Page No.
Date

Probablity and Stats

Muskaan Punjabi
B - 57.

(1) $x \sim Poi(\theta) \Rightarrow \frac{x - \theta}{\sqrt{\theta}} \sim N(0,1)$

$P(x \leq 200) = 0.025$
$P(x < 200.5) = 0.025$
$P\left(z < \frac{200.5 - \theta}{\sqrt{\theta}}\right) = 0.025$

$P(x \geq 200) = 0.025$
$P(x > 199.5) = 0.025$
$P\left(z > \frac{199.5 - \theta}{\sqrt{\theta}}\right) = 0.025$

$-1.96 = \frac{200.5 - \theta}{\sqrt{\theta}}$

$1.96 = \frac{199.5 - \theta}{\sqrt{\theta}}$

$\Rightarrow \theta - 1.96\sqrt{\theta} - 200.5 = 0$
$\theta = 230.24$

$\Rightarrow \theta + 1.96\sqrt{\theta} - 199.5$
$\theta = 173.67$

$\therefore$ CI for θ at 95% $\Rightarrow (173.67, 230.24)$

(2) $\bar{x} = \sum_{i=1}^{n} x_i \left(\frac{1}{n}\right)$

$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

(i)

(i) $E(\bar{x}) = E\left(\frac{1}{n} \sum_{1}^{n} x_i\right)$
$= \frac{1}{n} \sum (E x_i)$
$= \frac{1}{n} (n\mu)$
$= \mu$

$Var(\bar{x}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)$
$= \frac{1}{n^2} \left(\sum Var(x_i)\right)$
$= \frac{n\sigma^2}{n^2}$
$= \frac{\sigma^2}{n}$

$\therefore x \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Page No.			
Date			

ii. $E(s^2) = E\left(\frac{1}{n-1} \Sigma (x_i - \bar{x})^2\right)$ .... $\left(\Sigma (x_i-\bar{x})^2 = \Sigma (x_i^2 + \bar{x}^2 - 2x_i\bar{x})\right.$

$= \Sigma x_i^2 - 2\Sigma x_i \bar{x} + n\bar{x}^2$

$= \Sigma x_i^2 - 2n\bar{x}^2 + n\bar{x}^2$

$= \Sigma x_i^2 - n\bar{x}^2 \Big)$

$= E\left(\frac{1}{n-1}(\Sigma x_i^2 - n\bar{x}^2)\right)$

$= \frac{1}{n-1}\left(\Sigma E(x_i^2) - nE(\bar{x}^2)\right)$

$= \frac{1}{n-1}\left(n\sigma^2 + n\cancel{\mu^2} - \cancel{n\mu^2} - \sigma^2\right)$

$= \frac{1}{\cancel{n-1}}(\sigma^2)\cancel{(n-1)} = \sigma^2$

iii. $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$

$Var\left(\frac{(n-1)s^2}{\sigma^2}\right) = Var\left(\chi^2_{n-1}\right)$

$\frac{(n-1)^{\cancel{2}}}{\sigma^4} Var(s^2) = 2\cancel{(n-1)}$

$Var(s^2) = \frac{2\sigma^4}{n-1}$

iv. $n=10 \quad \sigma^2 = 100$, $E(s^2) = \sigma^2 = 100$

$\Rightarrow P[50 \le s^2 \le 100] = P\left[\frac{9(50)}{100} \le \chi^2_9 \le \frac{150(9)}{100}\right]$

$= P[4.5 \le \chi^2_9 \le 13.5] = \Rightarrow P(\chi^2_9 \le 13.5) - P(\chi^2_9 \le 4.5)$

$\Rightarrow 0.8587 - 0.1245$

$\Rightarrow 0.7342$.

③ $n=4$

i. $L(p) = \left[{}^4C_0\, p^0 (1-p)^4\right]^{86} \times \left[{}^4C_1\, p^1 (1-p)^3\right]^{75} \times \left[{}^4C_2\, p^2 (1-p)^2\right]^{16}$

$\times \left[{}^4C_3\, p^3 (1-p)\right]^2 \times \left[{}^4C_4\, p^4 (1-p)^0\right]^1$

$= \text{const.} \times (1-p)^{344} \times p^{75} \times (1-p)^{225} \times p^{32} \times (1-p)^{32}$
$\times p^6 \times (1-p)^2 \times p^4$

$= \text{const.} \times (1-p)^{603} \times p^{117}$

$\ln(L(p)) = \ln(\text{const}) + 603\ln(1-p) + 117\ln(p)$

$\frac{\partial \ln(L(p))}{\partial p} = \frac{-603}{1-p} + \frac{117}{p} \Rightarrow \frac{603}{1-p} = \frac{117}{p}$

$603p = 117 - 117p$

$\hat{p} = \frac{117}{720} = 0.1625$

ii.

X	0	1	2	3	4
O_i	86	75	16	2	1
E_i	88.55	68.729	20.003	2.587	0.1255

$\Rightarrow$

X	0	1	≥ 2
O_i	86	75	19
E_i	88.55	68.729	22.7155

$\Sigma \frac{(O_i - E_i)^2}{E_i} = \chi^2_{df}$ $\quad df = 3-1-1 = 1$

$\chi^2_1 = 1.2533$, χ^2_1 at 5% (L.O.S) $= 3.841$

Page No. | Date

As $\chi^2_{cal} < \chi^2_{tab}$, Binomial model is not a good fit.

④ $f(x) = \dfrac{c\beta^3}{(x+\beta)^4}$, $x>0$, $\beta>0$

i. $\Rightarrow \int_0^\infty f(x)dx = 1 \Rightarrow \int_0^\infty \dfrac{c\beta^3}{(x+\beta)^4} dx = 1$

$$1 = c\beta^3 \left[\frac{(x+\beta)^{-3}}{-3}\right]_0^\infty = \frac{c\beta^3}{-3}\left[\frac{1}{(x+\beta)^3}\right]_0^\infty$$

$$1 = \frac{c\beta^3}{-3}\left[0 - \frac{1}{\beta^3}\right] = \frac{c}{3}$$

$$\Rightarrow \boxed{c = 3}$$

$\therefore$ $f(x) = \dfrac{3\beta^3}{(x+\beta)^4}$ is pareto distribution with $\alpha = 3$.

$E(x) = \dfrac{\beta}{\alpha-1} = \dfrac{\beta}{2}$ $\quad$ $Var(x) = \dfrac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} = \dfrac{3\beta^2}{4}$

ii. $\bar{x} = \mu$

$\Rightarrow \dfrac{\Sigma x_i}{n} = \dfrac{\beta}{2} \Rightarrow \boxed{\hat{\beta} = 2\bar{x}_n}$

$$E(\hat{\beta}) - \beta = E(2\bar{x}_n) - \beta$$
$$= 2E(\bar{x}_n) - \beta$$
$$= \cancel{2}\left(\frac{\beta}{\cancel{2}}\right) - \beta = 0$$

$\therefore$ $\hat{\beta}$ is unbiased estimator of β.

$$MSE(\hat{\beta}) = Var(\hat{\beta}) + \{Bias(\hat{\beta})\}^2$$

$$\ldots [Bias(\hat{\beta}) = 0]$$

$$Var(\hat{\beta}) = Var(2\bar{x}_n)$$

$$= 4\,Var(\bar{x}_n) = 4\left(Var\left(\frac{\Sigma x}{n}\right)\right)$$

$$= \frac{4}{n^2}\Sigma Var(x) = \frac{\cancel{4}}{n^{\cancel{2}}}(\cancel{n})\left(\frac{3}{\cancel{4}}\beta^2\right) = \frac{3\beta^2}{n}$$

$$MSE(\hat{\beta}) = \frac{3\beta^2}{n} + 0 = \frac{3\beta^2}{n}$$

As $n \to \infty$, $MSE \to \infty$, $\therefore$ is consistent.

iii. $$Bias[b\bar{x}_n] = E[b\bar{x}_n] - \beta$$

$$= b\,E(\bar{x}_n) - \beta$$

$$= b\left(\frac{\beta}{2}\right) - \beta = \beta\left(\frac{b}{2} - 1\right)$$

$$Var(b\bar{x}_n) = b^2 Var(\bar{x}_n)$$

$$= \frac{b^2}{n^{\cancel{2}}}\left[\cancel{n}\left(\frac{3\beta^2}{4}\right)\right] = \frac{3b^2\beta^2}{4n}$$

$$MSE(b\bar{x}_n) = \frac{3b^2\beta^2}{4n} + \beta^2\left(\frac{b}{2} - 1\right)^2 \Rightarrow \frac{6b\beta^3}{4n} + \beta^2\left(\frac{b}{2} - 1\right)\left(\frac{2x}{2}\right.$$

$$\frac{d}{db}MSE(b\bar{x}_n) = \frac{6b\beta^2}{4n} - \cancel{2}\beta^2\left(\frac{b}{2} - 1\right)\left(\frac{1}{\cancel{2}}\right)$$

$$0 = \frac{3}{2}\frac{b\beta^{\cancel{2}}}{n} - \cancel{\beta^2}\left(\frac{b}{2} - 1\right)$$

$$0 = \frac{3b}{2n} - \frac{b}{2} - 1 \Rightarrow 1 = b\left[\frac{3}{2n} - 1\right] \Rightarrow b = \frac{1}{\left(\frac{3}{2n} - 1\right)}$$

Page No.		
Date		

$$0 = \frac{6b\cancel{\beta^2}}{4n} + \cancel{\beta^2}\left(\frac{b}{2} - 1\right)$$

$$1 = \frac{3b}{2n} + \frac{b}{2} - 1$$

$$1 = \frac{b}{2}\left[\frac{3}{n} + 1\right] \Rightarrow \boxed{\frac{2}{(1+3/n)} = b}$$

We know that $\hat{\beta} = 2\bar{x}_n$ is an unbiased estimator of β as seen in (ii).

$b\bar{x}_n$ is negatively biased for $b \leq 2$, $n \geq 1$ when $b = \frac{2}{(1+3/n)}$. Also, $n \to \infty$, $b \to 2$

$\therefore$ estimator is consistent.

$X \sim U(a,b) \Rightarrow E(x) = \bar{x} = \frac{1}{\cancel{b-a}} \to = \frac{b+a}{2}$

$$2\bar{x} = b + a \quad — ①$$

$$Var(x) = \frac{(b-a)^2}{12} \Rightarrow \quad 2\sqrt{3}s = b - a$$

$$b = 2\sqrt{3}s + a \quad — ②$$

$$a = b - 2\sqrt{3}s \quad — ③$$

Substituting ② in ①

$$\cancel{2}\bar{x} = \cancel{2}\sqrt{3}s + \cancel{2}a$$

$$\underline{\hat{a} = \bar{x} - \sqrt{3}s}$$

ii. Substituting ③ in ①

$$\cancel{2}\bar{x} = \cancel{2}b - \cancel{2}\sqrt{3}s$$

$$\underline{\hat{b} = \bar{x} + \sqrt{3}s}$$

Page No.			
Date			

Let sample of five observations be 1,2,3,4,50

$\bar{x} = \frac{60}{5} = 12$ & $s = \cancel{1.29} \; 21.272$

$\hat{a} = \bar{x} - \sqrt{3}s$ $\hat{b} = \bar{x} + \sqrt{3}s$

$= 12 - \sqrt{3}(21.272)$ $= 12 + \sqrt{3}(21.272)$

$\hat{a} = -24.844$ $\hat{b} = 48.84$

Potential weakness of MOM to estimate a and b for $U(a,b)$ is invalid value, as we can see that there's a value (50) in sample greater than $\hat{b}$ which contradicts $U(a,b)$.

$f(x) = \frac{1}{b} \Rightarrow L = \prod_{i=1}^{n} \frac{1}{b} = \frac{1}{b^n}$

$0 \leq x_1 < x_2 \dots x_n \leq b$

$\therefore$ For L to be maximum, b^n should be minimum, for b to be minimum,

$\boxed{b = X_{max}}$

$\therefore$ MLE is X_{max}

$H_0: p = 0.1$ $H_a: p > 0.1$

$P[\text{Type } 1]_{p=0.1} = P[X \geq 3] + P[X=2]P[X' \geq 3] + P[X=1]P[X' \geq 4] + P[X=0]P[X' \geq 5]$

$= (1 - 0.8891) + (0.2824)(0.0043) + (0.6590 - 0.2824)(1 - 0.9744) + (0.8891 - 0.6590)(1 - 0.8891) = 0.14727 \simeq 15\%$

$$p = 0.3$$

$$= P[X \geq 3] + P[X=2]P[X' \geq 3] + P[X=1]P[X \geq 4] + P[X=0]P[X \geq 5]$$

$$= 0.7472 + (0.0138)(0.2763) + (0.085 - 0.0138)(0.5075) + (0.2528 - 0.085)(0.7472)$$

$$= 0.91253 = \underline{91.253\%}$$

$$P[\text{Type II error}] = P[\text{Accept } H_0 \mid H_0 \text{ is false}]$$

$$= 1 - 0.91253 = 0.08747$$

$$= \underline{8.747\%}$$

$$\frac{\hat{p} - p}{\sqrt{\frac{\hat{p}\hat{q}}{n}}} \sim N(0,1)$$

$$\hat{p} = \frac{40}{120} = \frac{1}{3} \qquad \hat{q} = \frac{2}{3}$$

$$Z_{10\%} = 1.2816 \Rightarrow \hat{p} - 1.2816\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= \frac{1}{3} - 1.2816\left[\sqrt{\frac{(1/3)(2/3)}{120}}\right]$$

$$= 0.2782$$

$H_0: p = 0.278 \qquad H_a: p > 0.278$

$$z = \frac{x - np}{\sqrt{npq}} = \frac{39.5 - 120(0.278)}{\sqrt{120(0.278)(1-0.278)}} = 1.2511$$

$Z_{tab} = 1.2816$, As $Z_{cal} < Z_{tab}$, we have insufficient evidence to reject H_0 at 10%.

Page No.
Date

vi. The C.I states that there is 10% chance of $p \le 0.2782$.
The hypothesis test states that p could be less than or equal to 0.278 with probability more than 10% (approximately 10.5%).

vii. $H_0: \mu_1 = \mu_2$ $\qquad H_a: \mu_1 \neq \mu_2$

$$\frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

$\bar{x}_1 = 65$ $\qquad \bar{x}_2 = 70$ $\qquad S_2 = 70$ $\qquad n_1 = 12$
$\qquad S_1 = 54$ $\qquad n_2 = 15$

$S_p^2 = 4027.04 \Rightarrow S_p = 63.459$

$$\frac{65 - 70 - 0}{63.459\sqrt{\frac{1}{12}+\frac{1}{15}}} = \underline{-0.2034}$$

$t_{cal} = -0.2034$ $\qquad t_{tab} = \pm 2.060$

As t_{cal} lies between t_{tab}, we have insufficient evidence to reject H_0 at $\underline{\underline{5\%}}$.

7. i. $\bar{x} = 30$ $\qquad S_x^2 = 64.3125$
(Public)
$\bar{y} = 35.055$ $\qquad S_y^2 = 67.4027$

ii. The condition that needs to be true for testing equal mean is the population variances of both samples are equel.
we can look at the sample variances of both data being close enough, to assume that their populations variances are same.

iii. $H_0: \mu_x = \mu_y$ $\quad H_a: \mu_y > \mu_x$

$$Z_c = \frac{\bar{y} - \bar{x} - (\mu_y - \mu_x)}{S_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \quad \ldots \left(S_p = \sqrt{\frac{(n_x-1)S_x^2 + (n_y-1)S_y^2}{n_x + n_y - 1}} \right)$$

$$S_p = 8.1153$$

$$Z_c = \frac{35.055 - 30}{8.1153\sqrt{\frac{2}{9}}} = 1.321$$

$$Z_{tab} = 1.746$$

$\therefore$ As $Z_{cal} < Z_{tab}$ i.e. $\left(t_{0.05,16}\right)$, we have insufficient evidence to reject H_0.

(8) i. Parameter $= \theta$ $\quad$ Estimator $: \hat{\theta}$

Unbiased estimator $\Rightarrow$ $\underline{E(\hat{\theta}) = \theta}$

ii. $\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$

iii. $\text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta}) + [\text{Bias}(\hat{\theta})]^2 = E(\hat{\theta} - \theta)^2$

iv. $Bias(\tilde{\theta}) = E(\tilde{\theta}) - \theta = +ve$

$$\therefore \underline{E(\tilde{\theta}) > \theta}$$

$MSE(\hat{\theta}) > MSE(\tilde{\theta})$: $\tilde{\theta}$ is better, efficient parameter compared to $\hat{\theta}$ even though $\tilde{\theta}$ has got +ve bias. For a parameter to be efficient, MSE must be low.

v. If by increasing the sample size, i.e. 'n'.
$n \to \infty$
$MSE \to 0$, If this is true, the ~~pa~~ estimator can be termed as consistent.

vi. Two methods of estimating θ =>
- Method of Moments (MOM)
- Maximum Likelihood Estimate (MLE).

(a) i. $t_k = \dfrac{\bar{x} - \mu}{S/\sqrt{n}}$, $\bar{x}$ = Sample Mean, μ = Population Mean, S = Sample Variance, n = Sample Size

$= \dfrac{N(0,1)}{\sqrt{\chi^2_k / k}}$

k = degree of freedom; χ^2_k and $N(0,1)$ are independent

ii. $E(t_k) = \cancel{}\ 0$, $k > 1$

$Var(t_k) = \dfrac{k}{k-2}$, $k > 2$

iii. $n = 10$, $\bar{x} = 50$, $S^2 = 48.667$

a. $t_k = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{50 - \mu}{\sqrt{48.667/10}}$

$$-3.25 \le \frac{50-\mu}{\sqrt{48.667/10}} \le 3.25$$

$$50 - 3.25\sqrt{\frac{48.667}{10}} \le \mu \le 50 + 3.25\sqrt{\frac{48.667}{10}}$$

$\therefore$ Confidence Interval for population mean at 99% CI $\Rightarrow$ (42.8303, 57.1696)

b. $X \sim N\left(0, \frac{9}{7}\right) \Rightarrow \bar{x} \sim N\left(0, \frac{9}{70}\right)$

$$\cancel{Z_c = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \Rightarrow \frac{50-\mu}{\sqrt{9/70}}}$$

$$\Rightarrow \frac{\bar{x}-0}{\sqrt{9/70}} \sim N(0,1)$$

$$t_{n-1} \sim N\left(0, \frac{9}{7}\right)$$

$$\Rightarrow \frac{\bar{x}-\mu}{S/\sqrt{n}} \sim N\left(0, \frac{9}{7}\right)$$

$$\Rightarrow \frac{\frac{\bar{x}-\mu}{S/\sqrt{n}} - 0}{\sqrt{9/7}} \sim N(0,1)$$

$$\Rightarrow \frac{\bar{x}-\mu}{S/\sqrt{n}} \times \frac{1}{\sqrt{9/7}} \sim N(0,1)$$

$$\Rightarrow \frac{\bar{x}-\mu}{\sqrt{9S^2/7n}} \sim N(0,1)$$

Page No. | Date

$$-2.58 < \frac{\bar{x}-\mu}{\sqrt{9s^2/7n}} < 2.58$$

$\Rightarrow$ $50 \pm 2.58\sqrt{\frac{9(48.667)}{7(10}}$ = Confidence Interval at 99%.

$\Rightarrow$ $\underline{(43.54,\ 56.45)}$

(10) $X \sim Poi(\lambda) \Rightarrow \Sigma X \sim Poi(n\lambda)$ $n = 1000$

$\Rightarrow \frac{\Sigma X - n\lambda}{\sqrt{n\lambda}} \dot{\sim} N(0,1)$

$H_0: \lambda = 3$ $H_a: \lambda \neq 3$

i. $P[\text{type I}] = P[\text{Rejecting } H_0 \mid H_0 \text{ is true}]$

$$= P[\Sigma X > 3100 \mid \lambda = 3]$$

$$= P\left[Z > \frac{3100-3000}{\sqrt{3000}}\right]$$

$$= 1 - P[Z < 1.8257]$$

$$= P[Z > 1.8257] = 0.03394$$

$$= 3.394\%$$

ii. $P[\text{type 2}] = P[\text{Accepting } H_0 \mid H_0 \text{ is false}]$

$$= P[\Sigma X < 3100 \mid \lambda \neq 3]$$

$$= P\left[Z \leq \frac{3100 - n\lambda}{\sqrt{n\lambda}}\right]_{\lambda \neq 3}$$

$$= P\left[Z < \frac{3100 - 1000(\lambda)}{\sqrt{1000(\lambda)}} \,\middle|\, \lambda \neq 3\right]$$

iii. Power of test = $1 - P[\text{type II}]$

$$= 1 - P\left[Z < \frac{3100 - 1000\lambda}{\sqrt{1000\lambda}} \,\middle|\, \lambda \neq 3\right]$$

$$= P\left[Z > \frac{3100 - 1000\lambda}{\sqrt{1000\lambda}} \,\middle|\, \lambda \neq 3\right]$$

iv. $\hat{\lambda} \pm 2.5758\sqrt{\frac{\hat{\lambda}}{n}}$ ... $\hat{\lambda} = \frac{x}{n} = \frac{2900}{1000} = 2.9$

$\Rightarrow 2.9 \pm 2.5758\sqrt{\frac{2.9}{1000}}$

$\Rightarrow 2.9 \pm 0.1387 \Rightarrow 99\%$ CI $= (2.7613, 3.0387)$

(11) Initial data:

$n = 12$

$\Sigma x = 213200$ $\quad\quad \bar{x} = 17767$

$\Sigma x^2 = 4\,919\,860\,000$ $\quad\quad s = 10144$

After replacing two temporary employees salary:

$n = 12$

$\Sigma x = 213200 - 11000 - 48000 + 27500 + 31500$

$= 213200$

$\bar{x} = \frac{213200}{12} = \underline{17767}$

$\Sigma x^2 = 4\,919\,860\,000 - 11000^2 - 48000^2 + 27500^2 + 31500^2$

$= 4233360\,000$

$s = \sqrt{\frac{\Sigma x^2 - 12(\bar{x})^2}{n-1}} = \underline{6363}$

sample

$\therefore$ We observed that the mean remains same, but standard deviation changes as there's less difference between two va.

Page No.
Date

(12) $x_i \sim U(0,\theta)$

$f(x) = \frac{1}{\theta - 0}, \quad \theta > 0$

i. $L(\theta) = \begin{cases} \frac{1}{\theta^n}, & \theta > 0 \\ 0, & \text{otherwise} \end{cases}$

$0 \leq x_1 x_2 \dots x_n \leq \theta$, To maximise $\frac{1}{\theta^n}$, θ^n needs to be minimized.

$\theta \geq X_{max}$. $\therefore$ $\underline{L = X_{max}}$.

$\ln L(\theta) = -n\ln\theta \Rightarrow \frac{\partial \ln(L(\theta))}{\partial\theta} = \frac{-n}{\theta}$

$\frac{\partial^2 \ln(\theta)}{\partial\theta} = \frac{n}{\theta^2} > 0$

As double derivative of $L(\theta)$ is always greater than 0, CRLB for the variance for the unbiased estimator of θ is not possible.

$\hat{\theta}(c) = cY$, $Y = X_{max}$

ii. $P[Y \leq y] = P[X_{max} \leq y]$
$= P[X_1, X_2, \dots, X_n \leq y]$
$= P[X_1 \leq y]\, P(X_2 \leq y) \dots P[X_n \leq y]$
$= P[X_i \leq y]^n$
$= \left[\int_0^y \frac{1}{\theta}\, dx\right]^n = \left(\left[\frac{x}{\theta}\right]_0^y\right)^n = \left(\frac{y}{\theta}\right)^n = F_Y(y)$

$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(\frac{y}{\theta}\right)^n = \frac{ny^{n-1}}{\theta^n}, \quad \underline{0 < y < \theta}$

$$E(y^k) = \int_0^{\theta} y^k \frac{n}{\theta^n} y^{n-1} \, dy$$

$$= \frac{n}{\theta^n} \int_0^{\theta} y^{n+k-1} \, dy$$

$$= \frac{n}{\theta^n} \left[\frac{y^{n+k}}{n+k} \right]_0^{\theta} = \frac{n}{\theta^n (n+k)} \left[\theta^{n+k} \right]$$

$$E[y^k] = \frac{n\theta^k}{n+k} , \quad k > 0$$

iii. $$\text{Bias}[\hat{\theta}(c)] = E(\hat{\theta}(c)) - \theta$$

$$= E(cY) - \theta$$

$$= c\left[\frac{n\theta^1}{n+1}\right] - \theta = \cancel{\frac{n\theta - n\theta - \theta}{n+1}}$$

$$= \frac{cn\theta}{n+1} - \theta$$

$$\text{Bias}[\hat{\theta}(c)] = \theta\left[\frac{cn}{n+1} - 1\right]$$

$$\text{Var}(\hat{\theta}(c)) = \text{Var}(cY) = E(Y^2) - [E(Y)]^2$$

$$= \left(\frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2}\right) c^2$$

$$\text{MSE}(\hat{\theta}(c)) = \text{Var}(\hat{\theta}(c)) + \left(\text{Bias}[\hat{\theta}(c)]\right)^2$$

$$= c^2\left(\frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2}\right) + \theta^2\left[\frac{cn}{n+1} - 1\right]^2$$

$$= \frac{c^2 n\theta^2}{n+2} - \cancel{\frac{c^2 n^2 \theta^2}{(n+1)^2}} + \cancel{\frac{c^2 n^2 \theta^2}{(n+1)^2}} + \theta^2 - \frac{2cn\theta^2}{n+1} = \frac{c^2 n\theta^2}{n+2} - \frac{2cn\theta^2}{n+1} + \theta^2$$

Page No.
Date

iv. $\text{Bias}[\hat{\theta}(c)] = \theta\left[\frac{cn}{n+1} - 1\right] = 0$

$$\frac{cn}{n+1} - 1 = 0$$

$$\frac{cn}{n+1} = 1 \Rightarrow \boxed{c_u = \frac{n+1}{n}}$$

$$\cancel{c_n = n+1 \Rightarrow} \quad c(n-1) = 1$$

$$\cancel{c_u = \frac{1}{n-1}}$$

v. $\text{MSE}[\hat{\theta}(c)] = \frac{c^2 n\theta^2}{n+2} - \frac{2nc\theta^2}{n-1} + \theta^2$

$$\frac{d}{dc}\text{MSE}(\hat{\theta}(c)) = \frac{\cancel{2}c\cancel{n}\cancel{\theta^2}}{n+2} - \frac{\cancel{2n}\cancel{\theta^2}}{n-1} = 0$$

$$\frac{c}{n+2} = \frac{1}{n-1}$$

$$\boxed{c_m = \frac{n+2}{n+1}}$$

vi. We will prefer $\hat{\theta}(c_m)$ for estimating θ, as low MSE is preferred over high MSE.

As $n \to \infty$, $\text{MSE}[\hat{\theta}(c_m)] \to 0$