Page No. 1
Date

Roll no. 69
Div B

Probablity & Stats. Assignment II

1. $X \sim Poi(\theta)$ $\quad\quad \hat{\lambda} = 200$

$\frac{X-\theta}{\sqrt{\theta}} \approx N(0,1)$ $\quad\quad \hat{\lambda} = \frac{\Sigma x_i}{n} = 200$

$\therefore \frac{\hat{\lambda}-\lambda}{\sqrt{\lambda/n}} \approx N(0,1)$

$P(Z_{2.5\%} < Z < Z_{97.5\%}) = 0.95$

$P(-1.96 < Z < 1.96) = 0.95$

$P\left(-1.96 < \frac{\hat{\lambda}-\lambda}{\sqrt{\lambda/n}} < 1.96\right) = 0.95$

$P\left(\hat{\lambda} - 1.96(\sqrt{\lambda/n}) < \lambda < \hat{\lambda} + 1.96\sqrt{\lambda/n}\right) = 0.95$

$\therefore$ 95% CI of λ when $\hat{\lambda} = 200$ and $n = 1$

95% CI of $\lambda = \left(200 - 1.96\sqrt{\frac{200}{1}}, 200 + 1.96\sqrt{\frac{200}{1}}\right)$

$= (172.28141, 227.7185)$

Page No.
Date

2. $\overline{x} = \sum_{1}^{n} x_i$

$$s^2 = \frac{1}{n-1} \sum_{1}^{n} (\cancel{\Sigma} x_i - \overline{x})^2$$

i) $\mathcal{E}(\overline{x}) = \boxed{\mathcal{E}(\Sigma x_i / n)}$

$$= \frac{\Sigma E(x_i)}{n}$$ - [since x_i are independent and identically distributed]

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

$$\therefore E(\overline{x}) = \mu$$

$$V(\overline{x}) = V\left(\frac{\Sigma x_i}{n}\right)$$

$$= \frac{1}{n^2} V(\Sigma x_i)$$ - [since x_is are iids]

$$= \frac{1}{n^{\cancel{2}}} \times \cancel{n} \sigma^2$$

$$= \frac{\sigma^2}{n}$$

$$\therefore V(\overline{x}) = \frac{\sigma^2}{n}$$

ii) $E(s^2) = \sigma^2$ - tpt.

$$E(s^2) = E\left[\frac{1}{n-1} \Sigma (x_i - \overline{x})^2\right]$$

$$= \frac{1}{n-1} \times E(\Sigma(x_i - \overline{x}))^2$$

Page No.

Date

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

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

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

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

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

$$\frac{1}{n-1}\left[nE(\sigma^2+\mu^2) - n\left(\frac{\sigma^2}{n}+\mu^2\right)\right]$$

hence proved.

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

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

$$= \sigma^2$$

hp.

with mean $n-1$

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

$$E \cdot \frac{s^2}{\sigma^2} = 1$$

$$E \cdot (s^2) = \sigma^2$$

Variance :

$$V\left(\frac{(n-1)s^2}{\sigma^2}\right) = V(\chi^2_1) = 2(n-1)$$

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

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

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

(iv) $n = 10$

$\sigma^2 = 100$

$E(s^2) = \sigma^2 = 100$

$\therefore P(50 < s^2 < 150) = ?$

$\cancel{P(50 < x)}$ $\quad P(s^2 < 150) - P(s^2 < 50)$

$= P\left(\frac{9s^2}{100} < \frac{150 \times 9}{100}\right) - P\left(\frac{9s^2}{100} < \frac{50 \times 9}{100}\right)$

$= P(\chi^2_{n-1} < 13.5) - P(\chi^2_9 < 4.5)$

$= 0.8587 - 0.1245$

$= \underline{\underline{0.7342}}$

3. Let the no. of claims be X.

$X \sim Bin(4, p)$

(i) $L = \prod_{i=1}^{180} f(x_i)$

$= (P(X=0))^{86} \cdot P(X=1)^{75} \cdot P(X=2)^{16} \cdot P(X=3)^{2} \cdot P(X=4)^{1}$

$= ({}^4C_0 \cdot p^0 \cdot (1-p)^4)^{86} \cdot ({}^4C_1 p^1 (1-p)^3)^{75} \cdot ({}^4C_2 \cdot p^2 (1-p)^2)^{16} \cdot ({}^4C_3 p^3 (1-p)^2)^2 \cdot ({}^4C_4 \cdot p^4)^1$

$= \text{constant} \cdot p^{117} \cdot (1-p)^{603}$

$\ln L = 117 \ln p + 603 \ln (1-p) + \ln c.$

derivative

$\frac{d \ln L}{dp} = \frac{117}{p} + \frac{603}{(1-p)}(-1) + 0$

$\frac{117}{p} - \frac{603}{(1-p)} = 0$

$\times \left[\begin{array}{l} 117(1-p) + 603p = p - p^2 \\ 117 - 117p + 603p = p - p^2 \\ 117 - 486p = p - p^2 \end{array} \right.$

$\frac{117}{p} = \frac{603}{(1-p)}$

$117 - 117p = 603p$

$117 = 720p$

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

$\boxed{\hat{p} = 0.1625}$

(ii) Goodness of fit test :

H_0 = no. of claims conforms to the binomial model.
H_1 = no. of claims does not conform to the binomial model

	0	1	2	3	4	
O_i	86	75	16	2	1	$\Sigma 180$
E_i	88.542	68.724	19.998	2.574	0.108	

$(\Sigma O_i \times P(X=x))$

$p = 0.1625$
$n = 4$
$q = 0.8375 = (1-p)$
$P(X=0) \Rightarrow {}^4C_0 \cdot p^0 \cdot (1-p)^4 = 0.4919$
$P(X=1) \Rightarrow {}^4C_1 \cdot p^1 \cdot (1-p)^3 = 0.3818$
$P(X=2) = {}^4C_2 \cdot p^2 \, (1-p)^2 = 0.1111$
$P(x=3) = {}^4C_3 \cdot p^3 \cdot (1-p)^1 = 0.0143$
$P(X=4) = {}^4C_4 \cdot p^4 \cdot (1-p)^0 = 0.0006$

$\because$ (3 & 4's) $E_i < 5$,

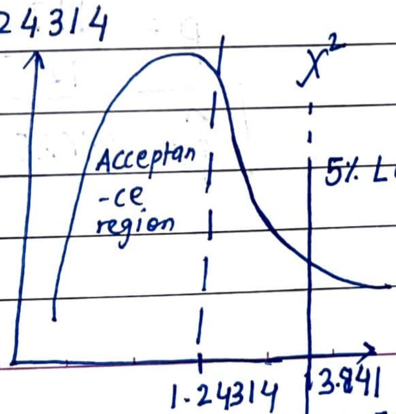
O_i	86	75	19
E_i	88.542	68.724	22.68
$(O_i - E_i)^2/E_i$	0.0729	0.573	0.597

$$TS = \Sigma \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{5-1-1-2}$$

$$= 1.24314 \sim \chi^2_1$$

$\chi^2_{cal} \rightarrow 1.24314$

$\chi^2_{tab} \rightarrow$



no claim conform

$\rightarrow$ we have insufficient evidence to reject H_0 @ 5% LOS $\therefore$

4. $$f(x) = \frac{c\beta^3}{(x+\beta)^4} \quad ; \; x > 0, \; \beta > 0$$

c is a constant & β is a parameter

(i) Comparing the above $f(x)$ to pdf of Pareto dist. (2 parameter version)

For pareto,

$$PDF = \frac{\alpha \lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}$$

Thus, β corresponds to α.

→ Thus, $c = 3$ (here c corresponds to α)

$$E(X) = \frac{\lambda}{\alpha - 1}$$

$$E(X) = \frac{\beta}{2}$$

$$V(X) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} = \frac{3\beta^2}{4}$$

(j) $E(X) = \beta/2$

$E(\bar{X}) = \bar{X} = \beta/2$

$\hat{\beta} = 2\bar{X}$

⇒ $$\text{Bias} = E(\hat{\beta}) - \beta$$

$$= E(2\bar{X}) - \beta$$

$$= 2E(\bar{X}) - \beta$$

$$= 2E(\Sigma x_i / n) - \beta$$

$$= \frac{2}{n} \Sigma E(x_i) - \beta$$

$$= \frac{2}{\cancel{n}} \cdot \frac{\cancel{n}\beta}{2} - \beta \Rightarrow \beta - \beta = 0$$

Bias = 0, hence unbiased.

Page No.
Date

(ii) $MSE = V(\hat{B}) + Bias^2$

$= V(\hat{B}) + 0^2$

$= V(2\bar{X})$

$= 2^2 V(\bar{X})$

$= 4V(\Sigma x_i/n)$

$= \frac{4}{n^2} \Sigma V(x_i)$

$= \frac{4}{n^2} \times \frac{3\beta^2 \times n}{4}$

$MSE = \frac{3\beta^2}{n}$ = As n increases, MSE tends to 0

hence, $\hat{\beta}$ is consistent.

(iii) $\hat{\beta} = b\bar{X}$

$E(\hat{\beta}_2) = E(b\bar{X})$

$= bE\left(\frac{\Sigma x_i}{n}\right)$

$= \frac{b}{n} \Sigma E(x_i)$

$= \frac{b}{\cancel{n}} \times \frac{\beta}{2} \times \cancel{n}$

$= \frac{b\beta}{2}$

$V(\hat{\beta}_2) = V(b\bar{X})$

$= b^2 V(\Sigma x_i/n)$

$= \frac{b^2}{n^2} \Sigma V(x_i)$

$= \frac{b^2}{n^{\cancel{2}}} \times \frac{3}{4} \times \beta^2 \times \cancel{n}$

$V(\hat{\beta}_2) = \frac{3}{4} \frac{b^2\beta^2}{n}$

$$MSE = V(\beta_2) + (Bias)^2$$

$$= \frac{3b^2\beta^2}{4n} + \left(\frac{b\beta}{2} - \beta\right)^2$$

$$= \frac{3b^2\beta^2}{4n} + \left(\frac{b\beta - 2\beta}{2}\right)^2$$

$$= \frac{3b^2\beta^2}{4n} + \left(\frac{b^2\beta^2 - 2(b\beta)(2\beta) + 4\beta^2}{4}\right)$$

$$= \frac{3b^2\beta^2 + nb^2\beta^2 + 4n\beta^2 - 4n\beta^2 b}{4n}$$

$$MSE = \frac{\beta^2}{4n}(3b^2 + nb^2 + 4n - 4nb)$$

$$\frac{dM}{db} = \frac{\beta^2}{4n}(6b + 2nb - 4n)$$

$$\frac{dM}{db} = 0 = \frac{\beta^2}{4n}(6b + 2nb - 4n)$$

$$6b + 2nb - 4n = 0$$

$$3b + nb - 2n = 0$$

$$b(3+n) = 2n$$

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

$$b = 2/(3+n/n)$$

$$\frac{d^2M}{db^2} = \frac{\beta^2}{4n}(6+2n)$$

$$\frac{\beta^2}{4n}(6+2n) > 0$$

hence, $\frac{d^2M}{db^2}$ is minimum

Thus, $b = \frac{2}{(1+3/n)}$ minises the MSE.

Page No.
Date

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

$X \sim U(a,b)$

i) $E(X) = \frac{a+b}{2}$

$\bar{X} = \frac{a+b}{2}$

$2\bar{X} = a+b$ — ① $\therefore b = 2\bar{X} - a$ — ②

$V(X) = \frac{1}{12}(b-a)^2$

$s^2 = \frac{1}{12}(b-a)^2$

$12s^2 = (b-a)^2$

$\sqrt{12}\, s = b-a$

$2\sqrt{3}\, s = b-a$

$2\sqrt{3}\, s = 2\bar{X} - a - a$ — from ②

$2\sqrt{3}\, s = 2\bar{X} - 2a$

$\therefore \sqrt{3}\, s = \bar{X} - a$

$\boxed{\hat{a} = \bar{X} - \sqrt{3}\, s}$

ii) $b = 2\bar{X} - a$

$b = 2\bar{X} - (\bar{X} - \sqrt{3}\, s)$

$b = 2\bar{X} - \bar{X} + \sqrt{3}\, s$

$\boxed{\hat{b} = \bar{X} + \sqrt{3}\, s}$

iii) Sample of observations: 1, 2, 9, 25, 100

$\therefore \bar{X} = \frac{137}{5} = 27.4$

$s = 41.704$

$\therefore$ ~~$a = 34.552$~~

~~$b =$~~

$\hat{a} = 27.4 - \sqrt{3}(41.704)$

$\boxed{\hat{a} = -44.8334}$

$\hat{b} = \bar{X} + \sqrt{3}\, s = \boxed{99.633 = \hat{b}}$

hence we can conclude that the value of $\hat{b}$ is very close to sample value of the given, but the value of $\hat{a}$ is very far from the true sample value so it is a potential weakness of method of moments.

iv) $U(a,b)$ where $a = 0$
$\therefore X \sim U(0,b)$

$$L = \prod_{i=1}^{n} f(x_i) \rightarrow (1/b-a) = 1/b-0 = 1/b$$

$$= \prod_{i=1}^{n} \left(\frac{1}{b}\right)$$

$$L = \left(\frac{1}{b}\right)^n$$

$$\log L = n \log b^{-1} = -1(n \log b)$$

$$\log L = \underline{\underline{-n \log b}}$$

$$\frac{d \log L}{db} = \frac{-n}{b}$$

$$-\frac{n}{b} = 0$$, this makes $b \rightarrow \infty$ which is not possible

and also

$$\frac{d^2 \log L}{db^2} = d\left(-\frac{n}{b}\right) = \frac{n}{b^2} > 0$$

which we have a minimum likelihood estimate.

$\therefore$ we can put together that $\hat{b} = \max(X_i)$

Q6. $H_0 : p = 0.1$

$H_1 : p > 0.1$

p = proportion of hits of missiles, each targeted @ a similar asteroid.

(i) P(Type I Error) = P(rejecting Ho / Ho is true)

Let X be the no. of hits in the first 12 missiles and Y be the no. of hits in the second set of 12 missiles.

$$P(\text{Type 1 error}) = P(X \geq 3 / p = 0.1) + P(X = 0 / p = 0.1) \cdot P(Y \geq 5 / p = 0.1) + P(X = 1 / p = 0.1) \cdot P(Y \geq 4 / p = 0.1) + P(X = 2 / p = 0.1) \cdot P(Y \geq 3 / p = 0.1)$$

$= (1 -$

Q6 cont.

iv) $n = 120$ missiles

$$\frac{X/n - P}{\sqrt{\frac{P(1-P)}{n}}} \sim N(0,1)$$

$x = 40$

$\hat{p} = \frac{40}{120} = 0.3333$ $\quad Z_{10\%} = 1.2816$

$\therefore CI = 0.3333 - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \times 1.2816$

$Z_{cal} = 0.27815$

v) $H_0: p = 0.278$

$H_1: p > 0.278$

$\therefore \frac{x - nP}{\sqrt{npq}} \sim N(0,1)$

$\therefore$ using continuity correction

$$= \frac{39.5 - 33.36}{\sqrt{24.086}} = 1.25108$$

$\therefore$ we donot have sufficient evidence to reject H_0 at 10% level of significance as $Z_{cal} < Z_{tab}$

vi)

vii) $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

$n_1 = 12$ $n_2 = 15$

$x_1 = 65$ $x_2 = 70$

$\sigma_1 = 54$ $\sigma_2 = 70$

$$s^2p = \frac{1}{25}\left[11(2916) + 14(4900)\right] = 4027.04$$

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

$$TS = \frac{(65-70)-0}{4027.04(0.15)} = -0.2034$$

$t_{tab} > t_{cal}$

CI at 95% = $(2.2634, 1.8566)$

$\therefore$ we have insufficient information to reject H_0.

Q7. (i) Public $\quad\quad$ Private

$\bar{X}_1 = 30$ $\quad\quad$ $\bar{X}_2 = 35.056$

$s_1^2 = 64.3125$ $\quad\quad$ $s_2^2 = 67.4027$

$n_1 = 9, \; n_2 = 9$

ii) 2 sided t-test can be applied in case when the samples come with equal variances. hence we are testing : $H_0 : \sigma_1^2 = \sigma_2^2$ v/s $H_1 : \sigma_1^2 \neq \sigma_2^2$

Test statistic is $\dfrac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{n_1-1, n_2-1}$

$$= \frac{64.3125 / 59.91}{67.4027 / 57.166} \sim F_{8,8}$$

$$= \frac{1.0734}{1.17906}$$

$$TS = 0.9103 \sim F_{8,8}$$

$F_{8,8}$ values at 5% levels (assumed to be normal dist ∴ 5% assumed), since it's 2 sided taking 2.5% + 97.5%.

→ 4.433 and 0.2256 . Since the value of the T.S. b/w the above values, we have insufficient evidence to reject the hypothesis & conclude that the popln. variances are equal.

(iii) $H_0 : \mu_{pr} : \mu_{pu}$

$H_1 : \mu_{pr} > \mu_{pu}$

$$TS \Rightarrow \frac{(\bar{X}_2 - \bar{X}_1) - (\mu_2 - \mu_1)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Date

$$= \frac{(35.055 - 30) - (0)}{8.1152\sqrt{2/9}}$$

$\therefore$ Test Statistic $= 1.32151$

where $S_p^2 = \frac{8(64.3125) + 8(67.4027)}{16}$

$S_p = \sqrt{65.8576}$

$S_p = 8.11526$

$\therefore$ Statistic $= 1.31251$

Tabulated val. , $t_{16, 5\%} = 1.746$

$\because$ $T.S_{cal} < T.S_{tab}$, we have insufficient evidence to reject H_0

Q8. (i) Unbiased Estimator

→ An unbiased estimator is an estimator where the expected value is equal to the value of parameter of the distribution.

→ Estimator is said to be biased if bias is equal to zero for all values of parameter θ.

(ii) Biased Estimator

→ An estimator which delivers an estimate which is consistently different than the parameter to be estimated.

⇒

(iii) MSE → Mean square Error of a estimate is defined by $MSE(\hat{\theta}) = Var(\theta) + Bias^2$

(iv) Both would be considered to be effecient when MSE

Cont. Q8. minimises such that when $n \to \infty$, $MSE = 0$

(iv) The estimator having less MSE is more consistent.

(vi) (i) Method of moments
(ii) Maximum Likelihood Estimate

Q9.
(i) The variable t_K is defined as: $\dfrac{N(0,1)}{\sqrt{\chi^2_K / K}}$

where k is the degree of freedom and 2 random variables $N(0,1)$ and χ^2_K are independent.

(ii) For $k > 2$
$E(t_K) = 0$
$V(t_K) = \dfrac{k}{k-2}$

(iii) a) Using t-test tables:

$$\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

$CI = 99\%$
$1 - \alpha = 99\%$
$\alpha = 1\%$ $\therefore$ $\alpha/2 = 0.5\%$ (2 sided)

$$P\left(t_{9,0.5\%} < \frac{\bar{X} - \mu}{s/\sqrt{n}} < t_{9,98.5\%}\right) = 0.99$$

$$P\left(\bar{X} - \frac{s}{\sqrt{n}} \cdot t_{9,0.5\%} < \mu < \bar{X} + \frac{s}{\sqrt{n}} \cdot t_{9,98.5\%}\right) =$$

Page No.
Date

99% CI $= (42.83, 57.17)$

b) from (ii), $t_{n-1} \sim N\left(0, \sqrt{\frac{n-1}{n-3}}\right)$

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

i.e. $$\frac{\bar{x}-4}{S/\sqrt{n}} \sim N\left(0, \sqrt{\frac{9}{7}}\right)$$

$$\frac{\bar{x}-4}{S/\sqrt{7n/9}} \sim N(0,1)$$

Critical values of level of confidence 2.58.

$$CI = \left(50 - 2.58\sqrt{\frac{48.667}{10} \times \frac{9}{7}}, \; 50 + 2.58\sqrt{\frac{48.667 \times 9}{10 \times 7}}\right)$$

$$CI = (43.54, 56.45)$$

Q10. i) $n = 1000$ $\quad H_0 : n\lambda > 3100$

$\lambda = 3$ $\quad H_1 : n\lambda \leqslant 3100$

$X \sim Poi(n\lambda)$

$X \sim Poi(3000)$

ii) $P(\text{Type I Error}) = P(\text{Reject } H_0 / H_0 \text{ is true})$

$= P(X > 3100 / X \sim Poi(3000))$

$$P(X > 3100) = P\left(Z > \frac{3100 - 3000}{\sqrt{3000}}\right)$$

$$= P(Z > 1.8257)$$

$$= 1 - P(Z < 1.8257)$$

$$= 1 - 0.657$$

$$P(X > 3100) = 0.343$$

ii) Type II error is the event of accepting hypothesis when it is false.

iii) Power of test is P (rejecting H_0 / H_0 is false)

$$= P[\; X > 3100 \;/\; X \sim Poi(n\lambda) - N(n\lambda - n\lambda)]$$

$$= 1 - P\left(z < \frac{3100 - n\lambda}{\sqrt{n\lambda}}\right)$$

The val. of power of test will depend on the value of parameter under hypothesis. [H_0 is false
↳ para meter - normal

iv) $\hat{\mu}$ is the estimated value,

$$\therefore \hat{\mu} \sim N(\mu, \hat{\mu}/n)$$

$$\therefore \frac{\hat{\mu} - \mu}{\sqrt{\mu/n}} \sim N(0,1)$$

$$P\left(-2.578 < \frac{2900 - \mu}{\sqrt{2900/1000}} < 2.578\right) = 0.99$$

$$\therefore 99\% \;\; CI = (2.7619, \; 3.0387)$$

11. $n = 12$ employees.

$\Sigma x = 2{,}13{,}200$

$\Sigma x^2 = 4{,}919{,}860{,}000$

$\bar{X}$ = sample mean = 17,767

s = Sample Std. Dev. = 10,144

new $\Sigma x_1 = 213200 - 11000 - 48000 + 27500 + 31500$

$= 213200$

$$\therefore \bar{x}_1 = \frac{213200}{12} = 17{,}767$$

$$\Sigma x_1^2 = 4919860000 - (11000)^2 - (48000)^2 + (27500)^2 + (31500)^2$$

$$= 4243360000$$

Date

$$s_1^2 = \frac{1}{n-1}\left[\Sigma x_1^2 - n(\bar{x})^2\right] = 41396775.64$$

we donot observe any change in the sample mean, but a change in the sample variance is observed i.e. it decreases; signfying that the new data deviates less than the mean of the population.

12) i) The CRLB result holds true under very general conditions except when the conditions like the range of the distribution is involving parameter like in this case. This is caused due to the discontinuity in the derivative.

ii) we have $Y = \max X_i$

Since each X_i lies between 0 & θ,

For $0 < y < \theta$

$$P[Y<y] = P[\max X_i < y]$$
$$= P\left[\bigcap_{i=1}^{n} (X_i < y)\right]$$
$$= \prod_{i=1}^{n} P(X_i < y) \quad [\because X_i\text{s are independent}]$$
$$= \prod_{i=1}^{n}\left[\int_0^y \frac{1}{\theta}dx\right]$$
$$= \left(\frac{y}{\theta}\right)^n$$

Thus the pdf of y will be,

$$g_y(y) = \frac{d}{dy}P(Y<y) = \frac{n}{\theta}\cdot y^{n-1} \quad \text{for } 0<y<\theta$$

now, for any non-negative real no.

$\longrightarrow$

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

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

$= \frac{n}{n+k} \cdot \frac{\theta^{n+k}}{\theta^n} - 0$

$= \frac{n\theta^k}{n+k}$

ii) Bias of estimator $\hat{\theta}(c)$ is given as:

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

$= E[cY - \theta]$

$= c \cdot E(Y) - \theta$

$= c \cdot \frac{n\theta}{n+1} - \theta \quad (h=1)$

$= \left(\frac{cn}{n+1} - 1\right)\theta$

Mean sq. error of the estimator $\hat{\theta}(c)$ is given as

$MSE[\hat{\theta}(c)] = E[\{\hat{\theta}(c) - \theta\}^2]$

$= E[\{cY - \theta\}^2]$

$= E[c^2Y^2 - 2c\theta Y + \theta^2]$

$= c^2E(Y^2) - 2c\theta E(Y) + \theta^2$

$\therefore = c^2 \frac{n\theta^2}{n+2} - c\frac{2n\theta^2}{n+1} + \theta^2$ (using iii with $k=1, k=2$)

iv) For $\hat{\theta}(c)$ to be an unbiased estimator of θ, we need, $\text{Bias}(\hat{\theta}(c, y)) = 0$

$\left(\frac{c_y n}{n+1} - 1\right) \cdot \theta = 0$

$\boxed{c_y = \frac{n+1}{n}}$

Page No.
Date

v) In order to minimise MSE $[\hat{\theta}(c)]$, we need

$$0 = \frac{d}{dc} MSE[\hat{\theta}(c)]$$

$$\frac{d}{dc}\left[c^2 \cdot \frac{n\theta^2}{n+2} - \frac{c2n\theta^2}{n+1} + \theta^2\right]$$

$$= \frac{2cn\theta^2}{n+2} - \frac{2n\theta^2}{n+1}$$

Thus, $C_m = \frac{2n\theta^2}{n+1} \cdot \frac{n+2}{2n\theta^2} = \frac{n+2}{n+1}$

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

$$= \frac{2n\theta^2}{n+2} > 0$$

$\downarrow$

minimum

vi) In order to minimise order in estimater, it is preferred to get for an estimator which has lower mean square error among different competing estimators. So here $\hat{\theta}(C_n)$ will be preferred over $\theta(C_y)$.

As n becomes large $\hat{\theta}(C_y) = \frac{n+1}{n} \cdot Y \to Y$

Similarly $\hat{\theta}(C_m) = \frac{n+2}{n+1} Y \to Y$

Thus the 2 estimators are same, they become one as $n \to \infty$.