MODEL QUESTION

Statistics (Theory)

(New Syllabus)

GROUP – A

1. i) Write down the result of $\frac{d}{dx}\left(1-e^{-\theta x}\right)$

OR, If M is the mode of a discrete probability distribution with mass function f then f(x) = ………….. at x = M. (Fill in the blank)

Ans: $\frac{d}{dx}(1-e^{-\theta x}) = \theta e^{-\theta x}$

OR, f(x) = maximum value at x = M.

ii) If two variables are proportional then they do not have any correlation. (Write true or false)

Ans: False.

iii) In case of ranking some individuals with respect to a characteristic tied ranks may occur. (Write true or false)

OR, Two ranks of an individual must be equal when some individuals are ranked with respect to two attributes. (Write true or false)

Ans: True

OR, False

iv) To study the trend of a time–series one needs to take

(a) only one observation (b) two observations

(c) only daily observations (d) observations for a sufficiently long time

(Write the correct answer)

Ans: (d) observations for a sufficiently long time.

v) A random variable X is such that P(X = 1) = P(X = 0) = $\frac{1}{2}$. What is the value of P(X > 0) ?

Ans: P(X > 0) = P(X = 1) = $\frac{1}{2}$

vi) If a discrete variable assumes values 5, 6, 7, 8, 9 and 10 with equal probabilities then its probability mass function is given by

$$f(x) = \begin{cases} \text{............, if, } x = 5,6,7,8,9,10 \\ 0, \text{otherwise} \end{cases} \quad \text{(Fill in the blank)}$$

OR, In case of a continuous random variable X

P(X = 2.22) = …………………. (Fill in the blank)

Ans: $$f(x) = \begin{cases} \frac{1}{6} \text{ if, } x = 5,6,7,8,9,10 \\ 0, \text{otherwise} \end{cases}$$

OR, P(X = 2.22) = 0

vii) For a binomial variable X

E(X) < Var (X). (Write true or false)

Ans: False (E(X) > V(X) for Bin.)

viii) If the parameter of a Poisson distribution is 7.83 then its mode is …………… (Fill in the blank)

Ans: Mode = [7.83] = 7

ix) If a digit is taken from a random number series, the probability that it is 9 and the probability that it is not 9 are not equal. (Write true or false)

OR, Can any digit repeat in a random number series ?

Ans: True.

OR, yes, any digit can repeat.

x) If the observed values of a sample are given, the value of a statistic becomes fixed. (Write true or false)

Ans: True.

GROUP – B

2. i) When only two attributes A and B having two forms each are considered, how are the marginal frequencies and the total frequency related ?

Ans: Sum of the two marginal frequencies for each of two attributes A and B = Total frequency.

ii) Define bivariate data with a suitable example.

OR, When are two variables said to be positively correlated ?

Ans: See, "Giri or Roychowdhury".

OR, When the values of one of the variables increase (or decrease) with the increase (or decrease) in the values of the other variable, then the two variables are positively correlated.

iii) Show the case when the value of Spearman's rank correlation coefficient will be +1.

OR, What do you mean by rank correlation ?

Ans: Spearman rank correlation coefficient is given by the formula $r_s = 1 - \frac{\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$

for n individuals between two sets of ranks. And hence r_s = +1 when d_i = 0 for all i = $1(1)^n$ i.e. $Q_i = R_i$ i.e. the two ranks are equal for each individual.

OR, See "Giri or Roychowdhury".

iv) If for a binomial distribution m = 10 and p = $\frac{5}{13}$ then what is the value of the coefficient of variation ?

OR, If for a binomial distribution the expected value is 4 and variance is 3, what are the values of its parameters ?

Ans: C.V of B(10, $\left(10, \frac{5}{13}\right) = \frac{6}{\mu} \times 100\% = \frac{\sqrt{mpq}}{mp} \times 100\%$

$$= \frac{\sqrt{10 \times \frac{5}{13} \times \frac{8}{13}}}{10 \times \frac{5}{13}} \times 100\%$$

$$= \frac{20/13}{50/13} \text{x} 100\% = 40\%$$

OR, Here mp = 4 & mpq = 3 $\quad \therefore$ q = $\frac{3}{4}$

$\therefore$ p = $1 - \frac{3}{4} = \frac{1}{4}$ $\qquad \therefore \; m = \frac{4}{p} = \frac{4}{1/4} = 16$

$\therefore$ Parameters are m = 16 & p = $\frac{1}{4}$

v) Show that a normal distribution is symmetric about its mean.

Ans: The p.d.f. of Normal distribution with parators μ & σ^2 is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \; ; \qquad -\infty < x < \infty$$

$= 0$, otherwise.

Now, $f(\mu + x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}x^2/\sigma^2}$

& $f(\mu - x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}x^2/\sigma^2}$

$\therefore f(\mu + x) = f(\mu - x)$

Hence the distribution is symmetric about μ, where μ is the mean of the distribution.

vi) State the meaning of a 'Population' in Statistics.

Ans: See "Giri or Roychowdhury"

GROUP – C

3. i) Determine the value of $\int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx$

[Given that $r\left(\frac{1}{2}\right) = \sqrt{\pi}$]

OR, Find the first order raw moment of the probability distribution having density function, given by

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, \text{if } x > 0, \theta > 0 \\ 0, \text{otherwise} \end{cases}$$

Ans: $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$

$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$ putting $\frac{x-\mu}{\sigma} = z$

$= 2\int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{--\frac{z^2}{2}} dz$; since $e^{-\frac{z^2}{2}}$ is an even function.

$= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-t} \frac{dt}{\sqrt{2t}}$ [putting $\frac{z^2}{2} = t$

$$zdz = dt; \quad dz = \frac{dt}{\sqrt{2t}}.$$

$$= \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-t}\, t^{\frac{1}{2}-1} dt$$

$$= \frac{1}{\sqrt{\pi}} . \sqrt{\frac{1}{2}} = \frac{\sqrt{\pi}}{\sqrt{\pi}} = 1.$$

OR, First order raw moment is given by

$$\mu_1^{/} = E(X) = \int_0^\infty x\, f(x)\, dx$$

$$= \int_0^\infty x . \frac{1}{\theta} e^{-x/\theta} dx$$

$$= \frac{1}{\theta} \int_0^\infty \theta t .\, e^{-t} . \theta dt \qquad x/\theta = t; \quad dx = \theta\, dt$$

$$= \theta . \int_0^\infty e^{-t} .\, t^{2-1} dt$$

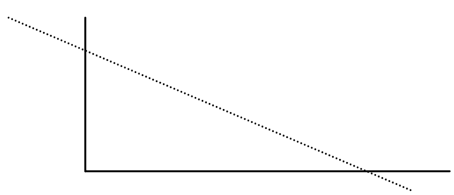
$$= \theta . \sqrt{2} = \theta .\ 1 = \theta.$$

ii) In the 2 x 2 ease of two attributes A and B discuss the cases of association.

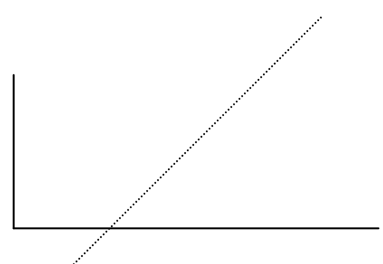
OR, Draw scatter diagrams in cases of perfect correlation between two variables.

Ans: See "Giri or Roychowdhury"

OR,



Perfect positive correlation



Perfect negative association

iii) If the bases and scales of two correlated variables are changed, determine the effect on correlation coefficient.

OR, If two variables x_1 and x_2 have a common variance and correlation coefficient

$r \neq 0$, express r in terms of θ so that $x_1 + 2x_2$ and $x_1 + \theta x_2$ $(|\theta| \le 1)$ become uncorrelated.

Ans: Consider the pair of values $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$ on the variables (x, y), and $u_i = \frac{x_i - a}{b}$ & $v_i = \frac{y_i - c}{d}$; $i = 1(1)^n$; u & v be new variables shifting base & scale. Then the correlation coefficient between x & y is given by

$$\gamma_{xy} = \frac{Cov(x,y)}{s_x\, s_y}, \text{ where } Cov(x, y) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x}) \times (y_i - \bar{y})$$

$$s_x = s.d(x) = \sqrt{\frac{1}{n}\sum (x_i - \bar{x})^2} \quad \&$$

$$s_y = s.d(y) = \sqrt{\frac{1}{n}\sum (y_i - \bar{y})^2}$$

Now, $Cov(u, v) = \frac{1}{n}\sum_{i=1}^{n} (u_i - \bar{u})(v_i - \bar{v})$

$$= \frac{1}{n}\sum_{i=1}^{n} \left(\frac{x_i - a}{b} - \frac{\bar{x} - a}{b}\right)\left(\frac{y_i - c}{d} - \frac{\bar{y} - c}{d}\right)$$

$$= \frac{1}{n}\frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{b.d} = \frac{Cov(x,y)}{b.d}$$

$$s_u = \sqrt{\frac{1}{n}\Sigma(u_i - \bar{u})^2} = \frac{1}{|b|}\sqrt{\frac{1}{n}\Sigma(x_i - \bar{x})^2} = \frac{s_x}{|b|}$$

and $$s_v = \sqrt{\frac{1}{n}\Sigma(v_i - \bar{v})^2} = \frac{1}{|d|}\sqrt{\frac{1}{n}\Sigma(y_i - \bar{y})^2} = \frac{s_y}{|d|}$$

Hence the correlation coefficient between u & v is given by

$$\gamma_{uv} = \frac{Cov(u,v)}{s_u.s_v} = \frac{Cov(x,y)}{b.d} \times \frac{|b|.|d|}{s_x.s_y}$$

$$= \frac{|b|.|d|}{b.d} = \pm \gamma_{xy}$$

Hence numerical value of the correlation coefficient remains unaltered with the change of base & scale.

OR,

Let $v(x_1) = v(x_2) = \sigma^2$

$\therefore r = \frac{Cov(x_1, x_2)}{\sigma.\sigma}$ or $Cov(x_1, x_2) = r.\sigma^2$

Now, $x_1 + 2x_2$ and $x_1 + \theta x_2$ are uncorrelated.

$\therefore$ Cov(x_1 + $2x_2$, x_1 + θx_2) = 0

or, Cov(x_1, x_1) + Cov(x_1, θx_2) + Cov($2x_2$, x_1) + Cov($2x_2$, θx_2) = 0

or, $\sqrt{\sigma\left(x_1^2\right)}$ + θ Cov(x_1, x_2) + 2 Cov(x_2, x_1) + 2θ v(x_2) = 0

or, $\sigma^2 + \theta.r\,\sigma^2 + 2.r\,\sigma^2 + 2\theta\,\sigma^2 = 0$

or, $1 + \theta r + 2r + 2\theta = 0$

or, $r(\theta + 2) = -(1 + 2\theta)$

$\therefore r = -\dfrac{1+2\theta}{\theta+2}$

iv) Derive the two normal equations in case of obtaining a linear regression line on the basis of bivariate data.

OR, If the normal equations are given for obtaining a linear regression line then determine the regression line.

Ans: See “Giri or Roychowdhury”.

OR, See “Giri or Roychowdhury”.

v) Describe the method of trend determination by the moving average method when the period–length is an even integer.

OR, State the disadvantages of determining trend by the method of moving averages.

Ans: See “Giri or Roychowdhury”.

OR, See “Giri or Roychowdhury”.

vi) Give the definition of a probability mass function of a random variable X and express $P(X = x \mid X \leq x_0)$ by using such function.

OR, If an unbiased coin is tossed thrice, find the probability distribution of the number of tails.

Ans: For a discrete random variable X, if there exists a function f(x) such that P(X = x) = f(x) satisfying the conditions – (i) $f(x) \geq 0\ \forall x$ and (ii) $\sum_{x} f(x) = 1$, then f(x) is called the probability mass function (p.m.f.) of X.

$P(X = x \mid X \leq x_0)$

$$= \frac{P(X = x, X \leq x_0)}{P(X \leq x_0)} = \frac{P(X = x)}{P(X \leq x_0)} \quad \text{if } x \leq x_0$$

$$= \frac{f(x)}{\sum_{x \leq x_0} f(x)} \quad \&\ \ 0, \text{ otherwise.}$$

OR, The sample space is

S = { HHH, THH, HTH, HHT, TTT, HTT, THT, TTH }

Let X be the r.v denoting the no. of tails.

Then the possible values of X are 0, 1, 2,3 and the probability distribution of X is given in the Table.

Values of X = x	Prob. P(X = x)
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$
Total	1

vii) Determine the median of a uniform (α, β) distribution.

OR, Determine the variance of a uniform (α, β) distribution.

Ans: The p.d.f. of the distribution is

$$f(x) = \frac{1}{p-\alpha}, \quad \infty < x < \beta$$

$$= 0, \text{ othewise.}$$

If μ_e be the median. Then

$$P(X \leq \mu_e) = P(X > \mu_e) = \frac{1}{2}$$

$$\therefore P(X \leq \mu_e) = \frac{1}{2}$$

$$\text{or, } \int_{\alpha}^{\mu_e} f(x)dx = \frac{1}{2} \quad \text{or, } \int_{\alpha}^{\mu_e} \frac{1}{\beta.\alpha}dx = \frac{1}{2}$$

$$\text{or, } \frac{1}{\beta-\alpha}(\mu_e - \alpha) = \frac{1}{2}$$

$$\therefore \text{Median, } \mu_e = \frac{\beta-\alpha}{2} + \alpha = \frac{\alpha+\beta}{2}.$$

$$\text{OR,} \quad E(X) = \int_{\alpha}^{p} x\, f(x)\, dx = \int_{\alpha}^{p} x \frac{1}{\beta-\alpha} dx$$

$$= \frac{1}{2}.\frac{p^2-\alpha^2}{p-\alpha} = \frac{\alpha+p}{2}$$

$$E(X^2) = \int_{\alpha}^{\beta} x^2.\frac{1}{\beta-\alpha}dx = \frac{\beta^3-\alpha^3}{3(\beta-\alpha)} = \frac{\alpha^2+\alpha\beta+\beta^2}{3}$$

Hence variance of X is

$$V(X) = E(X^2) - E^2(X)$$

$$= \frac{\alpha^2+\alpha\beta+\beta^2}{3} - \left(\frac{\alpha+\beta}{2}\right)^2$$

$$= \frac{4\alpha^2+4\alpha\beta+4\beta^2-3\alpha^2-3\beta^2-6\alpha\beta}{12}$$

$$= \frac{\alpha^2+\beta^2-2\alpha\beta}{12} = \frac{(\alpha-\beta)^2}{12}.$$

viii) If a Poisson variable X is such that 2P = (X = 5) = P(X = 6) then what is the value of P (X > 0) ?

OR, If a binomial (10, p) variable X is such that P(X = 5) = P(X = 6) then what is the value of p ?

Ans: We assume that $X \sim P(\lambda)$ $\therefore$ p.m.f.; $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$; $x = 0(1)\infty$

$\therefore$ 2P(X = 5) = P(X = 6)

$\Rightarrow 2.\frac{e^{-\lambda}\lambda^5}{5!} = \frac{e^{-\lambda}\lambda^6}{6!}$ $\therefore \lambda = 12$.

$\therefore$ P(X > 0) = 1 – P(X = 0)

$= 1 - e^{-\lambda} = 1 - e^{-12} = ..$

OR,

$X \sim B(10, p)$ $\therefore$ p.m.f. is $f(x) = \binom{10}{x} p^x q^{10-x}$; $x = 0(1)|0$

$\therefore$ P(X = 5) = P(X = 6)

$\Rightarrow \binom{10}{5} p^5 q^5 = \binom{10}{6} p^6 q^4$

or, $\frac{q}{p} = \frac{5!\ 5!}{6!\ 6!} = \frac{5}{6}$

or, $\frac{1}{p} - 1 = 5/6$ $\therefore$ $\frac{1}{p} = 11/6$ $\therefore p = \frac{6}{11}$

ix) In case of fitting a binomial distribution to a frequency distribution of a variable X, how will you estimate p and use it to find out an estimate of P(X = x) ?

Ans: Using the observed frequency distribution of r.v. X we can find the value of A.M. as $\bar{x} = \frac{\sum_i x_i f_i}{\sum_i f_i}$. Then we estimate the unknown p by the method of moments as

$E(X) \hat{=} \bar{x}$ where X ~ B (n, p)

i.e. np = $\bar{x}$

Hence moment estimate of p is given by $\hat{p} = \frac{\bar{x}}{n}$, where $\bar{x} = \text{A.M.}$

Now, p.m.f. of X is f(x) = P(X = x) = $\binom{n}{x} p^x \; q^{n-x}$; x = 0(1)n

$$\therefore \frac{f(x)}{f(x-1)} = \frac{n-x+1}{x}.\frac{p}{q}; \quad x \neq 0.$$

So using the estimate of p as $\hat{p}$ we can find.

$$P(X = x) = \frac{n-x+1}{x}.\frac{\hat{p}}{1-\hat{p}} \times f(x-1); \;\; x \neq 0.$$

Using P(X = 0) = f(0) = $(1 - \hat{p})^n$.

x) Describe the process of selecting 3 boys from a group of 10 boys by using a random number series.

OR, "In a sample survey more accurate result is obtained and sampling error may be gauged". Explain.

Ans: See "Giri or Roychowdhury".

xi) Give the definition of minimum variance unbiased estimator. Show that if f be the number of successes out of n Bernoulli trials with success probability p then $\frac{f}{n}$ is unbiased for p.

Ans: **M V U E :** An unbiased estimator T of a parameter θ is said to be MVUE if T has the minimum variance among all other unbiased estimators of θ.

Here f ~ B(n, p). Hence E(f) = np.

$$\therefore E\left(\frac{f}{n}\right) = p \quad \therefore \frac{f}{n} \text{ is an u.e of p.}$$

GROUP – D

4. i) For a Poisson (λ) distribution the 2^{nd}, 3^{rd} and 4^{th} central moments are $\mu_2 = \lambda$, $\mu_3 = \lambda$ and $\mu_4 = 3\lambda^2 + \lambda$ respectively. Use these to measure skewness and kurtosis of the distribution. Hence comment on the skewness and kurtosis of the distribution.

OR, Given that for a normal (μ, σ^2) distribution the mean is 65.5 inches and P(Z > 60.5) 0.9. Find the interval ($\mu - 3\sigma$, $\mu + 3\sigma$)

[Given that for the standard normal variable Z, P (A < 1.28) = 0.9]

Ans: The measure of skewness is given by [$s = \sqrt{m_2}$]

$g_1 = \frac{\mu_3}{\mu_2^{3/2}}$ & that of kurtosis is given by

$$g_2 = \frac{\mu_4}{\mu_2^2} - 3.$$

Here $g_1 = \frac{\lambda}{\lambda^{3/2}} = \frac{1}{\sqrt{\lambda}} > 0$ since $\lambda > 0$ for any Poisson distribution.

Hence Poisson distribution is positively skewed.

Now, $g_2 = \frac{3\lambda^2 + \lambda}{\lambda^2} - 3 = \frac{1}{\lambda} > 0.$ $\quad \therefore \lambda > 0.$

Hence Poisson distribution is leptokurtic.

OR,

Assume that X ~ N(μ, σ^2), where $\mu = 65.5$.

Here P(X > 60.5) = 0.9

or $P\left(\frac{X - 65.5}{\sigma} > \frac{60.5 - 65.5}{\sigma}\right) = 0.9$

or $1 - P(Z \leq -5/\sigma) = 0.9$

or $1 - \Phi(-5/\sigma) = 0.9$

or $\Phi(5/\sigma) = 0.9 = \Phi(1.28)$

$\therefore \frac{5}{\sigma} = 1.28 \qquad \therefore \sigma = \frac{5}{1.28}$

$\therefore 3\sigma = \frac{15}{1.28} = 11.72$

$\therefore (\mu - 3\sigma, \ \mu + 3\sigma) = (53.78, 77.22)$

ii) If there independent estimators T_1, T_2 and T_3 be unbiased for a parameter 0 and their variances are in the ratio 2 : 1 : 3 then which one of the following estimators would you prefer most for θ and why ?

$\frac{T_1 + T_2 + 2T_3}{4}, \frac{T_1 + T_2 + T_3}{3}, \frac{2T_1 + T_2}{3}.$

OR, On the basis of a random sample (x_1, x_2, ….,x_n) drawn without replacement from a finite population with variance σ^2 find the expectation of the sample variance and hence suggest an unbiased estimator for σ^2.

Ans: Let $V(T_1) = 2\sigma^2$, $V(T_2) = \sigma^2$, $V(T_3) = 3\sigma^2$

$$E\left(\frac{T_1 + T_2 + 2T_3}{4}\right) = \frac{E(T_1) + E(T_2) + 2E(T_3)}{4} = \frac{\theta + \theta + 2\theta}{4} = \theta$$

$$E\left(\frac{T_1 + T_2 + T_3}{3}\right) = \frac{E(T_1) + E(T_2) + E(T_3)}{3} = \frac{\theta + \theta + \theta}{3} = \theta$$

$$E\left(\frac{2T_1 + T_2}{3}\right) = \frac{2E(T_1) + E(T_2)}{3} = \frac{2\theta + \theta}{3} = \theta$$

$$V\left(\frac{T_1 + T_2 + 2T_3}{4}\right) = \frac{V(T_1) + V(T_2) + 4V(T_3)}{16} \quad \therefore T_1, T_2, T_3 \text{ are independent.}$$

$$= \frac{2\sigma^2 + \sigma^2 + 4 \times 3\sigma^2}{16} = \frac{15}{16}\sigma^2$$

$$V\left(\frac{T_1 + T_2 + T_3}{3}\right) = \frac{V(T_1) + V(T_2) + V(T_3)}{9} = \frac{2\sigma^2 + \sigma^2 + 3\sigma^2}{9}; \quad \therefore T_1, T_2, T_3 \text{ are independent.}$$

$$= \frac{6}{9}\sigma^2 = \frac{2}{3}\sigma^2$$

$$V\left(\frac{2T_1 + T_2}{3}\right) = \frac{4V(T_1) + V(T_2)}{9}; \quad T_1, T_2, T_3 \text{ are independent.}$$

$$= \frac{4 \times 2\sigma^2 + \sigma^2}{9} = \frac{9}{9}\sigma^2 = \sigma^2.$$

Here $V\left(\frac{T_1 + T_2 + T_3}{3}\right)$ is minimum and hence we prefer $\frac{T_1 + T_2 + T_3}{3}$ most.

OR, See "Giri or Roychowdhury".