①

$P(x) = P(X=x) = f(x)$

Module – I

Discrete P.D (or) Pmf — Probability mass f^n

1) $f(x) \geq 0$
2) $\sum f(x) = 1$ $\downarrow$ $k = ?$

Continuous P.D (or) Pdf — Probability density f^n

1) $f(x) \geq 0$
2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$ $\downarrow$ $k = ?$

Discrete Pmf

1) Mean $(\mu) = E(x) = \sum x\, f(x)$
2) Variance $(\sigma^2) = \sum x^2 f(x) - \mu^2$
3) S.D $(\sigma) = \sqrt{\sigma^2}$

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

Continuous Pdf

1) mean $(\mu) = E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$
2) Variance $(\sigma^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$

Mathematical Expectation

$$\left\{\begin{array}{l} E(X) = \sum x\, f(x) \\ E(X^2) = \sum x^2 f(x) \\ E(X^n) = \sum x^n f(x) \end{array}\right.$$

$E(a) = a$

$E(aX) = a\,E(X)$

$E(X+Y) = E(X) + E(Y)$

$V(a) = 0$

$V(aX) = a^2 V(X)$

$V(X+Y) = V(X) + V(Y)$

Module – II

Binomial distribution

$P(X=x) = n_{C_x}\, p^x q^{n-x}, \; x = 0, 1, \ldots n$

1) Mean $= \cancel{A}\; np$
2) Variance $= npq$
3) S.D $= \sqrt{npq}$
4) Mode $= \begin{cases} \boxed{(n+1)p} & n \text{ is integer} \\ I.P \text{ of } (n+1)p & n \text{ is not an int} \end{cases}$

Poisson distribution

$P(X=x) = \dfrac{e^{-\lambda}\cdot\lambda^x}{x!}, \; x = 0, 1, \ldots \infty$

1) Mean $= \lambda$
2) Variance $= \lambda$
3) S.D $= \sqrt{\lambda}$
4) Mode $= \lambda$

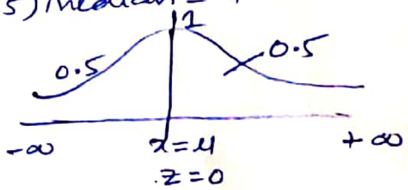
Normal distribution

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

(or)

$z = \dfrac{x-\mu}{\sigma}$

1) Mean $= \mu$
2) Variance $= \sigma^2$
3) S.D $= \sigma$
4) Mode $= \mu$
5) Median $= \mu$



Normal curve

Module – III

Correlation coefficient

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

$X = x - \bar{x}$

$\bar{x} = \dfrac{\sum x_i}{n}$

$Y = y - \bar{y}$

$\bar{y} = \dfrac{\sum y_i}{n}$

Rank correlation coefficient

$$\rho = 1 - \frac{6\sum D^2}{N(N^2-1)}$$

$D = R_1 - R_2$

N = no. of observations

Rank correlation coefficient for repeated ranks

$$\rho = 1 - \frac{6\left[\sum D^2 + \frac{1}{12}(m^3 - m) + \frac{1}{12}(m^3 - m) + \cdots\right]}{N(N^2-1)}$$

$D = R_1 - R_2$

N = No. of observations

Scanned with CamScanner

Module – III

I – Method of least squares / Fitting

Regression lines of y on x	Regression lines of x on y
$y = a + bx$	$x = a + by$
Normal eqns	Normal eqns
$\Sigma y = na + b\Sigma x$	$\Sigma x = na + b\Sigma y$
$\Sigma xy = a\Sigma x + b\Sigma x^2$	$\Sigma xy = a\Sigma y + b\Sigma y^2$

II – By taking deviation from actual Mean

Regression line of y on x	Regression line of x on y
$y - \bar{y} = b_{yx}(x - \bar{x})$	$x - \bar{x} = b_{xy}(y - \bar{y})$
where $b_{yx} = \frac{\Sigma XY}{\Sigma X^2}$	where $b_{xy} = \frac{\Sigma XY}{\Sigma Y^2}$
$X = x - \bar{x}$, $Y = y - \bar{y}$	$X = x - \bar{x}$, $Y = y - \bar{y}$
$\bar{x} = \frac{\Sigma x_i}{n}$, $\bar{y} = \frac{\Sigma y_i}{n}$	

Angle b/w the regression lines $\tan\theta = \frac{1-r^2}{r}\left(\frac{\sigma_x \cdot \sigma_y}{\sigma_x^2 + \sigma_y^2}\right)$

Module – IV

1. Sampling distribution

i) Population Mean $(\mu) = \frac{1}{N}\Sigma X_i$

ii) Population Variance $(\sigma^2) = \frac{1}{N}\Sigma(X_i - \mu)^2$

iii) Population Standard deviation $(\sigma) = \sqrt{\sigma^2}$

Population	Sample
Size = N	Size = n
Mean = μ	Mean = $\bar{x}$
Variance = σ^2	Variance = S^2
S.D = σ	S.D = S

Sampling with replacement

N^n Samples can be drawn

1) Mean of the Sampling distribution

$$\mu_{\bar{x}} = \frac{1}{N^n}\Sigma \bar{x}_i$$

2) Variance of sampling distribution

$$\sigma^2_{\bar{x}} = \frac{1}{N^n}\Sigma(\bar{x}_i - \mu_{\bar{x}})^2$$

3) S.D $(\sigma_{\bar{x}}) = \sqrt{\sigma^2_{\bar{x}}}$

Sampling without replacement

N_{C_n} Samples can be drawn

1) Mean of the Sampling distribution

$$\mu_{\bar{x}} = \frac{1}{N_{C_n}}\Sigma \bar{x}_i$$

2) Variance of Sampling distribution

$$\sigma^2_{\bar{x}} = \frac{1}{N_{C_n}}\Sigma(\bar{x}_i - \mu_{\bar{x}})^2$$

Note:–

$\mu = \mu_{\bar{x}}$

$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$ (with replacement)

$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$ (without replacement)

2. Test of Hypothesis – I

Large Sample $(n \geq 30)$

1) Test of Hypothesis for Single mean:– $Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

H_0:– $\mu = 50$, H_1:– $\mu \neq 50$

2) Test of Hypothesis for diff of means:– $Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

H_0: $\bar{x}_1 = \bar{x}_2$, H_1: $\bar{x}_1 \neq \bar{x}_2$

3) Test of hypothesis for Single proportion:– $Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$

H_0: $P = 0.5$, H_1: $P \neq 0.5$

$p = \frac{x}{n}$, $P = \%$

$Q = 1 - P$

4) Test of Hypothesis for diff of proportions:– $Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

H_0: $P_1 = P_2$, H_1: $P_1 \neq P_2$

$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}$, $Q = 1 - P$

$p_1 = \frac{x_1}{n_1}$, $p_2 = \frac{x_2}{n_2}$

Procedure

1) Null Hypothesis (H_0):–

2) Alternative Hypothesis (H_1):–

3) Level of Significance (α):–

$\alpha = 5\%$, $Z_\alpha = 1.96$

0.05

0.95

95%

4) Test Statistic (Z):– formula

5) Conclusion:–

$|Z| < Z_\alpha$, we accept H_0

$|Z| > Z_\alpha$, we reject H_0

Module – V

$\boxed{\text{Small Samples } n < 30}$

1) i) t – distribution for single mean :– $t = \dfrac{\bar{x} - \mu}{S/\sqrt{n}}$

$H_0 :- \mu = 58$, $H_1 :- \mu \neq 58$

where $S^2 = \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$

L.o.S :- $t_{\alpha/2}$ at $\nu = n-1$ d.o.f

1) ii, t – distribution for diff of means :– $t = \dfrac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

$H_0 :- \bar{x}_1 = \bar{x}_2$, $H_1 :- \bar{x}_1 \neq \bar{x}_2$

where $S^2 = \frac{1}{n_1+n_2-2}\left[\Sigma (x_i - \bar{x}_1)^2 + \Sigma (x_i - \bar{x}_2)^2\right]$

(or)

$S^2 = \frac{1}{n_1+n_2-2}\left[(n_1-1)S_1^2 + (n_2-1)S_2^2\right]$

L.o.S :- $t_{\alpha/2}$ at $\nu = n_1 + n_2 - 2$ d.o.f

Conclusion :– $|t| < t_{\alpha/2}$, we accept H_0

$|t| > t_{\alpha/2}$, we reject H_0

---

2) F – distribution :– $F = \dfrac{S_1^2}{S_2^2}$ $(S_1^2 > S_2^2)$, $F = \dfrac{S_2^2}{S_1^2}$ $(S_2^2 > S_1^2)$

$H_0 : S_1^2 = S_2^2$, $H_1 : S_1^2 \neq S_2^2$

where $S_1^2 = \frac{1}{n_1-1} \Sigma (x_i - \bar{x}_1)^2$

$S_2^2 = \frac{1}{n_2-1} \Sigma (x_i - \bar{x}_2)^2$

L.o.S :– F_α at $(\nu_1 = n_1 - 1, \nu_2 = n_2 - 1)$ d.o.f

Conclusion :– $|F| < F_\alpha$, we accept H_0

$|F| > F_\alpha$, we reject H_0

---

3) χ^2 – distribution :– $\chi^2 = \Sigma \dfrac{(O_i - E_i)^2}{E_i}$
(chi-square)

O_i = observed frequency (Given frequency)
E_i = Expected frequency

$H_0 :$ $O_i = E_i$, $H_1 :$ $O_i \neq E_i$

$\boxed{\text{To find Expected frequencies } E_i}$

Method – I

$E_i = \dfrac{\Sigma O_i}{n}$

L.o.S :– χ^2_α at $\nu = n-1$ d.o.f

Method – II

E_i for $B.D = N \cdot n_{C_x} p^x q^{n-x}$

E_i for $P.D = N \cdot \dfrac{e^{-\lambda} \lambda^x}{x!}$

L.o.S :– χ^2_α at $\nu = n-1$ d.o.f

Method – III

$E_i = \dfrac{\text{row total} \times \text{column total}}{\text{Grand total}}$

L.o.S :– χ^2_α at $\nu = (r-1)(c-1)$ d.o.f

Conclusion :– $|\chi^2| < \chi^2_\alpha$, we accept H_0

$|\chi^2| > \chi^2_\alpha$, we reject H_0