditional probability of the other *i.e.*, for two events A and B,

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

Where P(B/A) represents the conditonal probability of occurence of B when the event A has already happened.

Generally if $A_1, A_2 \ldots\ldots A_n$ are n independent events, then

$$\begin{aligned} P(A_1, A_2, \ldots\ldots A_n) \\ &= P(A_1) P(A_2) \ldots\ldots P(A_n) \end{aligned}$$

Baye's Theorem—

An event A corresponds to a number of exhaustive events $B_1, B_2 \ldots\ldots B_n$. If $P(B_i)$ and $P(A/B_i)$ are given, then

$$P\left(\frac{B_i}{A}\right) = \frac{P(B_i)\, P(A/B_i)}{\Sigma P(B_i)\, P(A/B_i)}$$

Proof : By the multiplication law of probability

$$P(AB_i) = P(A)\, P(B_i/A) = P(B_i)\, P(A/B_i)$$

$$\therefore\ P\left(\frac{B_i}{A}\right) = \frac{P(B_i)\, P(A/B_i)}{P(A)}$$

Since the event A corresponds to $B_1, B_2 \ldots\ldots B_n$, we have by the addition law of probability

$$\begin{aligned} P(A) &= P(AB_1) + P(AB_2) + P(AB_3) \\ &\quad + \ldots\ldots + P(AB_n) \\ &= \Sigma P(AB_i) \\ &= \Sigma P(B_i)\, P(A/B_i) \end{aligned}$$

$$P\left(\frac{B_i}{A}\right) = \frac{P(B_i)\, P(A/B_i)}{\Sigma P(B_i)\, P(A/B_i)}$$

which is know as the theorem of inverse probability.

Example : There are three bags : First containing 1 white, 2 red, 3 green balls. Second 2 white, 3 red and 1 green balls and third 3 white, 1 red and 2 green balls. Two balls are drawn from a bag chosen at random. These are found to be one white and another red. What is the probability that the balls are come from the second bag ?

Solution : Let B_1, B_2, B_3 pertain to the 1st, 2nd and 3rd bag chosen and A : The two balls are red and white

$$P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}$$

$P(A/B_1)$ = P (a white and a red balls are drawn from frist bag)

$$P(A/B_2) = \frac{2}{5}$$

and $$P(A/B_3) = \frac{1}{5}$$

By Baye's Theorem,

$$P\left(\frac{B_2}{A}\right) = \frac{P(B_2)\,P(A/B_2)}{P(B_1)\,P(A/B_1) + P(B_2)\,P(A/B_2) + P(B_3)\,P(A/B_3)} = \frac{6}{11}$$

Corr : If P_1, P_2 be the probabilities of happening of the two events, then—

(i) the probability that the first event happens and second fails is $P_1(1 - P_2)$

(ii) the probability that atleast one of the events happens is $1 - (1 - p_1)(1 - p_2)$

(iii) the probability that both the events happens is P_1P_2.

Random Variable—If a real variable X be associated with the outcome of random experiment, then since the values variable for instance, if a random experiment E cosists of tossing a pair of dice, the sum X of the two number which turn up have the value 2, 3, 4, …… 12 depending on chance.

If a ranodm variable takes a fintie set of values, it is called a **discrete variate**. On the other hand, if it assumes infinite number of uncountable values, it is called a continuous variate.

Discrete Probability Distribution—Suppose a discrete variate X is the outcome of some experiment. If the probability that X takes the values x_i is P_1, then

$$P(X = x_i) = P_i, \quad \text{for } i = 1, 2, \ldots\ldots,$$

and $$\sum_{i=0}^{n} P_i = 1$$

Distribution Function—The distribution $F(x)$ of the discrete variate X is defined by

$$F(x) = P(X = x) = \sum_{i=1}^{n} P(x_i),$$

where x is any integer.

Example : The probability density function of a variate X is—

X :	0	1	2	3	4	5	6
P(X) :	K	3K	5K	7K	9K	11K	13K

(i) Find $P(X < 4)$, $P(X \geq 5)$, $P(3 < X \leq 6)$.

(ii) What will be the minimum value of K so that $P(X \leq 2) > 3$

Solution : (i) Since

$$\sum_{i=0}^{6} P(x_i) = 1,$$

therefore

$$K + 3K + 5K + 7K + 9K + 11K + 13K = 1$$

$$\Rightarrow \quad K = \frac{1}{49}$$

$$\therefore \quad P(X < 4) = K + 3K + 5K + 7K = 16K = \frac{16}{49}$$

$$P(X \geq 5) = 11K + 13K = 24K = \frac{24}{49}$$

$$P(3 < X \leq 6) = 9K + 11K + 13K = 33K = \frac{33}{49}$$

(ii) $$P(X \leq 2) = K + 3K + 5K = 9K > 3$$

$$\therefore \quad K > \frac{1}{3}$$

Thus, min value of

$$K = \frac{1}{3}.$$

Continuous Probability Function—When a variate X takes every value in an inerval, it gives rise to continuous distribution defined by the variate like heights or weights are continuous distributions.

The probability of a continuous variate X is defined by a function $f(x)$ such that the probability of the variate x failing in the small interval $x - \frac{1}{2}$ dx to $x + \frac{1}{2} dx$ is $f(x)\, dx$.

Symbolically it can be expressed as

$$P\left[x - \frac{1}{2} dx \le x \le x + \frac{1}{2} dx\right] = f(x)\, dx$$

Then $F(x)$ is called the probability density function and continuous curve,

$$y = f(x) \text{ is}$$

the density function,

$$\int_{-\infty}^{\infty} f(x)\, dx = 1$$

Distribution Function—

If $\quad F(x) = P(X \le x)$

$$= \int_{-\infty}^{\infty} f(x)\, dx,$$

then $F(x)$ is defined as the cummulative distribution function or simply the distribution function of the continuous variate X. It is the probability that the value of the variate X will be less than x.

Properties of Distribution Function F(*x*) :

(i) $F'(x) = F(x) \ge 0$ non decreasing function

(ii) $F(\infty)$

(iii) $F(\infty) = 1$

(iv) $P(a \le x \le b) = \int_a^b f(x)\, dx$

$$= \int_{-\infty}^{b} f(x)\, dx - \int_{-\infty}^{a} f(x)\, dx$$

Example : If the function defined as follows a density function ?

(i) $\quad f(x) = e^{-x} \quad x \ge 0$

$\quad\quad\quad\;\; = 0 \quad\quad x < 0$

(ii) If so, determine the probability that the variate having this density will fall in the interval (1, 2).

(iii) Also find the cummulative distribution function F(2).

Solution : (i) Here $f(x) \ge 0$ for all x in (1, 2) and

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{0} f(x)\, dx + \int_{0}^{\infty} f(x)\, dx = 0 + \int_0^{\infty} e^{-x}\, dx$$

$$= -e^{-x} \Big]_0^{\infty} = 1$$

Hence the function satisfy the requirements for a density function.

(ii) Required Probability

$$= P(1 \le x \le 2)$$

$$= \int_1^2 e^{-x}\, dx = e^{-1} - e^{-2}$$

$$= 0{\cdot}368 - 0{\cdot}135$$

$$= 0{\cdot}233$$

(iii) Cummulative Probability function,

$$F(2) = \int_{-\infty}^{2} f(x)\, dx$$

$$= \int_{-\infty}^{0} 0{\cdot}dx + \int_0^2 e^{-x}\, dx$$

$$= 1 - e^{-2}$$

$$= 0{\cdot}865$$

Binomial Distribution

It is concerned with trials of repetitive nature in which only the occurence or non occurence, success or failure, acceptance or rejection, yes or no of a particular event is of interest.

If we perform a series of independent trials such that for each trial, p is the probability of a success and q that of failure, the probability of r success in series of n trials is given by,

$${}^nC_r \cdot p^r q^{n-r} \text{ and } \sum_{r=0}^{n} {}^nC_r\, p^r q^{n-r} = 1$$

Example : The probability that a pen manufactured by a company, will be defective is $\frac{1}{10}$. If 12 such pens are manufactured, find the probability that—

(a) exactly two will be defective.

(b) atleast two will be defective.

(c) none will be defective.

Solution : Probability of defective pen

$$= \frac{1}{10} = 0{\cdot}1$$

$\therefore$ Probability of a non defective pen

$$= 1 - 0{\cdot}1 = 0{\cdot}9$$

(a) Probability that exactly two will be defective $= {}^{12}C_2\, (0{\cdot}1)^2\, (0{\cdot}9)^{11}$

$$= 0{\cdot}2301$$

(b) Probability that at least two will be defective $= 1 -$ [Probability that either no or, one will be defective]

$= 1 - {}^{12}C_0\,(0{\cdot}9)^{12} - {}^{12}C_1\,(0{\cdot}1)\,(0{\cdot}9)^{11}$
$= 0{\cdot}3412$

(c) Probability that none will be defective

$= {}^{12}C_{12}\,(0{\cdot}9)^{12} = 0{\cdot}233$

Mean and Variance of the Binomial Distribution

$$\text{Mean,} \quad u = \sum_{r=1}^{n} r\,p(r) = np;$$

$$\text{Variance,} \quad \sigma^2 = \sum_{r=1}^{n} [r^2\,p(r) - u^2] = npq$$

Poisson Distribution

The Poisson distribution is a limiting case of Binomial distribution. It is derived when of success p is very small. The probability of r success in n trials be

$$p(X = r) = \frac{m^r\,e^{-m}}{r\,!}$$

Where, $m = np \rightarrow$ mean of binomial distribution.

Example : If the probability of a bad reaction from a certain injection is 0·001. Determine the chance that out of 2000 individuals more than two will get a bad reactions.

Solution : $m = np$
$= 2000\,(0{\cdot}001)$
$= 2$

Probability that more than 2 will get a bad reaction

= 1 – [Probability that no one gets a bad reacton + Probability that one gets a bad reaction + Probability that two gets bad reaction]

$$= 1 - \left[e^{-m} + \frac{m^1 e^{-m}}{1\,!} + \frac{m^2 e^{-m}}{2\,!}\right]$$

$$= 1 - \left[\frac{1}{e^2} + \frac{2}{e^2} + \frac{2}{e^2}\right] = 1 - \frac{5}{e^2}$$

$= 0{\cdot}32$

Comparison of Frequency Distribution

1. **Measures of Central Tendency—**

(i) **Mean**—If $x_1, x_2 \ldots\ldots x_n$ be n items, then their arithmetic mean denoted by $\bar{x}$ is given by

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

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

In case the data is arranged in a frequency distribution in which $x_1, x_2, x_3 \ldots\ldots x_n$ are the values of the variable and $f_1, f_2, f_3 \ldots\ldots f_n$ are the corresponding frequencies, then

Arithmetic mean,

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \ldots\ldots + f_nx_n}{f_1 + f_2 + \ldots\ldots + f_n}$$

$$= \frac{\Sigma\, f_ix_i}{f_1 + f_2 + \ldots\ldots f_n}$$

$$= \frac{\Sigma f_ix_i}{\Sigma f_i}$$

Calculation of Mean Short-Cut Methods

When the magnitude of the frequencies and the variable are large, their products turn out to be large, then to reduce labour following method is useful.

Methods of Assumed Mean—Take arbitrary number n, called the assumed mean or the provsional mean, somewhere in the middle of the highest and the lowest values of the variable in the given data,

Shifting the origin to the point, the formula

$$\bar{x} = \frac{\Sigma f_ix_i}{\Sigma f_i}$$

can be written as

$$\bar{x} - a = \frac{\Sigma f_i(x_i - a)}{\Sigma f_i}$$

$$= a + \frac{\Sigma f_id_i}{\Sigma f_i}$$

Where $d_i = a$ is called the diversion of variate x from the assume mean A.

(ii) **Median**—If the items of a series are arranged in ascending or descending order of magnitude, then value of middle item is called the **MEDIAN**. Thus the median has the same number of items above it as below it.

If n is the number of items, then the median is the $\left(\dfrac{n+1}{2}\right)^{\text{th}}$ item in case n is odd.

If n is even, then the median is average of the two middle items.

In case of grouped data, Median (M_d)

$$= l + \frac{\left(\frac{n}{2} - C\right)}{f} \times i$$

where l = lower limit of the class in which the median lies

f = frequency of this class

C = cummulative frequency of the class preceding the median class

i = class interval and n is the total frequency

(iii) **Mode**—The mode is defined as that value of the variable which occurs more frequently *i.e.*, the value of the maximum frequency.

For a grouped distribution, Mode

$$= l + \frac{f_m + f_1}{(f_m - f_1)\,(f_m - f_2)} \times i$$

where l = lower limit of the modal class (*i.e.*, class having the maximum frequency)

f_m = maximum frequency

f_1 and f_2 = frequencies of the classes preceeding and following the modal class.

In symmetrical distribution, the mean, median and mode coincide but in other distribution they are connected as

Mean – Mode = 3 (Mean – Median)

Measures of Dispersion

In a frequency distribution it is equally important to know the variates are cluttered around or scattered away from the point of central tendency. Such variation is called the DISPERSION. Some commonly used measures of dispersion—

(i) Range

(ii) Quartile deviation

(iii) Average or mean deviation

(iv) Standard deviation

(i) **Range**—It is difference between the greatest and least values of the variable *i.e.*, Range = Greatest value – Least value

(ii) **Quartile Deviation**—Quartile are the sizes of the items which divide the series into four equal parts. The second quartile is the MEDIAN. The first and third quartiles also known as LOWER and UPPER QUARTILES a mathematically represented by the formulas

$$Q_1 = 1 + \frac{\left(\frac{n}{4} - C\right)i}{f}$$

$$Q_2 = \frac{l + \left(\frac{3n}{4} - C\right)i}{f}$$

where symbols have their usual meaning.

The difference between upper and lower quantities *i.e.*, $Q_3 - Q_4$ is called the QUARTILE RANGE and heat of this range is known as QUARTILE DEVIATION. It is a better measure than range as it accounting 50% of the data.

(iii) **Average Deviation or Mean Deviation**—It is defined as the arithmetic average of the absolute values of the deviations *i.e.*, all deviation tape pointer, from the mean mode as median.

$$\text{Mathematically A·D}_m = \frac{\Sigma \mid f(x - \bar{x}) \mid}{\Sigma f}$$

Where $A.D_m$ = Average deviation from the mean $\bar{x}$.

(iv) **Standard Deviation**—It is the square root of the mean of square of the differences of the various from the mean Mathematically,

$$\sigma = \sqrt{\left[\frac{\Sigma \mid f_i\,(x_i - \bar{x})^2}{N}\right]}$$

Where, N = total frequency = Σf_i

If the deviations are measured from any other value instead of mean, it is called the root mean square deviation and its square is known as the variance.

Co-efficient of Variation

The ratio of the standard deviation to the mean, *i.e.*, $(\sigma/\bar{x})$ is known as the co-efficient of variation. This is a ratio having no dimension and is used for comparing. The variations between the two groups with different means.

Relation Between Measures of Dispersion

(i) Quartile Deviation $= \frac{2}{3}$ (Standard deviation)

(ii) Mean Deviation—$\frac{1}{5}$ (Standard deviation)

Correlation—In a divariate distribution, if the classes in one variable are associated by classes in the other, then variables are called **CORRELATED**. Correlation is called positive otherwise negative. If the ratio of two variable deviation is constant, then correlation is said to be perfect.

Co-efficient of Correlation—The numerical measures of correlation is called the co-efficient of correlation and is defined as—

$$r = \frac{\Sigma XY}{n\sigma_x\sigma_y}$$

$$= \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$$

$$\left[\because \ \sigma_x^2 = \frac{\Sigma x^2}{n};\ \sigma_y^2 = \frac{\Sigma y^2}{n}\right]$$

where, X = deviation from mean, $\bar{x} = x - \bar{x}$

Y = deviation from mean, $\bar{y} = y = \bar{y}$

σ_x = S.D. of x series

σ_y = S.D. of y series

n = number of values of the two variables

Methods of Calculation

(a) **Direct Method**—Substituting the value of σ_x and σ_y in the above formula, we get

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

$$\text{or, } \frac{n\Sigma xy - \Sigma x\Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2] \times [n\Sigma y^2 - \Sigma x^2\}]}}$$

REGRESSION—Regression is the estimation or production of unknown values of one variable from known value of another variable. It measrues the nature and extent correlation.

Line of Regression—If the reactor diagram indicates some relationship between variables X and Y, then the dots of the reactor diagram will be concentrated round a curve. This curve is called the curve of regression. When the curve is straight line, then it is called line of regression.

Equation of the line of regressin is,

$$y - \bar{y} = \frac{r\sigma_y}{\sigma_x}(x - \bar{x})$$

and, equation of the line of regression of x on y is

$$x - \bar{x} = \frac{r\sigma_x}{\sigma_y}(y - \bar{x})$$

Regression co-efficent

Regression co-efficient is defined as the slope of lines of regression, *i.e.*, regression coefficient of line x on $y = r\frac{\sigma_x}{\sigma_y}$

Solved problems

1. X is a Poisson variable and it is found that the probability that $x = 2$ is two thirds of the probability that $x = 1$. Find the probability that $x = 0$ and the probability that $x = 3$. What is the probability that exceeds ?
2. If a random variable has Poisson distribution such that P(1) = P(2), find
   (i) mean of the distribution.
   (ii) P(4)
3. Fit a Poisson distribution to the set of observations :

x :	0	1	2	3	4
f :	122	60	15	2	1

4. If in a lot of 500 solenoids, 25 are defective; find the probability of 0, 1, 2, 3 defective solenoids in random sample of 20 solenoids.
5. A product of 0·5% defective and is packed in cartons of 100. What percentage contains not more than 3 defectives ?
6. If 10 per cent of the rivets produced by a machine are defective, then find the probability that out of 5 rivets chosen at random.
   (i) none will be defective.
   (ii) one will be defective.
   (iii) at least two will be defective.
7. The probability that on entering student will graduate is 0·4. Determine the probability that out of 5 students,
   (i) None, (ii) one, and (iii) at least one will graduate.
8. Frequency of accidents per shift in a factory is as shown in the following table :

Accident per shift	0	1	2	3	4
Frequency	180	92	24	3	1

   Calculate, the mean number of accidents per shifts.
9. The incidence of occupational disease in an industry is such that the workmen have a 10% chance of suffering from it. What is probability that in a group of 7, 5 or more will suffer from it ?

10. A car hire firm has two cars which it hires out day by day. The numbers of demands for a car on each day is distributed as Poisson distribution with mean 1·5. Calculate the proportion of days.

    (i) On which there is no demand

    (ii) On which demand is refused ($e^{-1\cdot5} = 0\cdot2231$).

11. A manufacturers knows that the condensers he makes contain on the average 1% defectives. He packs them in boxes of 100. What is the probability that a box picked at random will contain 3 or more faulty condensers ?

12. A certain screw making machine produces an average of 2 defective screws out of 100, and packs them in boxes of 500. Find the probability that a box contains 15 defective screws.

13. A sortie of 20 aeroplanes is sent on an operational flight. The chances that an aeroplane fails to return is 5%. Find the probability that—

    (i) One plane does not return

    (ii) at most 5 planes do not return.

14. The probability that a bomb dropped from a plane will strike the target is $\frac{1}{5}$. If six bombs are dropped, then find the probability that—

    (i) exactly two will strike the target.

    (ii) at least two will strike the target.

15. If an average 1 vessel in every 10 is wreeked; find the probability that out of 5 vessels expected to arrive, at least 4 will arrive safely.

16. If the probability that a new born child is a male is 0·6, then find the probability that in a family of 5 children, there are exactly 3 boys.

17. If the chance that one of the ten telephone lines is busy at an instant is 0·2, then—

    (i) What is the chance that 5 of the lines are busy ?

    (ii) What is the probability that all the lines are busy ?

18. If $(x) = \begin{cases} \frac{1}{2}(x+1), & 1 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$

    represents the density of a random variabel X, find $\Sigma(x)$ and var. (x).

19. Find the standard deviation for the following discrete distribution—

x :	8	12	16	20	24
$p(x)$:	1/8	1/6	3/8	1/4	1/12

20. A random variable x has the following probability function—

Values of x :	–2	–1	0	1	2	3
$p(x)$:	0·1	k	0·2	$2k$	0·3	k

Find the value of k of and calculate mean and variance.

EXPLANATIONS

1. Since $p(2) = \frac{2}{3}p(1)$

$$\therefore \quad \frac{m^2}{2!}e^{-m} = \frac{2}{3}me^{-m}$$

$$\Rightarrow \quad m = \frac{4}{3}$$

$$p(0) = e^{-m} = e^{-4/3} = 0\cdot2636$$

$$p(3) = \frac{m^3e^{-m}}{3!} = \frac{\left(\frac{4}{3}\right)^3}{6} \times e^{-4/3} = 0\cdot1041$$

$$p(x > 3) = 1 - p(0) - p(1) - p(2) - p(3)$$

$$= 1 - 0\cdot2636 - \frac{4}{3} \times 0\cdot2626 - \frac{4}{3} \times \frac{2}{3} \times 0\cdot2636 - 0\cdot1041$$

$$= 0\cdot0465$$

2. (i) Probability of r success in Poisson distribution $= \frac{m^r}{r\,!}e^{-m}$

where, m = mean = $n.p.$

$$\frac{m^1}{1\,!}e^{-m} = \frac{m^2}{2\,!}e^{-m}$$

$$me^{-m} = \frac{m^2}{2!}e^{-m}$$

$$m = 2$$

(ii) $$p(4) = \frac{m^r}{r\,!}e^{-m} = \frac{2^4}{4\,!}e^{-2} = \frac{16}{24}e^{-2} = \frac{2}{3}e^{-2}$$

3. $$\text{Mean} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{60 + 36 + 6 + 4}{200} = 0\cdot5$$

Mean of Poisson distribution,

$$m = 0{\cdot}5$$

$\therefore$ Theoretical frequency for r success

$$= \frac{Ne^{-m}\,(m)^r}{r\,!}$$

$$= \frac{200\,e^{-0.5}\,(0{\cdot}5)^r}{r\,!}$$

where, $r = 0, 1, 2, 3, 4$

Theoretical frquencies are

x:	0	1	2	3	4 ($e^{-s} = 0{\cdot}61$)
f:	121	61	15	2	0

4. Probability of defective

$$= \frac{25}{500}$$
$$= 0{\cdot}05$$

Probability of non defective

$$= 1 - 0{\cdot}05$$
$$= 0{\cdot}95$$

Since 20 sample are randomly chosen, therefore

Probability 20 sample are round only chosen

Probability of no defective $= {}^{20}C_0 \times (0{\cdot}95)^{20} = 0{\cdot}3585$

Probability of 1 defective $= {}^{20}C_1 \times (0{\cdot}05) \times (0{\cdot}95)^{19} = 0{\cdot}3773$

Probability of 2 defective $= {}^{20}C_2 \times (0{\cdot}05)^2 \times (0{\cdot}95)^{18} = 0{\cdot}1887$

Probability of 3 defective $= {}^{20}C_3 \times (0{\cdot}05)^3 \times (0{\cdot}95)^{17} = 0{\cdot}0596$

5. Probability of defective $= \frac{5}{100} = 0{\cdot}005$

Probability of non-defective $= 1 - 0{\cdot}005 = 0{\cdot}995$

Since, there are 100 product, hence required probability for not more than 3 defective

$= {}^{100}C_0 \times (0{\cdot}995)^{100} + {}^{100}C_1 \times (0{\cdot}005) \times (0{\cdot}995)^{99} + {}^{100}C_2 \times (0{\cdot}005)^2 \times (0{\cdot}995)^{98} + {}^{100}C_3 \times (0{\cdot}005)^3 \times (0{\cdot}995)^{97}$

$= 0{\cdot}9983$

Required percentage $= 99{\cdot}83$

6. Probability of defective

$$= \frac{10}{100} = 0{\cdot}1$$

Probability of non-defective

$$= 0{\cdot}9$$

Since 5 rivets chosen at random, therefore

(i) In case of none will be defective

Required probability

$$= {}^5C_0 \times (0{\cdot}1)^0 \times (0{\cdot}9)^5$$
$$= 0{\cdot}59049$$

(ii) Probability of one will be defective

$$= {}^5C_1 \times (0{\cdot}1)^1 \times (0{\cdot}9)^4$$
$$= 0{\cdot}32805$$

(iii) Probability of at least 2 will be defective

$$= 1 - {}^5C_0 \times (0{\cdot}9)^5 - {}^5C_1 \times (0{\cdot}1) \times (0{\cdot}9)^4$$
$$= 0{\cdot}08146$$

7. Probability of engineering graduate student $0{\cdot}4$

So probability of non graduate engineering student $= 0{\cdot}6$

Since there are 5 students, hence

(i) Probability of none graduate student

$$= {}^5C_0 \times (0{\cdot}4)^0 \times (0{\cdot}6)^5$$
$$= 0{\cdot}08$$

(ii) Probability of one graduate student

$$= {}^5C_1 \times (0{\cdot}4)^1 \times (0{\cdot}6)^4$$
$$= 0{\cdot}26$$

(iii) Probability of at least one graduate student $= 1 - {}^5C_0 \times (0{\cdot}4)^0 \times (0{\cdot}6)^5$

$$= 1 - 0{\cdot}08 = 0{\cdot}32$$

8. Mean $= \frac{\Sigma f_i x_i}{\Sigma f_i}$

$$= \frac{0 \times 180 + 1 \times 92 + 2 \times 24 + 3 \times 3 + 4 \times 1}{180 + 92 + 24 + 3 + 1}$$

$$= \frac{153}{300}$$
$$= 0{\cdot}51$$

9. Probability of suffering diseases

$$= \frac{10}{100} = 0{\cdot}1$$

Here, $n = 7$

$\therefore$ Mean, $m = np$

$$= 7 \times 0{\cdot}1$$
$$= 0{\cdot}7$$

$$p(x \geq 5) = p(5) + p(6) + p(7)$$

$$= \frac{m^5}{5\,!}e^{-m} + \frac{m^6}{6\,!}e^{-m} + \frac{m^7}{7\,!}e^{-m}$$

$$= e^{-0.7}\left[\frac{(0.7)^5}{5!}+\frac{(0.7)^6}{6!}+\frac{(0.7)^7}{7!}\right]$$
$$= 0.0008$$

10. Mean; $m = 1.5.\ e^{-1.5}$

$= 0.2231$

(i) Probability on which there is no demand,

$$p(0) = e^{-m}$$
$$= e^{-1.5}$$
$$= 0.2231$$

(ii) Since firm has 2 cars, therefore probability on which there is no demand

$$= 1 - p(0) - p(1) - p(2)$$
$$= 1 - 1.2231 - 1.5 \times e^{-1.5} - \frac{1}{2}\times 1.5 \times e^{1.5}$$
$$= 0.1913$$

11. Probability of defective codenser,

$$p = \frac{1}{100}$$
$$= 0.01$$

Mean, $m = np$

$$= 100 \times \frac{1}{100}$$
$$= 1$$
$$p(x \geq 3) = 1 - p(0) - p(1) - p(2)$$
$$= 1 - e^{-m} - me^{-m} - \frac{m^2 e^{-m}}{2!}$$
$$= 1 - e^{-1} - e^{-1} - \frac{1}{2}e^{-1}$$
$$= 0.08$$

12. Probability of defective screws

$$= \frac{2}{100} = \frac{1}{50}$$

Since, $n = 500$,

therefore

Mean, $m = np$

$$= 500 \times \frac{1}{50}$$
$$= 10$$

Probability of 15 defective out of 500

$$= \frac{mre^{-m}}{r!}$$
$$= (10)^{15} \times \frac{e^{-10}}{15!}$$
$$= 0.347 \qquad [\because\ n = 100]$$

13. Probability that an aeroplane fails to returen

$$= \frac{5}{100} = \frac{1}{20}$$

Probability that an aeroplane returns

$$= 1 - \frac{1}{20} = \frac{19}{20}$$

Since 20 planes are sended, hence

(i) Probability that one plane does not return

$$= {}^{20}C_1\left(\frac{1}{20}\right)\left(\frac{19}{20}\right)^{19}$$

(ii) Probability that at most 5 planes does not return

$$= {}^{20}C_0\left(\frac{1}{20}\right)^0\left(\frac{19}{20}\right)^{20} + {}^{20}C_1\left(\frac{1}{20}\right)^1\left(\frac{19}{20}\right)^{19} + {}^{20}C_2\left(\frac{1}{20}\right)^2\left(\frac{19}{20}\right)^{18}$$
$$+ {}^{20}C_3\left(\frac{1}{20}\right)^3\left(\frac{19}{20}\right)^{17} + {}^{20}C_4\left(\frac{1}{20}\right)^4\left(\frac{19}{20}\right)^{16} + {}^{20}C_5\left(\frac{1}{20}\right)^5\left(\frac{19}{20}\right)^{15}$$
$$= \sum_{r=1}^{5} {}^{20}C_r\left(\frac{1}{20}\right)^r\left(\frac{19}{20}\right)^{20-r}$$

14. Probability of non striking the target by bomb $= \frac{4}{5} = 0.8$

Probability of striking the target by bomb

$$= \frac{1}{5} = 0.2$$

Since 6 bombs are dropped, hence

(i) Required probability when at least 2 will strike the target

$$= 1 - {}^6C_0\,(0.8)^6 - {}^6C_1\,(0.2)^1\,(0.8)^5$$
$$= 0.345$$

(ii) Probability of exactly 2 bombs striking the target

$$= {}^6C_2 \times (0.2)^2 \times (0.8)^4$$
$$= 0.245$$

15. Probability of one vessel wreeked out of

$$10 = \frac{1}{10}$$

Probability of no vessel wreeked out of

$$10 = \frac{9}{10}$$

Required probability when at least 4 arrive safely out of 5

$$= {}^5C_4\left(\frac{9}{10}\right)^4\left(\frac{1}{10}\right)^1 + {}^5C_5\left(\frac{9}{10}\right)^5$$

$$= \frac{5 \times 9^4}{10^5} + \frac{9^5}{10^5}$$

$$= \frac{45927}{50000}$$

16. Probability that new born child is male

$= 0{\cdot}6$

Probability that new born child is female

$= 0{\cdot}4$

$\therefore$ Probability of 3 boys out of 5

$= {}^5C_3 \times (0{\cdot}6)^3 \times (0{\cdot}4)^2$

$= 0{\cdot}3456$

17. Probability of one telephone busy out of ten

$= 0{\cdot}2$

(i) Required probability

$= {}^{10}C_5 \times (0{\cdot}2)^5 \times (0{\cdot}8)^5$

$= 0{\cdot}026$

(ii) Probability that all lines are busy

$= {}^{10}C_{10} \times (0{\cdot}2)^{10} \times (0{\cdot}8)^0$

$= 1{\cdot}004 \times 10^{-7}$

18.
$$\Sigma(x) = \int_{-1}^{+1} x f(x)\, dx$$

$$= \int_{-1}^{+1} x \times \frac{1}{2}(x+1)\, dx$$

$$= \frac{1}{2}\int_{-1}^{+1}(x^2 + x)\, dx$$

$$= \frac{1}{2}\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_{-1}^{+1}$$

$$= \frac{1}{2}\left[\frac{1}{3} + \frac{1}{2} + \frac{1}{3} - \frac{1}{2}\right]$$

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

$$\text{Var}(x) = \int_{-1}^{+1} (x - \upsilon)^2 f(x)\, dx$$

$$= \int_{-1}^{+1} \left(x - \frac{1}{3}\right)^2 \frac{1}{2}(x+1)\, dx$$

$$\frac{1}{18}\int_{-1}^{+1} (9x^2 - 6x + 1)(x+1)\, dx$$

$$= \frac{1}{18}\int_{-1}^{+1} (9x^3 + 3x^2 - 5x + 1)\, dx$$

$$= \frac{1}{18}\left[\frac{9x^4}{4} + x^3 - \frac{5x^2}{2} + x\right]_{-1}^{+1}$$

$$= \frac{1}{18}\left[\frac{9}{4} + 1 - \frac{5}{2} + 1 - \frac{9}{4} + 1 + \frac{5}{2} + 1\right]$$

$$= \frac{4}{18}$$

$$= \frac{2}{9}$$

19.
$$\text{Mean, } \upsilon = \Sigma x \cdot p(x)$$

$$= 8 \times \frac{1}{8} + 12 \times \frac{1}{6} + 16 \times \frac{3}{8} + 20 \times \frac{1}{4} + 24 \times \frac{1}{12}$$

$$= 16$$

$$\therefore \quad \sigma^2 = \Sigma(xu)^2, p(x)$$

$$= (8-16)^2 \times \frac{1}{8} + (12-16)^2 \times \frac{1}{6} + (16-10)^2 \times \frac{3}{8} + (20-16)^2 \times \frac{1}{4} + (24-16)^2 \times \frac{1}{12}$$

$$= 8 + \frac{16}{6} + 4 + \frac{32}{6}$$

$$= 20$$

$\therefore$ Standard deviation

$= \sigma$

$= (20)^{1/2}$

$= 2\sqrt{5}$

20. For the value of k,

$$\Sigma p(x) = 1$$

$$\therefore 0{\cdot}1 + k + 0{\cdot}2 + 2k + 0{\cdot}3 + k = 1$$

$$\Rightarrow \quad 4k = 0{\cdot}4$$

$$\Rightarrow \quad k = 0{\cdot}1$$

$$\therefore \text{Mean} = \Sigma p(x)$$

$$= -2 \times 0{\cdot}1 - 1 \times 0{\cdot}1 + 0 \times 0{\cdot}2 + 1 \times 2 \times 0{\cdot}1 + 2 \times 0{\cdot}3 + 3 \times 0{\cdot}1 \quad [\text{taken } k = 0{\cdot}1]$$

$$= -0{\cdot}2 - 0{\cdot}1 + 0 + 0{\cdot}2 + 0{\cdot}6 + 0{\cdot}3$$

$$= 0{\cdot}8$$

$$\sigma^2 = \Sigma(x - u)^2\, p(x)$$

$$= (-2 - 0{\cdot}8)^2 \times 0{\cdot}1 + (-1 - 0{\cdot}8)^2 \times 0{\cdot}1 + (0 - 0{\cdot}8)^2 \times 0{\cdot}2 + (1 - 0{\cdot}8)^2 \times 0{\cdot}2 + (2 - 0{\cdot}8)^2 \times 0{\cdot}3 + (3 - 0{\cdot}8)^2 \times 0{\cdot}1$$

$$= 0{\cdot}1[(-2{\cdot}8)^2 + (-1{\cdot}8)^2 + (-0{\cdot}8)^2 \times 2 + (0{\cdot}2)^2 \times 2 + (1{\cdot}2)^2 \times 3 + (2{\cdot}8)^2]$$

$$= 0{\cdot}1[7{\cdot}84 + 3{\cdot}24 + 1{\cdot}28 + 0{\cdot}8 + 4{\cdot}31 + 7{\cdot}84]$$

$$= 2{\cdot}23$$

Solved Problems (Statistics)

1. Find the mode from the following data—

Age	Frequency
0 – 6	6
6 – 12	11
12 – 18	25
18 – 24	35
24 – 30	18
30 – 36	12
36 – 42	6

2. Find the standard deviation of the following two series which one shows greater variaton :

Series A	Series B
192	83
288	87
236	93
229	109
184	124
260	126
348	126
291	101
330	102
243	108

3. Calculate the median of following distribution giving marks obtained by students in a certain examination :

Marks	No. of Students
40 – 50	2
30 – 40	7
20 – 30	12
10 – 20	9
0 – 10	1

4. The following data related to sizes of shoes at a store during a given week. Find the median size of the shoes. Also calculate the quartile; 7th dectile and 8th percentile.

Size	Frequency
2	2
5·5	5
6	15
6·5	30
7	60
7·5	40
8	23
8·5	11
9	4
9·5	1

5. Calculate the median and quartiles of the following items :

14, 8, 6, 12, 15, 9, 18.

6. Psychological tests of intelligence and engineering ability were applied to 10 students. Here is a record of ungrouped data showing intelligence ratio (I.R.) and engineering ratio (E.R.), calculate the co-efficient of correlation.

Student	I.R.	E.R.
A	105	101
B	104	103
C	102	100
D	101	98
E	100	95
F	99	96
G	98	104
H	96	92
I	93	97
J	92	94

7. The number examined, the mean weight and S.D. in each group of examination by three medical examiners are given below. Find the mean weight and S.D. of the entire data when grouped together.

Mean Exam	Number Examined	Mean wt. (lbs)	SD (lbs)
A	50	113	6
B	60	120	7
C	90	115	8

8. Find the mean deviation of the following frequency distribution :

Class	Frequency
0 – 6	8
6 – 12	10
12 – 18	12
18 – 24	9
24 – 30	5

9. The daily earning of 60 workers in rupees are given below, calculate the average income per worker.

Rupees, x	Number of workers, f	Product
2	10	20
3	16	48
4	11	44
5	8	40
6	6	36
7	4	28
8	3	24
9	2	18
$\Sigma f = 60$	$\Sigma fx = 35$	258

10. Short cut method : The following figures are giving the heights in cms of 7 kids chosen at random. Calculate the simple arithmetic mean by direct method and shortcut method.

S. No.	Height, x
1	64
2	59
3	67
4	69
5	65
6	70
7	68

11. In a partially destroyed laboratory record; only the lines of regression of y on x and x on y are available as $4x - 5y + 33 = 0$ and $20x - 9y = 107$ respectively. Calculate $\bar{x}$, $\bar{y}$ and the coefficient of correlation between x and y.

12. The two regression equations of the variables x and y are $x = 19{\cdot}13 - 0{\cdot}87y$ and $y = 11{\cdot}64 - 0{\cdot}50x$. Find—

(i) Mean of x's

(ii) Mean of y's

(iii) Correlation co-efficient between x and y

EXPLANATIONS

1.

Age	Frequency	Cummulative Frequency
0 – 6	6	6
6 – 12	11	17
12 – 18	$25 = f_{-1}$	42
18 – 24	$35 = f$	77
24 – 30	$18 = f_1$	95
30 – 36	12	107
36 – 42	6	113

$$\therefore \text{Mode} = l + \frac{f - f_{-1}}{2f - f_{-1} - f_1} \times i$$

$$= 18 + \frac{35 - 25}{70 - 25 - 18} \times 6$$

$$= 20{\cdot}22$$

2.

Series A			Series B		
Size x	Dev. from ass. avg. 260	Sq. of dev. dx^2	Size y	Dev from 105 (dy)	Sq. of dev. dy^2
192	–68	4624	83	–22	484
288	28	784	87	–18	324
236	–24	576	93	–12	144
229	–81	961	109	4	16
184	–76	5776	124	19	361
260	0	0	126	21	441
348	88	7744	126	21	441
291	31	961	101	–4	16
330	70	4900	102	–3	9
243	–17	289	108	3	9
2601		$\Sigma dx^2 = 26615$	1059		2245

For A, Arithmetic mean,

$$\bar{x} = \frac{\Sigma x}{n} = \frac{2601}{10} = 260\cdot 1$$

Standard deviation,

$$\sigma = \sqrt{\frac{\Sigma dx^2}{n} - (\bar{x} - 0)^2}$$

$$= \sqrt{\frac{26615 - 10(0\cdot 1)^2}{10}} = \sqrt{2661\cdot 49}$$

$$= 57\cdot 6 \text{ (Approx)}$$

Coefficient of variation

$$= \frac{\sigma}{x} \times 100 = \frac{51\cdot 6}{260\cdot 1} \times 100 = 19\cdot 8$$

For B, Arithmetic mean

$$= \frac{1059}{10} = 105\cdot 9$$

Standard deviation

$$= \sigma = \sqrt{\frac{2245}{10} - (105\cdot 9 - 105)^2}$$

$$= 14\cdot 96$$

Co-efficient of variation

$$= \frac{14\cdot 96}{105\cdot 9} \times 10 = 14\cdot 1$$

Since the co-efficient of variation for series A is greater than that of B, hence series A show greater variation.

3.

Marks	Frequency	Cummulative Frequency
40 – 50	2	2
30 – 40	7	9
20 – 30	12	21
10 – 20	9	30
0 – 10	1	31

Median = size of $\left(\frac{n+1}{2}\right)^{\text{th}}$ *i.e.,* 16th item which lies in 20 – 30

Now,

$$\text{Median} = l_1 + \frac{l_2 - l_1}{f}(m - c')$$

Where, c' = total frequency = cummulative frequency of the median group.

$$\therefore \quad \text{Median} = 20 + \frac{30-20}{12}[16 - (31 - 21)]$$

$$= 20 + \frac{10}{12} \times 6 = 26 \text{ marks.}$$

4. Cummulative

2 7 22 52 112 152 175 186 190 191

$$\text{Median} = \text{Size of } \left(\frac{n+1}{2}\right)^{th}$$

or $\left(\frac{191+1}{2}\right)^{th}$ *i.e.* 96^{th} pair = 7

$$\text{Lower quartile} = \text{size of } \left(\frac{n+1}{4}\right)^{th}$$

or $\left(\frac{191+1}{4}\right)^{th}$ *i.e.,* 48^{th} pair = 6·5

$$\text{Upper quartile} = \text{size of } \left[\frac{3(n+1)}{4}\right]^{th}$$

or 144^{th} pair = 7·5

$$2^{nd} \text{ qurtile} = \text{size of } \left[\frac{2(n+1)}{5}\right]^{th}$$

or $76\cdot 8^{th}$ pair = 7

$$7^{th} \text{ dectile} = \text{size of } \left[\frac{7(n+1)}{10}\right]^{th}$$

or $134\cdot 5^{th}$ pair = 8

$$85^{th} \text{ percentile} = \text{size of } \left[\frac{85(n+1)^{th}}{100}\right]$$

or $163\cdot 2^{th}$ pair = 8

5. Arranging them in the ascending order of magnitude.

S. No.	Size in items
1	6
2	8
3	9
4	12
5	14
6	15
7	18

$$\text{Median} = \text{size of } \left(\frac{n+1}{2}\right)^{th}$$

or, 4^{th} item = 12

$$\text{Lower quartile} = \text{size of } \left(\frac{n+1}{4}\right)^{th}$$

or $\left(\frac{7+1}{2}\right)^{th}$ or 2^{nd} item = 8

Upper quartile = size of $\frac{3}{4}(n+1)^{th}$

or $\left(\frac{3\times 8}{4}\right)^{th}$ or 6^{th} item = 15

6.

Student	I. x	R. $x - x = X$	E. y	R $y - y = Y$	X^2	Y^2	XY
A	105	6	101	3	36	9	18
B	104	5	103	5	25	25	25
C	102	3	100	2	9	4	6
D	101	2	98	0	4	0	0
E	100	1	95	–3	1	9	–3
F	99	0	96	–2	0	4	0
G	98	–1	104	6	1	36	–6
H	96	–3	92	–6	9	36	18
I	93	–6	97	–1	36	1	6
J	92	–7	94	–4	49	16	28
Total	990	0	980	0	170	140	92

Mean of x, $\bar{x} = \dfrac{990}{10}$

$= 99$

and Mean of y, $\bar{y} = \dfrac{980}{10}$

$= 98$

$\Sigma X^2 = 170$

$\Sigma Y^2 = 140$

and $\Sigma XY = 92$

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \Sigma Y^2}}$$

$$= \frac{92}{\sqrt{170 \times 140}}$$

$$= 0{\cdot}59$$

7. Here $n_1 = 50$,

$\bar{x}_1 = 113$,

$\sigma_1 = 6$

$n_2 = 60$,

$\bar{x}_2 = 120$,

$\sigma_2 = 7$

$n_3 = 90$,

$\bar{x}_3 = 115$,

$\sigma_3 = 8$

$\therefore$ Mean of entire data,

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{n_1 + n_2 + n_3}$$

$$= \frac{50 \times 113 + 60 \times 120 + 90 \times 115}{50 + 60 + 90}$$

$$= 116$$

If σ is the S.D. of the entire data, then

$$N\sigma^2 = n_1\sigma_1{}^2 + n_2\sigma_2{}^2 + n_3\sigma_3{}^2 + n_1D_1{}^2 + n_2D_2{}^2 + n_3D^2{}_3$$

Where, $N = n_1 + n_2 + n_3$

$= 200$

$D_1 = \bar{x}_1 - \bar{x}$

$= -3$,

$D_2 = \bar{x}_2 - \bar{x}$

$= 4$,

$D_3 = \bar{x}_3 - \bar{x}$

$= -1$

$$\therefore\ 200\,\sigma^2 = 56 \times 36 + 60 \times 49 + 90 \times 64 + 50 \times 9 + 60 \times 16 + 90 \times 1$$

$$\Rightarrow \quad \sigma^2 = \frac{12000}{200}$$

$$= 60$$

$$\Rightarrow \quad \sigma = \sqrt{60}$$

$$= 7{\cdot}7461$$

8.

Class	Mid Value (x)	Frequency (f)	$d = x - a$	fd	$\lvert x - 14 \rvert$	$f \lvert x - 14 \rvert$
0 – 6	3	8	–12	–96	11	88
6 – 12	9	10	–6	–60	5	50
12 – 18	15	12	0	0	1	12
18 – 24	21	9	+ 6	54	7	63
24 – 30	27	5	+ 12	60	13	65
Total		44		–42		278

$$\therefore \quad \text{Mean} = a + \frac{\Sigma fd}{\Sigma f}$$

$$= 15 - \frac{42}{44}$$

$$= 14 \text{ (nearly)}$$

and average deviation

$$= \frac{\Sigma f \lvert x - \bar{x} \rvert}{\Sigma f}$$

$$= \frac{278}{44}$$

$$= 6{\cdot}3$$

9. We know,

$$\bar{x} = \frac{\Sigma fx}{\Sigma f}$$

Arithmetic mean

$$= \frac{258}{60}$$

$$= \text{Rs. } 4{\cdot}30$$

Arithmetic mean

$$= a + \frac{\Sigma fd}{\Sigma f}$$

$$= 5 + \frac{-42}{60} = 5 - \frac{7}{10}$$

$$= \text{Rs. } 4{\cdot}30$$

Here, $d = x - 5$

$\Sigma fd = -42$

10. Here number of children,

$n = 7$

and $\Sigma x = 462$

$\therefore$ Arithmetic mean

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

$$= \frac{462}{7}$$

$$= 66 \text{ cm}$$

Short cut method : Deviations from assumed mean 65 are

$-1, -6, 2, 4, 0, 5, 3$

$\therefore$ Arithmetic mean

$$= 65 + \frac{1}{n}\Sigma d$$

$$= 65 + \frac{1}{7}(7)$$

$$= 66 \text{ cm}$$

11. Since the regression lines pass through $(\bar{x}, \bar{y})$, therefore

$$4\bar{x} - 5\bar{y} + 33 = 0,$$

and $$20\bar{x} - 9\bar{y} = 107$$

Solving, we get $\bar{x} = 13$,

and $\bar{y} = 17$

Rewriting the line of regression of 4 on x,

$$y = \frac{4}{5}x + \frac{33}{5};$$

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

$$= \frac{4}{5}$$

Rewriting the line of regression of x and y as

$$x = \frac{9}{20}y + \frac{107}{9};$$

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

$$= \frac{9}{20}$$

$$\therefore \quad r^2 = \frac{4}{5} \times \frac{9}{20}$$

$$= 0{\cdot}36$$

or, $r = 0{\cdot}6$

12. Since the mean of x's and the mean of y's lie on the two regression lines, therefore

$$\bar{x} = 19{\cdot}13 - 0{\cdot}87\,\bar{y} \quad \ldots\text{(i)}$$

$$\bar{y} = 11{\cdot}64 - 0{\cdot}50\,\bar{x} \quad \ldots\text{(ii)}$$

Multiplying equation (ii) by 0·87 and subtracting from equation (i), we get

$$[1 - (0{\cdot}87)(0{\cdot}50)]\,\bar{x} = 19{\cdot}13 - (11{\cdot}64)(0{\cdot}87)$$

$$\therefore \quad \bar{x} = 15{\cdot}79$$

and $\quad \bar{y} = 11{\cdot}64 - (0{\cdot}50)(15{\cdot}79)$

$$= 3{\cdot}74$$

Thus regression co-efficient of y on x is $-0{\cdot}504$ and that of x on y is $-0{\cdot}87$.

Now since the co-efficient of correlation is the geometric mean between the two regression co-efficient, therefore

$$r = \sqrt{(-0{\cdot}50)(-0{\cdot}87)}$$

$$= \sqrt{0{\cdot}43}$$

$$= -0{\cdot}66$$

[**Note** : –ve sign is taken since both the regression co-efficient are –ve]

OBJECTIVE TYPE QUESTIONS

1. The mean and standard deviation of a binomial distribution are 10 and 2 respectively. The value of P is—
(A) 1·0 (B) 0·8
(C) 0·6 (D) 0·4

2. If the probability for A to fail in an examination is 0·2 and that for B is 0·3, then the probability that either A fails or B fails is—
(A) 0·5 (B) 0·44
(C) 0·06 (D) None of these

3. The probability that a teacher will give an unannounced test during any class is 1/5. If a student is absent twice then probability that misses atleast one test is—
(A) $\frac{2}{3}$ (B) $\frac{4}{5}$
(C) $\frac{7}{25}$ (D) $\frac{9}{25}$

4. The probability that atleast one of the events A and B occurs is 0·6. If A and B occur simultaneously with probability 0·2, then $P(\bar{A}) + P(\bar{B})$ is—
(A) 0·4 (B) 0·8
(C) 1·2 (D) 1·4

5. If M and N are two events, then probability that exactly one of them occurs is—
(A) $P(M) + P(N) + 2P(M \cap N)$
(B) $P(M) + P(N) - 2P(M \cap N)$
(C) $P(\overline{M}) + P(\overline{N}) + 2P(M \cap N)$
(D) $P(M \cap \overline{N}) + P(\overline{M} \cap N)$

6. If A and B are two mutually exclusive events, then P(A + B) is equal to—
(A) P(A) (B)
(B) P(A) + P(B)
(C) $P(A)\,P(B') + P(A')\,P(B)$
(D) $P(A)\,P(B') - P(A')\,P(B)$

7. The binomial distribution with mean 20 and standard deviation 4 is—
(A) $\left(\frac{1}{5}+\frac{4}{5}\right)^{100}$ (B) $\left(\frac{4}{5}+\frac{1}{5}\right)^{100}$
(C) $\left(\frac{4}{5}+\frac{1}{5}\right)^{50}$ (D) None of these

8. A purse contains 4 copper coins and 3 silver coins. The second purse contains 6 copper coins and 2 silver coins. A coin is taken out from any purse, the probability that it is a copper coin is—
(A) $\frac{11}{4}$ (B) $\frac{37}{36}$
(C) $\frac{3}{7}$ (D) $\frac{3}{4}$

9. In a Binomial distribution, if the mean is 9 and S.D. is $\sqrt{6}$, then values of n and p respectively are—
(A) $27, \frac{1}{3}$ (B) $81, \frac{1}{9}$
(C) $36, \frac{1}{4}$ (D) $18, \frac{1}{2}$

10. $P(A \cap B) = 0{\cdot}15$, $P(B') = 0{\cdot}10$, then P(A/B) is—

(A) $\frac{1}{3}$ (B) $\frac{1}{4}$

(C) $\frac{1}{5}$ (D) $\frac{1}{6}$

11. A and B toss 3 coins. The probability that they both obtain the same number of head is—

(A) $\frac{1}{9}$ (B) $\frac{3}{16}$

(C) $\frac{5}{16}$ (D) $\frac{3}{8}$

12. The probability of having atleast one tail in five throws with a coin is—

(A) $\frac{31}{32}$ (B) $\frac{1}{32}$

(C) $\frac{1}{5}$ (D) 1

13. India plays two P matches each with West Indies and Australia. In any match the probabilities of India getting points 0, 1 and 2 are 0·45, 0·05, and 0·5 respectively (Assuming that outcomes are independent) the probability of India getting atleast 7 points is—

(A) 0·8750 (B) 0·0875

(C) 0·0625 (D) 0·0250

14. Two events A and B have probabilities 0·25 and 0·50 respectively, the probability that both A and B occur simultaneously is 0·14. Then the probability that neither A nor B occur is—

(A) 0·39 (B) 0·375

(C) 0·89 (D) 0·86

15. If three disc are thrown simultaneously, then the probability of getting score of 5 is—

(A) $\frac{5}{216}$ (B) $\frac{1}{6}$

(C) $\frac{1}{36}$ (D) $\frac{1}{72}$

16. Let A and B be two independent events. The probability that both A and B occur is $\frac{1}{12}$ and the probability that neither A nor B occurs is $\frac{1}{2}$. The respective probabilities of A and B are—

(A) $\frac{1}{6}$ and $\frac{1}{2}$

(B) $\frac{1}{2}$ and $\frac{1}{6}$

(C) $\frac{1}{3}$ and $\frac{1}{4}$ or $\frac{1}{4}$ and $\frac{1}{3}$

(D) None of these

17. A box contains 100 tickets numbered 1, 2, 3, …… 100. Two tickets are chosen at random. If it is given that the maximum number on the two chosen tickets is not more than 10, then the probability that the minimum number on them is not less than 5 is—

(A) $\frac{1}{3}$ (B) $\frac{1}{2}$

(C) $\frac{152}{153}$ (D) None of these

18. A box contains 5 black and 5 red balls. Two balls are randomly picked one after another from the box, without replacement. The probability for both balls being red is—

(A) $\frac{1}{90}$ (B) $\frac{1}{5}$

(C) $\frac{19}{90}$ (D) $\frac{2}{9}$

19. Two dice are thrown. What is the probability that the sum of the numbers on the two dice is eight ?

(A) $\frac{5}{36}$ (B) $\frac{5}{18}$

(C) $\frac{1}{4}$ (D) $\frac{1}{3}$

20. An unbiased coin is tossed three times. The probability that the head turns up in exactly two cases is—

(A) $\frac{1}{9}$ (B) $\frac{1}{8}$

(C) $\frac{2}{3}$ (D) $\frac{3}{8}$

21. The probability that two friends share the same birth month is—

(A) $\frac{1}{6}$ (B) $\frac{1}{12}$

(C) $\frac{1}{144}$ (D) $\frac{1}{24}$

22. A box contains 5 black balls and 3 red balls. A total of three balls are picked from the box one after another, without replacing themback. The brobability of getting two black balls and one red ball is—

(A) $\frac{3}{8}$ (B) $\frac{2}{15}$

(C) $\frac{15}{28}$ (D) $\frac{1}{2}$

23. If 20 per cent managers are technocrats, the probability that a random committee of 5 managers consists of exactly 2 technocrats is—

(A) 0·2048 (B) 0·4000

(C) 0·4096 (D) 0·9421

24. A box contains three blue balls and four red balls. Another identical box contains two blue balls and five red balls. One ball is picked at random from one of the boxes and it is red. The probability that it came from the first box is—

(A) $\frac{2}{3}$ (B) $\frac{4}{9}$

(C) $\frac{4}{7}$ (D) $\frac{2}{7}$

25. The equilibrium data of component A in the two phases B and C are give below—

X (moles of A / moles of B)	Y (moles of A / moles of B)
1	0·5
2	4·125

The estimate of y for $x = 4$ by fitting a quadratic expression of a form $y = mx^2$ for the above data is—

(A) 15·5 (B) 16

(C) 16·5 (D) 17

26. Three identical dice are rolled the probability that the same number will appear on each of them is—

(A) $\frac{1}{6}$ (B) $\frac{1}{36}$

(C) $\frac{1}{18}$ (D) $\frac{3}{28}$

27. A box contains 6 red balls and 4 green balls, one ball is randomly picked and then a second ball is picked withut replacement of the first ball. The probability that both the balls are green is—

(A) $\frac{1}{15}$ (B) $\frac{2}{25}$

(C) $\frac{2}{15}$ (D) $\frac{4}{25}$

28. A pair of fair dice is rolled simultaneously. The probability that the sum of the numbers from the dice equals six is—

(A) $\frac{1}{6}$ (B) $\frac{7}{36}$

(C) $\frac{5}{36}$ (D) $\frac{1}{12}$

29. A box contains 8 balls, 2 of which are defective. The probability that none of the balls drawn are defective when two are drawn at random without replacement is—

(A) $\frac{15}{28}$ (B) $\frac{9}{16}$

(C) $\frac{7}{16}$ (D) $\frac{1}{8}$

30. Four fair coins are tossed simultaneously. The probability that at least one head turns up—

(A) $\frac{1}{16}$ (B) $\frac{15}{16}$

(C) $\frac{7}{8}$ (D) $\frac{1}{8}$

31. From a pack of regular playing cards, two cards are drawn at random. What is the probability that both cards will be kings, if the first card is not replaced ?

(A) $\frac{1}{26}$ (B) $\frac{1}{52}$

(C) $\frac{1}{169}$ (D) $\frac{1}{221}$

32. The following data about the flow of liquid was observed in a continuous chemical process plant—

Flow rate	Frequency
7·5 to 7·7	1
7·7 to 7·9	5
7·9 to 8·1	35
8·1 to 8·3	17
8·3 to 8·5	12
8·5 to 8·7	10

Mean flow rate of the liquid is—

(A) 8·00 litres / sec (B) 8·06 litres / sec
(C) 8·16 litres / sec (D) 8·26 litres / sec

33. A regression model is used to express a variable Y as a function of another variable X. This implies that—
(A) There is a casual relationship between Y and X
(B) A value of X may be used to estimate a value of Y
(C) Values of X exactly determine values of Y
(D) There is no casual relationship between Y and X

34. Manish has to travel from A to D changing bases at stops B and C enroute. The maximum waiting time at either stop can be 8 minutes each, but any time of waiting up to 8 minutes is equally likely at both places. He can afford up to 13 minutes of total waiting time if he is to arrive at D on time. What is the probability the Manish will arrive late at D ?
(A) $\frac{8}{13}$ (B) $\frac{13}{14}$
(C) $\frac{119}{128}$ (D) $\frac{9}{128}$

35. Arrivals at a telephone booth are considered to be Poisson, with an average time of 10 minutes between successive arrivals. The length of a phone call is distributed exponentially with mean 3 minutes. The probability that an arrival does not have to wait before service is—
(A) 0·3 (B) 0·5
(C) 0·7 (D) 0·9

Answer with Explanation

1. (C) Given $np = 10$

and $npq = 2^2$

$= 4$

$\Rightarrow \quad q = \frac{4}{10}$

$= \frac{2}{5}$

$\Rightarrow \quad p = 1 - q$

$= \frac{3}{5}$

2. (B) Required Probability

$= 1 -$ (neither A nor B fails)

$= 1 - p$(Both A & B are successful)

$= 1 - (1 - 0{\cdot}2)(1 - 0{\cdot}3)$

$= 0{\cdot}44$

3. (D) Required Probability

$= 1 - p$(student does not miss any test)

$= 1 - p$(the teacher does not given the test on both days)

$= 1 - \left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{5}\right)$

$= 1 - \frac{16}{25}$

$= \frac{9}{25}$

4. (C) $P(\overline{A}) + P(\overline{B})$

$= 2 - (P(A) + P(B))$

$= 2 - P(A \cup B) - P(A \cap B)$

$= 2 - 0{\cdot}6 - 0{\cdot}2$

$= 1{\cdot}2$

5. (D) Required Probability

$=$ P(only one M and N occurs)

$=$ P(M occurs but not N) + P(N occurs but not M)

$= P(M \cap \overline{N}) + P(\overline{M}) \cap N)$

(Note—that $M \cap \overline{N}$ and $N \cap \overline{M}$ *i.e.*, $M - N$ and $N - M$ are ME)

6. (B) Given $A \cap \phi = P(A \cap B)$

$= 0$

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

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

7. (B) Given Mean $= 20$

and S.D. $= 4$

$\therefore \quad np = 20$

and $npq = (\text{S.D.})^2$

$= 16$

From here, we get

$q = \frac{4}{5}$,

$p = \frac{1}{5}$

and $n = 100$

So the binomial distribution is ($p = q$)

$$n = \left(\frac{4}{5}+\frac{1}{5}\right)^{100}$$

8. (B)Required probability

$$= \frac{1}{2}\times\frac{4}{7}+\frac{1}{2}\times\frac{6}{8}$$

$$= \frac{2}{7}+\frac{3}{8}$$

$$= \frac{16+21}{56}$$

$$= \frac{37}{56}$$

9. (A) Given : $np = 9,$

and $\sqrt{npq} = \sqrt{6},$

$\Rightarrow$ $npq = 6$

$\Rightarrow$ $q = \frac{6}{9}$

$= \frac{2}{3}$

$\Rightarrow$ $p = 1-q$

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

$= \frac{1}{3}$

and $np = 9$

$\Rightarrow$ $n = \frac{9}{p}$

$= 27$

10. (D) $P(A/B) = \frac{P(A\cap B)}{P(B)}$

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

$$= \frac{0{\cdot}15}{1-0{\cdot}10}$$

$$= \frac{15}{90}$$

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

11. (C) Required Probability

$$= \frac{1}{8}\times\frac{1}{8}+\frac{3}{8}\times\frac{3}{8}+\frac{3}{8}\times\frac{3}{8}+\frac{1}{8}\times\frac{1}{8}$$

$$= \frac{1\times1+3\times3+3\times3+1\times1}{8\times8}$$

12. (A) It is a case of Bernoullian trials where success is "a trial comes up"

Here, $p = \frac{1}{2}$

and $q = \frac{1}{2}$

$\therefore$ Required Probability

$$= 1-p(0)$$

$$= 1-{}^5C_0\left(\frac{1}{2}\right)^5$$

13. (B) p(India gets 7 points)

= p(India wins 3 matches and drews exactly one of the four matches)

$$= {}^4C_3\,(0{\cdot}5)^3\,(0{\cdot}05)$$

$$= 4\times\left(\frac{1}{2}\right)^3\times\left(\frac{1}{20}\right)$$

$$= 0{\cdot}0250$$

p (India gets 8 points)

= p(India coins all the four matches)

$$= (0{\cdot}5)^4$$

$$= 0{\cdot}0625$$

$\therefore$ p(India gets atleast 7 points)

$$= 0{\cdot}0250 + 0{\cdot}0625$$

$$= 0{\cdot}0875$$

14. (A) p(neither A nor B)

$= 1 - P(\text{either A or B})$

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

$= 1 - \{(P(A) + P(B) - P(A\cap B)\}$

15. (C) Total number of equally likely events is 216 and favourable outcomes are :

(2, 2, 1), (2, 1, 2), (1, 2, 2), (3, 1, 1), (1, 3, 1), (1, 1, 3)

$\therefore$ Required Probability

$$= \frac{6}{216}$$

$$= \frac{1}{36}$$

16. (C) If $P(A) = x$

and $P(B) = y$

then $P(A\cap B) = \frac{1}{12}$

$\Rightarrow$ (neither A nor B occurs)

$= P(A')\,P(B')$

$= \frac{1}{2}$

$$\Rightarrow \quad P(A') \, P(B') = \frac{1}{2}$$

$$\text{or,} \quad (1-x)(1-y) = \frac{1}{2}$$

17. (A) Let A be the event "maximum number is not more than 10" and B. "minimum number is not less than 5"

$$\Rightarrow \quad P(A) = \frac{^{10}C_2}{^{100}C_2}$$

and $\quad P(A \cap B) = p$(number on selected ticketsis from 5 to 10)

$$= \frac{^{6}C_2}{^{100}C_2}$$

(since there are six tickets with such numbers)

∴ Required Probability

$$= P(B/A)$$

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

18. (D) The probability of drawing a red ball

$$= \frac{5}{10}$$

If the ball is not replaced, the box will have a ball,

So probability of drawing the red ball in next chance $= \frac{4}{9}$

Hence probability of drawing 2 balls

$$= \frac{5}{10} \times \frac{4}{9}$$

$$= \frac{2}{9}$$

19. (A) 20. (D)

21. (B) Probability that first friend is born in any month $= 100\%$

$$= 1$$

Probability that second friend is born in the same month as that of first friend

$$= 1 \times \frac{1}{12}$$

$$= \frac{1}{12}$$

22. (C) Let B refer to black balls and R to red balls consider the combinations BBR, BRB, RBB

$$\text{Probability of BBR} = \frac{5}{8} \times \frac{4}{7} \times \frac{3}{6}$$

$$= \frac{5}{28}$$

$$\text{Probability of BRB} = \frac{5}{8} \times \frac{3}{7} \times \frac{4}{6}$$

$$= \frac{5}{28}$$

$$\text{Probability of RBB} = \frac{3}{8} \times \frac{5}{7} \times \frac{4}{6}$$

$$= \frac{5}{28}$$

∴ Total probability of getting two black balls and one red ball $= \frac{5}{28} + \frac{5}{28} + \frac{5}{28}$

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

23. (A) Probability of technocrat manager

$$= \frac{20}{100} = \frac{1}{5} = p$$

Probability of non-tecnocrat manager

$$= \frac{80}{100} = \frac{4}{5} = q$$

Probability of a random committee of 5 with exactly 2 technocrats

$$= {}^{5}C_2 \, p^2 q^3$$

$$= \frac{5 \times 4}{1 \times 2} \left(\frac{1}{5}\right)^2 \left(\frac{4}{5}\right)^3$$

$$= 0{\cdot}2048$$

24. (D) Number of boxes

$$= 2$$

Probability of picking ball from one box

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

Total number of balls

$$= \text{Blue balls} + \text{Red balls}$$

$$= 3 + 4$$

$$= 7$$

Probability of red ball

$$= \frac{4}{7}$$

Thus probability of getting red ball from one box $= \frac{1}{2} \times \frac{4}{7}$

$$= \frac{2}{7}$$

25. (D)

26. (B) As (1, 1, 1), (2, 2, 2), (3, 3, 3), (4, 4, 4), (5, 5, 5), (6, 6, 6) are only favourable outcomes.

$\therefore$ Required probability

$$= \frac{6}{216}$$

27. (D) $\frac{{}^4C_1}{{}^{10}C_1} \times \frac{{}^3C_1}{{}^9C_1} = \frac{4}{10} \times \frac{3}{9}$

$$= \frac{2}{15}$$

28. (C) 29. (A) 30. (B)

31. (D) Probability of both cards being kings

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

$$= \frac{1}{221}$$

32. (C)

Flow rate (litre/sec)	Mean value of flow rate (x)	Frequency (f)	fx
7·5 – 7·7	7·6	1	7·8
7·7 – 7·9	7·8	5	36
7·9 – 8·1	8·0	35	280
8·1 – 8·3	8·2	17	139·4
8·3 – 8·5	8·4	12	100·8
8·5 – 8·7	8·6	10	86
		$\Sigma f = 80$	$\Sigma fx = 652{\cdot}8$

$\therefore$ Mean flow rate

$$= \frac{\Sigma fx}{\Sigma f}$$

$$= \frac{652{\cdot}8}{80}$$

$$= 8{\cdot}16$$

33. (B) 34. (A) 35. (A)