QUICK REVISION NOTES

CHAPTER-1

* Statistics : Latin word 'status'
  Italian word 'statista'
  German word 'statistik'
* Attribute : The qualitative information for a single object, single person/single commodity is called an attribute.
* Box Head : Entire upper part of the table.
* Caption : Upper part of the table whose describing column
* Stub : left part of the table whose describing rows.
* Continuous frequency distribution is called grouped frequency distribution.
* There is no difference between class limits and class boundaries in exclusive series.
* There are no class boundaries in inclusive series
* For less than : Adding preceeding frequencies
* For More than : Adding succeeding frequencies
* Histogram or Area Diagram : For equal or unequal C.I.
* Frequency Polygon : For equal C.I
* Ogives : For cumulative frequency distribution.
* Multiple line Chart : Relations between two or more than two variables with same units
* Multiple axis Chart : Relations between two or more than two variables with different units
* Vertical Bar Diagram : For quantitative information
* Horizontal Bar Diagram : For qualitative information
* Multiple or Grouped Bar chart : Comparision/relationship between different variables for a single unit.

* Divided Bar Chart : Comparision/relationship between different components for single variab

* There are four types of frequency curve:
  - Bell shaped curve or uni-modal curve
  - U-shaped curve or Bi-modal curve
  - J-shaped curve
  - Mixed curve

* In case of increasing profits → J-shaped curve

* In case of profits → Any of these/All of these/Bell-shaped curve

* In case of loss → None of these

* Central Angles = $\frac{\text{Actual Value} \times 360}{\text{Total value}}$

* logarithmic chart is also known as Ratio chart

* The difference between Upper class Boundary and lower class Boundary is known as class length.

* In 1985, out of total of 1570 workers of a factory, 1200 were members of a trade union. The number of women employee was 200, of which 175 did not belongs to a trade union. In 1990, 1800 employees who belonged to a trade union and 50 who did not belong to a trade union. of all the employees in 1990, 300 were women, of whom only 8 did not belong to a trade union. Present the following informatio in a suitable tabular form and also find the number of trade union women in 1985.

CHAPTER-2
CENTRAL TENDENCY

* Sum of the deviation from Arithmetic mean is equal to
* GM of the product of two variables is the product of their GM's.
* GM of the ratio of two variables is the ratio of their GM's
* Relations between AM, GM & HM:
  - AM = GM = HM (for equal observations)
  - AM > GM > HM (for unequal observations)
  - AM ≥ GM ≥ HM (without condition)
* $(GM)^2 = (AM).(HM)$
* AM: In case of direct observations or equal importance
* Weighted Mean: In case of unequal importance (weight)
* GM: In case of average by middle observations
* HM: In case of any condition (speed, Distance, Interest)
* Median can be calculated in unequal class intervals
* For open-end CI: Median
* In case of skewness distribution: Mean = Median = Mode
* In case of positively skewness distribution:
  Mean > Median > Mode
* In case of negatively skewness distribution
  Mean < Median < Mode
* Change of origin: Addition, subtraction, increasing, Decreasing
* Change of scale: Multiplication, Division
* AM, GM, HM, Median, Mode, Quartile, Decile, percentile is effected due to change of origin & change of scale.

An aeroplane flies from A to B at the rate of 500 km/h and comes back from B to A at the rate of 700 km/h. The average speed of the aeroplane is :

(a) 600 km/h
(b) 583.33 km/h
(c) 588.88 km/h
(d) 620 km/h

The average salary of a group of unskilled workers is ₹ 10000 and that of a group of skilled workers is ₹ 15000. If the combined salary is ₹ 12000, then what is the percentage of skilled workers :

(a) 40 %
(b) 50 %
(c) 60 %
(d) 80%

If there are two groups with 75 and 65 as harmonic means and containing 15 and 13 observation, then the combined HM is given by

(a) 65
(b) 70.36
(c) 70
(d) 71

Mean of 100 items is 49. It was discovered that three items which should have been 60, 70, 80 were wrongly read as 40, 20, 50 respectively. The correct mean is :

(a) 80
(b) 50
(c) 82.5
(d) 48

#

The average weight of students in a class of 35 students is 40 kg. If the weight of the teacher be included, the average rises by (1/2) kg, the weight of the teacher is :

(a) 40.5 kg
(b) 50 kg
(c) 41 kg
(d) 58 kg

The mean weight of a group of 10 items is 28 and that of another group of n items is 35. The mean of combined group of (10+n) items is found to be 30. The value of n is :

(a) 2
(b) 4
(c) 10
(d) 12

The arithmetic mean of a set of observations is $\bar{X}$. If each observation is divided by β and then it is increased by 12, then the mean of new series:

(a) $\frac{\bar{X}}{\beta}$ (b) $\frac{\bar{X}+12}{\beta}$

(c) $\frac{\bar{X}+12\beta}{\beta}$ (d) $\beta\bar{X}+12$

An auto ride in Delhi cost ₹1 for 1st km and 60 paise for each additional kilometre. The cost of each kilometre is incurred at the beginning of kilometre so that the rider pays for a whole kilometre. The average cost for 2.75 kilometre in ₹... is:

(a) 0.75 (b) 0.80

(c) 0.90 (d) 1.20

A cyclist covers first three kms at an average speed of 0 km/h. Another two kms at 3 km/h and the last two kms at 2 km/h. The average speed for the entire journey in km/h is:

(a) 2.43 (b) 3.43

(c) 4.43 (d) 5.43

The algebaric sum of the deviations of 20 observations measured from 30 is 2. Therefore, the mean of the observations is

(a) 30.7 (b) 30.1

(c) 29.7 (d) 29.1

CHAPTER-3
DISPERSION

* Variance is the square of the standard deviation
* Coeff. of variance = Coeff. of S.D × 100
* Use direct method of standard deviation in coeff. of variance
* Consistent $\propto \frac{1}{\text{Coeff. of variance}}$
* Variability $\propto$ Coeff. of variance
* QD : MD : SD = 10 : 12 : 15
* Change of origin : Addition, Subtraction, Increasing, Decreasing
* Change of scale : Multiplication, Division
* Standard deviation, Mean Deviation, Quartile Deviation, Range is not effected due to change of origin but is effected due to change of scale.
* Deviation always be positive
* Covariance may be negative, positive or zero
* Mean deviation is less than standard deviation
* Central Tendency : Including constant term, answer with positive sign.
* Dispersion : Excluding constant term, answer with '+ve' sign.
* Standard deviation of first 'n' natural numbers is : $\left(\frac{n^2-1}{6}\right)$
* The variance of first 'n' even natural numbers or, first 'n' odd natural numbers are : $\left(\frac{n^2-1}{3}\right)$.
* The variance of the A.P. $a, (a+d), (a+2d), \dots, (a+2nd)$ is $\left[\frac{n(n+1)}{3} d^2\right]$
* The mean deviation from the mean of the A.P. $a, (a+d), (a+2d), \dots, (a+2nd)$ is $\left[\frac{n(n+1)}{2n+1} d\right]$

Let s be the standard deviation of n observations. Each of the n observations is multiplied by a constant c. Then the standard deviation of the resulting number is

(a) s (b) $c.s$

(c) $s\sqrt{c}$ (d) c

If the standard deviation of $0, 1, 2, 3, \ldots, 9$ is K, then the standard deviation of $10, 11, 12, 13, \ldots, 19$ is

(a) $10K$ (b) $K+10$

(c) K (d) $K+\sqrt{10}$

The coeff. of variance of a series is 58. Its SD is 21.2. Its arithmetic mean is

(a) 36.6 (b) 22.6

(c) 26.6 (d) 36.1

If two variables x and y are related by $2x+3y-7=0$ and the mean and mean deviation about mean of x are 1 and 0.3 respectively, then the mean deviation of y about mean is

(a) -5 (b) 12

(c) SD (d) 4

If the mean and S.D. of x are a and b respectively, then the SD of $\frac{x-a}{b}$ is

(a) -1 (b) 1

(c) ab (d) a/b

For a set of 100 observations, taking assumed mean as 4, the sum of the deviation is (-11) cm and the sum of the squares of these deviation is 257 cm. The coeff. of variance is

(a) 41.13 % (b) 40.13 %

(c) 42.13 % (d) 40.87 %

CHAPTER-4
CORRELATION

* The quantitative relationship between two variables are called correlation.
* Correlation Coeff. lies between -1 & +1, which is denoted by 'r'
* Positive Correlation : Both the variables are increases or decreases with same direction.
* Perfect Positive Correlation : Both the variables are increases or decreases with same direction with same unit
* Negative Correlation : Both the variables are increases or decreases with opposite direction
* Perfect Negative Correlation : Both the variables are increase or decreases with opposite direction with same unit
* Ungrouped Data : Individual series
* Grouped Data : Continuous series
* Correlation can be obtained in ungrouped and grouped data both.
* Only Karl Pearson's (Assumed mean) method can be used in grouped data
* Product moment correlation is considered for both amount of correlation & nature of the correlation
* Spurious Correlation or Zero Correlation : No causal relation between two variables.
* Required condition for Rank correlation coeff., sum of the rank difference (ΣD) should be equal to zero (0).
* Correlation between two attributes : spearman's rank metho
* For finding correlation coeff., Only Karl Pearson's method (Assumed mean) can be used in Bivariate Frequency Distribution method.
* If the correlation coeff. (r) is equal to 0, then the variables are said to be independent.

* Correlation coeff. is independent due to change of origin and dependent due to change of scale.
* Coeff. of determination is the squares of the correlation coeff
* Coeff. of non-determination = 1 − Coeff. of determination
* If $r < 6PE$, then there is no significant and in this case correlation does not exist
* If $r > 6PE$, then there is a significant and in this case, correlation exist
* If the correlation coeff. between two variables is 1 ; then
$$y = a + bx \; ; \quad b > 0$$
* If the correlation coeff. between two variables is(-1); then
$$y = a + bx \; ; \quad b < 0$$
* If the correlation coeff. between two variables is (± 1), then
$$y = a + bx$$

If the covariance between two variables is 20 and the variance of one of the variable is 16, what would be the variance of the other variable

(a) less than 10 (b) More th 100
(c) More than 25 (d) less th 20

If the coeff. of correlation between two series x and y is 0.8 and their covariance is 20. If the variance of x series is 16. The standard deviation of y series is

(a) 6.25 (b) 39.06
(c) 31.25 (d) none of these

From the following data compute the correlation coeff. between x and y, given

: Arithmetic mean of x series is 25 and that of y = 18
: Sum of the products of deviations of x and y from their mean is 122
: Sum of the squares of deviation of x and y from their mean is 136 and 138
: No. of pairs of value = 15

(a) 0.98 (b) 0.89
(c) −0.89 (d) −0.98

The coeff. of concurrent deviation for p pair of observatio was found to be $\frac{1}{\sqrt{3}}$. If the number of concurrent deviation was found to be 6, then the value of p is :

(a) 8 (b) 7
(c) 10 (d) 9

If the relation between x and μ is $3x + 4\mu = -7$ and between y and v is $4v - 16y = -11$, $r_{xy} = -0.6$, then find $r_{\mu v}$

(a) 0.6 (b) −0.6
(c) 0.4 (d) −0.4

If $r = 0.6$ and $n = 64$, find out PE of the correlation coeff and determine the limits for the measures of population

(a) 0.053; 0.547 to 0.653 (b) 0.054; 0.846 to 0.946
(c) 0.034; 0.547 to 0.653 (d) None of these

A student calculates the value of r as 0.7 when the value of n is 5 and concludes that r is highly significant Is he correct

(a) Yes (b) No
(c) Not defined (d) none of these

CHAPTER-5
REGRESSION ANALYSIS

* The functional relationship between two variables are called regression equations and this process is known as regression analysis.
* b_{xy} & b_{yx} are called the regression coeff. of x on y and y on x respectively.
* Correlation coeff. is the geometric mean between regression coeff. i.e.,

$$r = \sqrt{b_{xy} \cdot b_{yx}}$$

* Both the regression coeffs (b_{xy} & b_{yx}) should have the same sign, i.e., either both are (+ve) or both are (-ve).
* If regression coeffs have negative sign, correlation coeff. will be negative and they have positive sign, correlation coeff. would be also positive.
* If one of the regression coeff is greater than unity then the other is less than unity (1).
* Arithmetic mean of the regression coeff. is greater than the geometric mean of the regression coeff. or the correlation coeff.

$$\left(\frac{b_{xy} + b_{yx}}{2}\right) > \sqrt{b_{xy} \cdot b_{yx}} \text{ or } r$$

Slope of the regression equation is the reciprocal form of the regression coeff. i.e.,

$$\text{Slope of x on y} = \frac{1}{b_{xy}}$$

* Regression equations intersect on their actual means
* If $r = 0$, then regression equations are perpendicular to each other.
* If $r = \pm 1$, then two regression lines coincide or identical.
* The difference between the observed value and the estimated value is called Error or residual value.

The likely production corresponding to a rainfall of 40cm from the following data

	Rainfall	Output (in quintals)
Average	30	50
S.D	5	10

$r = 0.8$

Find output.

(a) 66 Q (b) 56 Q
(c) 76 Q (d) 86 Q

For some data, the following results were obtained of the two variables x and y:

$\bar{x} = 53.2$; $\bar{y} = 27.9$; $b_{yx} = -1.5$; $b_{xy} = -0.2$.

The most probable value of y, when $x = 60$:

(a) 17.7 (b) 16.6
(c) 15.7 (d) 18.6

Given that the variance of $x = 9$ and the regression equations are $8x - 10y + 66 = 0$ and $40x - 18y = 214$.

Find: (i) mean value of x and y
(ii) coeff. of correlation
(iii) standard deviation of y

(a) (i) 13, 17 (ii) 0.8 (iii) 4
(b) (i) 17, 13 (ii) 0.6 (iii) 4
(c) (i) 13, 17 (ii) $\frac{6}{10}$ (iii) 4
(d) none of these

If $u = 2x + 5$ and $v = -3y - 6$ and regression coeff. of y on x is 2.4 what is the regression coeff. of v on u

(a) -3.6 (b) 3.6
(c) 2.4 (d) -2.4

If the regression line of y on x and x on y are given by $2x+3y+1=0$ and $5x+6y+1=0$, then arithmetic means of x and y

(a) 1, −1 (b) 1, 1

(c) −1, −1 (d) −1, 1

If the regression coeff. of y on x, the coeff of correlation between x and y and variance of y are $-\frac{3}{4}$, $-\frac{\sqrt{3}}{2}$ and 4 respectively. What is the variance of x

(a) $\frac{16}{3}$ (b) $\frac{4}{3}$

(c) 3 (d) $\frac{3}{4}$

Given the regression equation as $3x+y=13$ and $3x+4y=$ which one is the regression equation of x on y

(a) 1st equ. (b) 2nd equ.

(c) both (a) & (b) (d) none of these

If $2x+3y=8$ is the regression equation of y on x. Find the regression coeff. of y on x.

(a) $\frac{2}{3}$ (b) $-\frac{2}{3}$

(c) $\frac{3}{2}$ (d) $-\frac{3}{2}$

CHAPTER-6
PROBABILITY AND EXPECTED VALUE

* Probability: Chance of occurance of an event expressed numerically is called probability.
* Probability of an event which is certain to occur is 1 and the probability of an impossible event is 0.
* Experiment: Any operations that result in two or more than two outcomes is called an experiment.
* Events or Exhaustive Cases: One or more possible outcomes in an experiment is called an event.
* Equally likely: Events are said to be equally likely, if there is no reason to expect any one is preference to other.

  ually Exclusive: If in an experiment of an event rules out the happening of all others events in the same experiment, then these event are said to be mutually exclusive event
* All the mutually exclusive cases are independent events but all the independent events are not mutually exclusive.
* If the events are mutually exclusive then the two events cannot occur simultaneously in the same trials
* In case of mutually exclusive event: $P(A \cap B) = 0$
* In case of non-mutually exclusive event: $P(A \cap B) = P(A) \cdot P(B)$
* If A, B & C are mutually exclusive and exhaustive events, then $P(A) + P(B) + P(C) = 1$
* In case of independent event: $P(B/A) = P(B)$
* Event does not occur is called complementry event.
* (Event occur : Event does not occur) ⇒ odd in favour of
* (Event does not occur : Event occur) ⇒ odd against
* Single experiment and single event ⇒ Direct formula
* Single experiment and two or more than two events (Mutually exclusive or non-mutually exclusive) ⇒ Addition Theorem

* Two or more than two experiments and two or more than two events ⇒ Create sub cases with addition sign

* At least one ⇒ $\left[1-\left(\begin{array}{c}\text{Multiplication of the complementary of the}\\ \text{given events}\end{array}\right)\right]$

* for expected value:

2 Coins:

Events:	(2H or 0T)	(1H or 1T)	(0H or 2T)
	↓	↓	↓
Prob:	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

3 Coins:

Events:	(3H or 0T)	(2H or 1T)	(1H or 2T)	(0H or 3T)
	↓	↓	↓	↓
Prob:	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

1 Dice:

Events:	1	2	3	4	5	6
			↓			
Prob:			$\frac{1}{6}$			

Sum of 2 Dices:

Events:	2	3	4	5	6	7	8	9	10	11	12
	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
Prob:	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Find the chance of throwing at least one up in a single throw with three dices

(a) $\frac{91}{216}$ (b) $\frac{1}{216}$

(c) $\frac{125}{216}$ (d) $\frac{25}{216}$

The odd against a certain event are 5 to 3 and the odds in favour of another event, independent of the former are 7 to 5, find the chance that at least one of the event will happen.

(a) $\frac{70}{96}$ (b) $\frac{25}{96}$

(c) $\frac{71}{96}$ (d) $\frac{81}{96}$

Four persons are chosen at random from a group containing 3 men, 2 women and 4 children, the chance that exactly two of them will be children

(a) $\frac{11}{21}$ (b) $\frac{10}{21}$

(c) $\frac{9}{21}$ (d) $\frac{20}{21}$

A speak truth in 75% and B in 80% of the cases. In what percentage of cases are they likely to contradict each other narrating the same incidents

(a) 85% (b) 65%

(c) 35% (d) 75%

A piece of equipment will function only when all three parts A, B, C are working. The probability of part A failing during the first year is $1/6$, that of part B failing is $1/20$ and that of part C failing is $1/10$. What is the probability that the equipment will fail before the end of the year

(a) $\frac{53}{80}$ (b) $\frac{23}{80}$

(c) $\frac{57}{80}$ (d) $\frac{21}{80}$

The probability that a person A who is now 25 years old, lives for another 30 years is $2/5$ and the probability that a person B who is now 45 years old lives for another 30 years is $7/16$. Find the probability that at least one of those persons will be alive 30 years hence

(a) $\frac{27}{80}$ (b) $\frac{23}{80}$

(c) $\frac{53}{80}$ (d) $\frac{21}{80}$

Three horses A, B and C are in a race. A is twice likely to win as B and B is twice likely to win as C. What are the respective probabilities of winning

(a) $\frac{1}{2}, \frac{2}{7}, \frac{4}{7}$ (b) $\frac{3}{7}, \frac{4}{7}, \frac{1}{7}$

(c) $\frac{4}{7}, \frac{2}{7}, \frac{1}{7}$ (d) $\frac{1}{7}, \frac{2}{7}, \frac{4}{7}$

A player tosses three coins. He wins ₹ 5 if 3 heads appear, ₹3 if two heads appear, ₹1 if one head appear. On the other hand he losses ₹ 15 if 3 tails occur. Find the expected gain

(a) 1.25 (b) 0.25

(c) 0.75 (d) 3.25

In a business venture, a man can make a profit of ₹50000 or incur a loss of ₹ 20000 the probability of making profit or incurring loss, from the past experience, are known to be 0.75 and 0.25 respectively. What is his expected profit

(a) 35000 (b) 42000

(c) −42000 (d) 32500

A number is selected from the first 1000 natural number. What is the probability that it would be multiple of 5 or 9

(a) $\frac{711}{1000}$ (b) $\frac{311}{1000}$

(c) $\frac{289}{1000}$ (d) $\frac{279}{1000}$

A bag contains tickets numberd from 1 to 20. Two tickets are drawn, find the probability that both the numbers are prime

(a) $\frac{153}{190}$ (b) $\frac{28}{190}$

(c) $\frac{162}{190}$ (d) $\frac{128}{190}$

CHAPTER-7

THEORETICAL DISTRIBUTION

* Theoretical distribution is the frequency distribution in which frequency are obtained by mathematical computation.
* standard form of a Binomial distribution $(p+q)^n$
where, $p \rightarrow$ prob. of success, $q \rightarrow$ Prob. of failure, $n \rightarrow$ Total no. of tria
* Binomial distribution is biparametric distribution because it has two parameters n and p.
* In Binomial distribution, mean $(\mu) = np$ & variance $(\sigma^2) = npq$
* In Binomial distribution, mean is always greater than variance
* Binomial distribution may be Unimodal or Bimodal distributi
* Binomial distribution is symmetric, when $p = q = 0.5$
* Poisson distribution is uniparametric distribution b it has only one parameter mean.
* Poisson distribution may be Unimodal or Bimodal distribution
* In Poisson distribution, the probability of success is very close to zero.
* In Poisson distribution, mean and variance are equal. i.e.,

$$\text{mean}(m) = \text{variance}(\sigma^2) = np$$

* Binomial distribution and Poisson distribution both are discrete probability distribution.
* Normal distribution is also known as Z-dis.
* Normal distribution is continuous probability distribution.
* Normal distribution is biparametric distribution because it has two parameters mean and variance.
* ~~Normal distribution is unimodal distribution.~~
* In normal distribution, mean, median and mode are equal.
* In normal distribution, median is equidistant from lower and upper quantiles i.e., $\left(\frac{q_1 + q_3}{2}\right)$.

* In normal distribution, QD : MD : SD = 10 : 12 : 15
* The two tails of the normal curve extend indefinitely and never touch the horizontal line
* Normal curve is in bell-shaped. It shapes depend upon its parameters i.e., mean and variance.
* In normal distribution, curve is symmetrical.
* The total area of the normal curve is 1.
* Area between (-1 to 1) or $[(\mu-\sigma)$ to $(\mu+\sigma)] \rightarrow$ 68.27 % or 0.6827
* Area between (-2 to 2) or $[(\mu-2\sigma)$ to $(\mu+2\sigma)] \rightarrow$ 95.5 % or 0.955

  Area between (-3 to 3) or $[(\mu-3\sigma)$ to $(\mu+3\sigma)] \rightarrow$ 99.7 % or 0.997

  $(\mu-\sigma)$ and $(\mu+\sigma)$ are called Point of Inflexion

  $[(\mu-0.675\sigma)$ and $(\mu+0.675\sigma)]$ are called Quantiles $(Q_1 \& Q_3)$

  Mean deviation from mean $(MD_{\bar{x}})$ = Mean deviation from median (MD_{Me}) = Mean deviation from mode $(MD_{Mo}) = \frac{4}{5} SD$
* Probability density function: $\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}}$ $-\infty$ to ∞
* Probability density function is always greater than equal to 0

In a probability distribution, calculate mean, variance and standard deviation as a result of throwing a dice

(a) $\frac{7}{2}$, 2.92, 1.70 (b) $\frac{5}{2}$, 2.92, 1.70

(c) $\frac{7}{2}$, 2.92, 2.70 (d) none of these

In a Binomial distribution consisting of 5 independent trials, probability of 1 and 2 success are 0.4096 and 0.2048 respectively, find the parameter p of the distribution

(a) 1/3 (b) 1/4

(c) 1/5 (d) 1/6

A dice is tossed thrice. If getting at a four is considered a success, find the mean and variance of the probabilit (Binomial) distribution of the number of success

(a) $\frac{1}{2}, \frac{5}{12}$ (b) $\frac{2}{3}, \frac{5}{12}$

(c) $\frac{1}{2}, \frac{7}{12}$ (d) none of these

If the sum of mean and variance of a Binomial distribution is 4.8 for five trials, find the distribution.

(a) $\left(\frac{1}{4}+\frac{3}{4}\right)^5$ (b) $\left(\frac{1}{5}+\frac{4}{5}\right)^5$

(c) $\left(\frac{1}{3}+\frac{2}{3}\right)^5$ (d) none of these

The mortality rate for a certain dicease is 7 in 1000. What is the probability by Poisson distribution for just 2 deaths on account of this disease in a group of 400. (Given: $e^{-2.8} = 0.06$

(a) 0.06 (b) 23.52%

(c) 22.52 (d) none of these

It is given that 3% of the electric bulbs manufactured by a company are defective. Using Poisson distribution, find the probability that a sample of 100 bulbs will contain no defective bulb. (Given: $e^{-3} = 0.05$)

(a) 1.05 (b) 0.95

(c) 0.05 (d) 1.95

In a poisson distribution, $3P(X=2) = P(X=4)$, find $P(X=3)$.

(a) $6e^{-6}$ (b) $36e^{-6}$

(c) $36e^{6}$ (d) none of these

The scores made by candidate in a certain test are normally distributed with mean 500 and standard deviation 100. What percentage of candidate receives the scores between 400 & 600.

(a) 136.54 (b) 68.27

(c) 34.13 (d) 95.5

A certain type of wooden beam has a mean breaking strength of 1500 kgs. and a standard deviation of 100 kgs. Find the relative frequency of all such beams whose breaking strength between 1450 and 1600 kgs.
(Given: Area (-0.5 to 0) = 0.1915 & (0 to 1) = 0.3413)

(a) 1.5328
(b) 0.1498
(c) 0.5328
(d) none of these

If the two quantiles of a normal distribution are 47.30 and 52.70 respectively. What is the mode of the distribution.

(a) 50
(b) 100
(c) 25
(d) 75

Find the point of inflexion of the normal curve

$$f(x) = \frac{1}{4\sqrt{2\pi}} e^{\frac{-(x-10)^2}{32}} \quad \text{for } -\infty \text{ to } \infty$$

(a) 10 & 14
(b) 16 & 14
(c) 6 & 14
(d) 10 & 16

CHAPTER-8

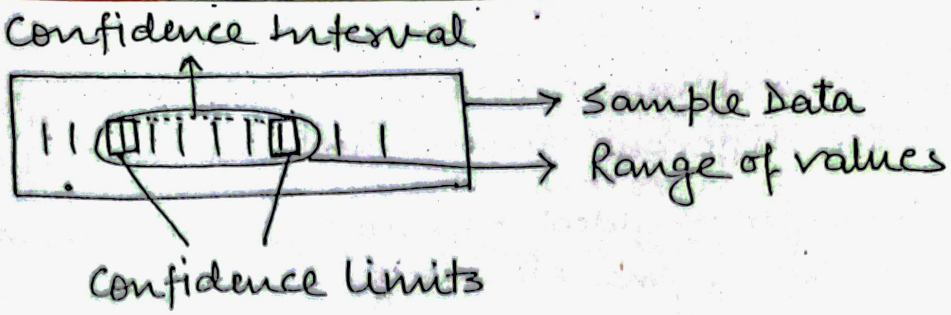
SAMPLING THEORY

* **Population**: Whole of the information which comes from statistical investigation is called population
* **Sample**: Selected part from the population is called sample
* **Parameter**: Any statistical measurement (mean, median, sd variance, etc) which comes from population data is called parameter.
* **Statistic**: Any statistical measurement (mean, media variance, etc) which comes from sample data called statistic.

Notation

	population	sample
Mean	μ	$\bar{x}$
SD	σ	s
Size	N	n
proportion	P	p

* **Estimation**: If the population data is defined by the sample data, then it is known an estimation.
* **Point Estimation**: A single statistic is used for making estimation of population parameter is called point estimation.
* **Interval Estimation**: Interval estimation is the range of values which is used for making estimation of population parameter.
* **Confidence limits**: Extreme values of the range of values are called confidence limits.
* Area between confidence limits is called confidence interval



* Confidence Interval based on confidence Level

Confidence Level (α)	Confidence Coeff. (Z_α)
90%	1.64
95%	1.96
98%	2.33
99%	2.58
100%	3

* $n > 30$; Large samples ; Z-test ; Z_α
  $n < 30$; small samples ; t-test ; t_α

* Population/sample proportion : It is the ratio of no. of elemen possessing the characteristics is to the total no. of elements in the population/sample.

* Degree of freedom (d.o.f.) = (size of samples) − 1
  or
  [(no. of rows − 1). (no. of columns − 1)]

* SRSWR : simple random sample with replacement (WR)
* SRSWOR : simple random sample without replacement (WOR)

* Sampling Distribution : with Replacement → N^n
  Without Replacement → ${}^N C_n$

* $\sqrt{\frac{N-n}{N-1}}$ is called finite population Multiplier or finite correction factor

* Statistical hypothesis is an assumption

* Stratified sampling provides separate estimates for populati means for different purposes and also an over all estimates
* The criteria for an ideal estimator is unbiasedness, consisten cy, sufficiency and efficiency.
* An unbiased estimator has an expected value equal to the true parameter value.
* Sample mean is the unbiased estimator of the population mean.
* A consistent estimator is said to be efficient estimator of a parameter, if its sampling variance is less than that of any other consistent estimator.
* Sample mean is the most efficient estimator of the populatic mean.
* A statistic which is unbiased and has also an minimum variance (most efficient) is called minimum variance unbiasec estimator (MVUE).
* Sample mean is the MVUE of the population mean.
* The standard deviation of the sampling distribution is known as standard error.
* A selection procedure of sample having no involvement of probability is known as purposive sampling or Judgement or subjective sampling.
* In systematic sampling selection of items is done at uni- form intervals of time.

If the population SD is known to be 5 for a population containing 80 units then the standard error of sample mean for a sample of size 25 without replacement is

(a) 5 (b) 0.20

(c) 1 (d) 0.83

A sample of size 3 is taken from a population of 10 members with replacement. If the sample observations are 1, 3 and 5, what is the estimate of the standard error of sample mean

(a) 1.96 (b) 1.15

(c) 1.25 (d) 2.28

A random sample of the heights of 100 students from a large population of students having SD as 0.35 m. shows an average height of 1.75 m.. What are the 95% confidence limits for the average height of all the students forming the populatio

(a) 1.68m., 1.82m. (b) 1.58m., 1.90m.

(c) 1.58m., 1.92m. (d) 1.50m., 2.0m.

If it is known that the 95% LCL & UCL to the population mean are 48.04 and 51.96 respectively, what is the value of the population variance when the sample size of 100.

(a) 8 (b) 100

(c) 12 (d) 12.50

The mean height obtained from a sample of size 100 taken randomly from a population is 164 cm. If the standard deviation of the height distribution of the population is 3 cm. set up 95% confidence limits for the mean height of the population

(a) 163.41, 164.59 (b) 153.41, 154.59

(c) 173.41, 174.59 (d) None of these

A random of 100 ball bearings selected from a large shipment of 2000 ball bearings has an average diameter of 0.354 cm. with a SD = 0.048 cm., find 95% confidence interval for the average diameter of these 2000 ball bearing

(a) 0.345, 0.363 (b) 0.445, 0.463

(c) 0.245, 0.263 (d) none of these

Out of consignment of 10000 tennis balls, 400 were selected at random and examined. It was found that 200 of these were defective. How many defective balls can you reasonably except to have in the whole consignment at 95% confidence level

(a) 4510, 5490 (b) 4765, 5035

(c) 4875, 5235 (d) none of these

A dice is thrown 9000 times and a throw of 3 or 4 obse 3240 times. Find the limits between which the probabilit a throw of 3 or 4 lies

(a) 0.345, 0.375 (b) 0.365, 0.395

(c) 0.360, 0.391 (d) 0.350, 0.401

It is known that population standard deviation is waiting time for new gas connection in a particular town is 20 days. How large sample should be chosen to the 95% confident that the average waiting time is within 3 days of true average

(a) 58 (b) 59

(c) 61 (d) 171

The incidence of a particular dicease in an area is such that 20% people of that area suffers from it. What size of sample should be taken so as to ensure that the error of estimation of the proportion should not be more than 5% with 95% confidence

(a) 246 (b) 236

(c) 226 (d) 286

A population consist of 12, 14 and 16. The possible number of size 2 which can be drawn without replacement are

(a) 3 (b) 6

(c) 9 (d) 12

A population consist of 1, 3, 5 and 7. The possible number of size 2 which can be drawn with replacement are

(a) 4 (b) 6

(c) 16 (d) 24

CHAPTER-9
INDEX NUMBERS

* Index number are devices for measuring differences in the magnitude of a group of related variables.
* 01 → Current year with respect to base year (Increasing differen
* 10 → Base year with respect to current year (Decreasing difference)
* Fisher's Index is the geometric mean between Laspeyre's Index and Paache's Index
* Bowley's Index is the arithmetic mean between Laspeyre's Inde and Paache's Index
* Bowley's Index is greater than Fisher's Index. i.e.,

$$\left(\frac{L+P}{2}\right) > \sqrt{L \times P}$$

* If do not given (Price Index or Quantity Index), then except Relatives Method, for all the Index use 'Price Index' and for Relatives Method use 'Quantity Index.

here are four types of Test of Adequacy
- Unit test
- Time Reversal test
- Factor Reversal test
- Circular Test

* All the index formulae satisfy unit test except simple aggregative method.
* Required condition for Time Reversal test ($P_{01} \times P_{10} = 1$); satisfy:
  - simple aggregative method
  - Fisher's Index
  - Marshall Edgeworth's Index
  - Kelly's Index
  - Simple geometric mean of price relatives method
  - Weighted geometric mean of price relatives method
* Required condition for Factor Reversal test ($P_{01} \times Q_{01} = V_{01}$)
  Satisfy: Fisher's Index

* Required condition for Circular test : $P_{01} \times P_{12} \times P_{20} = 1$
satisfy : simple geometric mean of Price relatives method

* Fisher's method is called best method.

* Fixed base Index = $\frac{\text{Current year Price Index} \times 100}{\text{Base (Fixed) year Price Index}}$

* Link Relatives = $\frac{\text{Current year Price Index} \times 100}{\text{Previous year Price Index}}$

* Chain Base Index = $\frac{\text{Link Relatives of the current year} \times \text{chain Base Index of the previous year}}{100}$

* Fixed Base Index = $\frac{\text{chain Base Index of the current year} \times \text{Fixed Base Index of the previous year}}{100}$

* Base Shifting = $\frac{\text{Current year Price Index} \times 100}{\text{Price Index of the year in which is to be shifted}}$

* Real Wages = $\frac{\text{Money Wages}}{\text{Price Index}} \times 100$

* Purchasing power of money is the reciprocal form of Price Index.

A textile worker in the city of Mumbai earns ₹ 3500 per month. The cost of living index for a particular month is 136.

Group	Expenditure	Group Index
Food	1400	180
Clothing	?	150
House Rent	?	100
Fuel & Lighting	560	110
Miscellaneous	630	80

The amount spent on house rent and clothing are

(a) 490, 420 (b) 420, 490

(c) 510, 420 (d) 490, 510

A worker earned ₹ 900 per month in 1990. The cost of living index increased by 70% between 1990 and 1993. How much extra income should the worker have earned in 1993, so that he could buy the same quantities as in 1990

(a) ₹ 7460 (b) ₹ 9460

(c) ₹ 7560 (d) ₹ 8464

During a certain period the cost of living index number goes up from 110 to 200 and the salary of the worker is also raised from ₹ 325 to ₹ 500. Does the worker

(a) gain (b) losses

(c) fully compensated (d) gain by 10%

Suppose a business executive was earning ₹ 2050 in the base period, what should be his salary in the current period, if his standard of living is to remain the same.

Given: $\Sigma W = 25$, $\Sigma IW = 3544$.

(a) ₹ 2906 (b) ₹ 2609

(c) ₹ 2806 (d) ₹ 2706