Probability & Statistics Assignment - 1

1] $\dfrac{\hat{\lambda}-\lambda}{\sqrt{\hat{\lambda}/n}} \sim N(0,1)$

$= \hat{\lambda} = \dfrac{\Sigma x_i}{n}$

$\hat{\lambda} = 200$

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

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

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

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

95% confident interval of λ when $\hat{\lambda} = 200$ & $n = 1$,

95% CI for $\lambda = (200 - 1.96\sqrt{200},\ 200 + 1.96\sqrt{200})$

$= (172.28141,\ 227.7185)$

2] X_i, $i = 1, 2, \ldots, n$ are i.i.d

mean $= \mu$

variance $= \sigma^2$

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

(i) $E(\bar{x}) = E\left(\dfrac{\Sigma x_i}{n}\right) = \dfrac{\Sigma E(x_i)}{n}$

$= \dfrac{n\mu}{n}$

$= \mu$

$$V(\bar{x}) = V\left(\frac{\sum X_i}{n}\right)$$

$$= \frac{1}{n^2} V(\sum X_i)$$

$$= \frac{1}{n^2} \sum V(X_i)$$

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

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

(ii) To show $E(S^2) = \sigma^2$

$$E(S^2) = E\left[\frac{1}{n-1}(\sum X_i - n\bar{x}^2)\right]$$

$$= \frac{1}{n-1}\left[E(\sum X_i^2) - nE(\bar{x}^2)\right]$$

$$= \frac{1}{n-1}\left[\cancel{E}\sum E(X_i^2) - nE(\bar{x}^2)\right]$$

Now, $V(X) = \sum E(X_i^2) - [E(X)]^2$

$$E(X_i^2) = \mu^2 + \sigma^2$$

$$V(\bar{x}) = E(\bar{x}^2) - [E(\bar{x})]^2$$

$$E(\bar{x}^2) = \mu^2 + \frac{\sigma^2}{n}$$

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

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

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

$$= \sigma^2$$

(iii) Now, $X \sim N(\mu, \sigma^2)$

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

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

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

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

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

(iv) We need, $P(0.5\sigma^2 < s^2 < 1.5\sigma^2) = P(s^2 < 1.5\sigma^2) - P(s^2 < 0.5\sigma^2)$

$n = 10$ & $\sigma^2 = 100$

$$P(s^2 < 1.5\sigma^2) = P(s^2 < 150)$$

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

$$P(\chi^2_9 < 13.5)$$

$$= 0.8587$$

$$P(s^2 < 0.5\sigma^2) = P(s^2 < 50)$$

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

$$= P(\chi^2_9 < 4.5)$$

$$= 0.1245$$

$$P(50 < s^2 < 150) = 0.8587 - 0.1245$$

$$= 0.7342$$

3] Let no. of claims be a random variable X

$X \sim Bin(4, p)$

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

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

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

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

$\log L = \log(\text{constant}) + 117 \log p + 603 \log(1-p)$

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

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

$\frac{dl}{dp} = 0$

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

$117 - 117p = 603p$

$720p = 117$

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

$\hat{p} = 0.1625$.

(iii) Goodness of fit test

$H_0 \rightarrow$ no. of claims conforms to the binomial model

$H_1 \rightarrow$ no. of claims does not confirm to the binomial model.

	0	1	2	3	4
O_i	86	75	16	2	1
E_i	88.542	68.724	19.998	2.574	0.108

$p = 0.1625$

$P(x=0) = {}^4C_0\, p^0 (1-p)^4 = 0.4818$

$P(x=1) = {}^4C_1\, p^1 (1-p)^3 = 0.3818$

$P(x=2) = 0.1111$

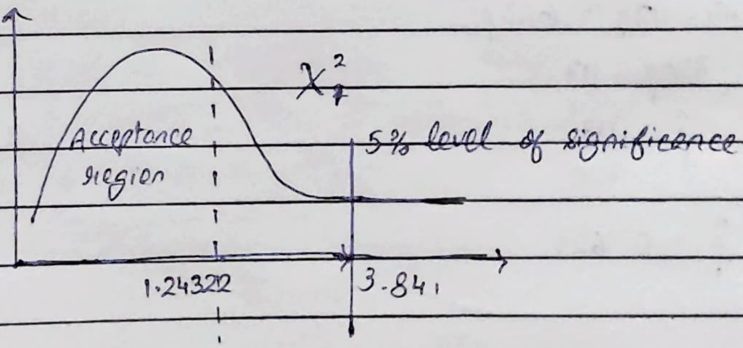
$P(x=3) = 0.0143$

$P(x=4) = 0.0006$

O_i	86	75	19
E_i	88.542	68.724	22.68

$$T.S. = \sum \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{5-1-1-2}$$

$$1.24322 \sim \chi^2_1$$



We have sufficient evident to reject H_0 at 5% level of significance.

$\therefore$ No. of claims conforms to the binomial model.

Page No.
Date

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

(i) Comparing the above $f(x)$ to P.D.F of Pareto distribution (2 parameter)

for pareto, P.D.F $= \frac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}$

Thus, β corresponds to α

So, $c = 3$ [as c corresponds to α here]

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

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

& $V(x) = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-1)} =$

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

(ii) $E(x) = \frac{\beta}{2}$

$E(x) = \bar{x}$

$\frac{\beta}{2} = \bar{x}$

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

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

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

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

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

$= \frac{2}{n} \times \frac{n\beta}{2} - \beta$

$= \beta - \beta$

$= 0$; $\therefore$ unbiased.

$$MSE = V(\hat{\beta}) + Bias^2$$
$$= V(\hat{\beta})$$
$$= V(2\bar{x})$$
$$= 4V(\bar{x})$$
$$= 4V\left(\frac{\sum X_i}{n}\right)$$
$$= \frac{4}{n^2} \sum V(X_i)$$
$$= \frac{4}{n^2} \times \frac{3\beta^2 n}{4}$$
$$= \frac{3\beta^2}{n}$$

As n increases, MSE tends to 0
$\therefore \beta$ is consistent

(iii) $\hat{\beta} = b\bar{x}$
$$E(\hat{\beta}_2) = E(b\bar{x})$$
$$= bE(\bar{x})$$
$$= bE\left(\frac{\sum x_i}{n}\right)$$
$$= \frac{b}{n} \sum E(X_i)$$
$$= \frac{b}{n} \times \frac{\beta n}{2}$$
$$= \frac{b\beta}{2}$$

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

$$= b^2 V\left(\frac{\sum X_i}{n}\right)$$

$$= \frac{b^2}{n^2} \sum V(X_i)$$

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

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

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

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

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

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

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

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

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

let $\frac{dM}{db} = 0$

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

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

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

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

$$b = 2/(3+n/n) = 2/(1+3/n)$$

Page No.
Date

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

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

$\therefore$ Minimum

$\therefore b = \frac{2}{(1+3/n)}$ minimises the MSE

(iv) $E(\hat{\beta}_2) = \frac{b\beta}{2}$

$$= \frac{2}{(1+3/n)} \times \frac{\beta}{2}$$

$$= \frac{n\beta}{n+3}$$

$$\text{Bias} = E(\hat{\beta}_2) - \beta$$

$$= \frac{n\beta - \beta(n+3)}{n+3}$$

$$= \frac{n\beta - n\beta - 3\beta}{n+3}$$

$$= \frac{-3\beta}{n+3}$$

Thus, we have a negative bias & hence ~~it~~ it underestimates the parameters

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

$$= \frac{\beta^2}{4n}[b^2(n+3) + 4n(1-b)]$$

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

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

$$= \beta^2\left[\frac{4n^2 + \cancel{3n}\, 12n - 4n^2}{n+3}\right]$$

$$= \frac{12n\beta^2}{n+3}$$

As $n \to \infty$, MSE does not tend to 0

$\therefore$ inconsistent

5] (i) $X \sim U(a, b)$

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

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

$$a = 2\bar{x} - b \quad \text{—— (i)}$$

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

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

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

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

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

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

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

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

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

$\hat{b} = \bar{X} + \sqrt{3}S$

(iii) Sample of observation : 1, 2, 8, 50, 150

$\bar{X} = 42.2$

$S = 63.57$

$\hat{a} = \bar{X} - \sqrt{3}S$

$= -67.906$

$\hat{b} = \bar{X} + \sqrt{3}S$

$= 152.3064$

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

(iv) $X \sim U(0, b)$

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

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

$\log L = -n \log b$

$\frac{dL}{db} = \frac{-n}{b}$

$\frac{dL}{db} = 0$; given $b = \infty$ which is not possible

&, $\frac{d^2L}{db^2} = \frac{n}{b^2} > 0$

$\therefore$ We have a minimum likelihood estimate

$\therefore \hat{b} = \max(x_i)$

6] $H_0 \to p = 0.1$

$H_1 \to p > 0.1$

(i) P(type I error) = P(rejecting H_0 | H_0 is true)

Let X be the no. of hits in first 12 missile & Y be the no. of hits in slep 12 missiles

$$P(\text{type I error}) = P(X \geq 3 | p = 0.1) + P(X = 0 | p = 0.1) P(Y \geq 5 | p = 0.1)$$
$$+ P(X = 1 | p = 0.1) P(Y \geq 4 | p = 0.1) + P(X = 2 | p = 0.1)$$
$$P(Y \geq 3 | p = 0.1)$$
$$= (1 - 0.8891) + (0.2824)(1 - 0.9957) + (0.6590 - 0.2824)$$
$$(1 - 0.9744) + (0.8891 - 0.6590)(1 - 0.8891)$$
$$= 0.14727 \approx 15\%$$

(ii) Probability of rejecting null hypothesis when $p = 0.3$

$$= P(X \geq 3 | p = 0.3) + P(X = 0 | p = 0.3) P(X \geq 5 | p = 0.3) + P(X = 1 | p = 0.3)$$
$$P(X \geq 4 | p = 0.3) + P(X = 2 | p = 0.3) P(Y \geq 3 | p = 0.3)$$
$$= (1 - 0.25828) + 0.0138(1 - 0.7237) + (0.0850 - 0.0138)(1 - 0.4925)$$
$$+ (0.2528 - 0.0850)(1 - 0.2528)$$
$$= 0.9125 \approx 91\%$$

(iii) Probability of type II error

Accepting H_0 when H_0 is false

$= 1 - 0.91253$

$= 0.08747$

$\approx 9\%$ (where $p = 0.3$)

(iv) 10 space agencies in aggregate fired 120 missiles & recorded 40 hits

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

$n = 120,\ X = 40$

$\hat{p} = \frac{40}{120} = \frac{1}{3} = 0.3333$

$Z_{10\%} = 1.2816$

lower bound of 90% right tailed confidence interval is,

$$\hat{p} - 1.2816\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.3333 - 1.2816\sqrt{\frac{0.3333 - 1.2816}{120}}$$

$$= 0.2782$$

(v) $p > 0.2782$ at 10% LOS

$H_0 \to p = \cancel{0.27}\ 0.278$ vs $H_1 \to p > 0.278$

One sample binomical model

$$\frac{X - np_0}{\sqrt{np_0 q_0}} \sim N(0,1)$$

$$\frac{39.5 - 120(0.278)}{\sqrt{120 \times 0.278 \times 0.722}}$$

$= 1.2511$

We don't have sufficient evidence to reject H_0 at 10% level of significance.

(vi) Lower bound of confidence implies that a 10% change of $p \le 0.2782$ whereas from the hypothesis test 'p' could be less than or equal to 0.2780 with probability more than 10%

(vii) $H_0: \mu_1 = \mu_2$ & $H_1: \mu_1 \neq \mu_2$

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

Given $\bar{x}_1 = 65$ $\bar{x}_2 = 70$

$S_1 = 54$ $S_2 = 70$

$n_1 = 12$ $n_2 = 15$

$S_p^2 = 1/25\,(11(2916) + 14(4900))$

$= 4027.04$

$$= \frac{65 - 70}{65.459\sqrt{\frac{1}{12} + \frac{1}{15}}}$$

$= -0.2034$

$\therefore$ There is ± 2.060 $(= t_{25,\,2.5\%})$.

So, we have insufficient evidence to reject H_0 at 5% in level.

Page No. | Date

7) (i) Public

$\bar{x} = 30$

$s_1^2 = 64.3125$

Private

$\bar{x}_2 = 35.0555$

$s_2^2 = 67.4027$

(ii) 2 sided F-test can be applied in case the samples come from populations with equal variance.

We are testing $H_0 \to \sigma_1^2 = \sigma_2^2$ vs $H_1 \to \sigma_1^2 \neq \sigma_2^2$

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

value of statistic is $\frac{64.31}{67.42} = 0.9542$

$F_{8,8}$ values at 5% levels are ~~0.25~~ 0.2256 & 4.433. Since, the value of the test statistics b/w the above values, we have insufficient evidence to reject the hypothesis & conclude that the population variances are equal.

(iii) $H_0 \to \mu_{Pri} = \mu_{Pu}$ & $H_1 \to \mu_{Pri} > \mu_{Pu}$

$$T.S = \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)}{s_p \sqrt{1/n_1 + 1/n_2}}$$

$$= \frac{35.0555 - 30}{8.1152\sqrt{2/9}}$$

$= 1.32151$

Tabulated value $= t_{16, 5\%}$

$= 1.746$

Since, T.S cal < tabulated value

$\therefore$ we can reject H_0.

8] (i) Unbiased estimator

- ~~Th~~ An unbiased estimator is an estimator where expected value is equal to the parameter.
- Estimator is said to be ~~unlock~~ unbiased if bias is equal to zero for all values of parameter θ.

(ii) Biased → A biased estimator is one which delivers an estimate which is consistently different from the parameter to be estimated.

(iii) The mean square error (MSE) of an estimator $\hat{\theta}$ is defined by,

$$MSE(\hat{\theta}) = var(\hat{\theta}) + Bias^2$$

(iv) The estimator having less MSE is more consistent

(v) They both would be considered to be consistent when MSE minimises such that when, $n \to \infty$, $MSE \to 0$

(vi) ~~is~~
- Method of Moments
- Maximum Likelihood Estimate.

9] (i) Variable t_K is defined as $t_K \Rightarrow \frac{N(0,1)}{\sqrt{\chi^2_K / K}}$

where, K denotes the degree of freedom

& the 2 random variables $N(0,1)$ & χ^2_K are independent.

(ii) For $K > 2$

$$E(t_K) = 0$$

$$V(t_K) = \frac{K}{K-2}$$

(iii) a] $\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1}$

$P\left(t_{9,0.5\%} < \frac{\bar{x}-\mu}{s/\sqrt{n}} < t_{9,99.5\%}\right) = 0.99$

$P\left(\bar{x} - t_{9,0.5\%}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{9,99.5\%}\frac{s}{\sqrt{n}}\right) = 0.99$

& $t_{9,0.5\%} = 3.250$

So, 99% Confidence interval for $\mu = \left(50 - 3.25\sqrt{\frac{48.667}{\sqrt{10}}}, 50 + 3.25\sqrt{\frac{48.667}{\sqrt{10}}}\right)$

$= (42.83, 57.17)$

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

Critical values level of confidence 2.58

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

$= (43.54, 56.45)$

10] $n = 1000$

$\lambda = 3$

$X \sim Poi(n\lambda)$

$X \sim Poi(3000)$

(i) P(Type I error) = P(reject H_0 | H_0 is true)

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

Using normal approx,

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

$= 1 - P(Z < 1.825)$

$= 3.39\%$

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

(iii) Power of test = P(rejecting H_0 | H_0 is false)

$= P[X > 3100 \mid H \sim Poi(n\lambda) - N(n\lambda, n\lambda)]$

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

The value of power of test will depend on the value of parameter under alternative hypothesis.

(iv) $\hat{\mu}$ is estimator of μ, $\hat{\mu}$ follows $N\left(\mu, \frac{\hat{\mu}}{2n}\right)$

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

or, $P\left(-2.57858 < \dfrac{2.9 - \mu}{\sqrt{2.9/1000}} < 2.5758\right) = 0.99$

Hence, the confidence interval for μ is $(2.7613, 3.0387)$

11] $n = 12$ $\quad \Sigma x^2 = 4919860000$

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

$S = 10144$

New sample mean $\rightarrow \Sigma x' = \Sigma x - 11000 - 48000 + 27500 + 31500$

$= \cancel{23}\ 213200$

$\bar{x}' = \frac{213200}{12}$

$= 17767$

New sample standard deviation $\rightarrow$

$\Sigma x^{2'} = 4919860000 - (11000)^2 - (48000)^2 + (27500)^2 + (31500)^2$

$= 4243360000$

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

$= \frac{1}{11}\left(4243360000 - 12(\cancel{1767}\ 17767)^2\right)$

$= \cancel{41396}\ 41396775.64$

$S' = 6434.032611$

The sample mean remains the same after the 2 ~~en~~ parameter employees as the 2 values removed were outliers & the 2 values included were around the sample mean.

The sample standard deviation reduces as the 2 values now included do not deviate very much from the sample mean causing the variance to reduce & hence the reduction of the sample standard ~~de~~ deviation.

12] (i) The Crámer-Rao Lower Bound (CRLB) result holds under very general conditions except where the range of the distributions involves the parameters, such as the uniform distribution in this case.
This is due to a discontinuity, so the derivative in the formula does not make sense.

(ii) We have $Y = \max X_i$

Since each X_i lies b/w 0 & θ, the support for Y will be 0 & θ

for $0 < y < \theta$

$$P(Y < y) = P[\max X_i < y]$$

$$= P[\prod_{i=1}^{n} (X_i < y)]$$

$$= \prod_{i=1}^{n} P(X_i < y) \quad \because X_i \text{ are independent}$$

$$= \prod_{i=1}^{n} \int_0^y \frac{dx}{\theta}$$

$$= \left(\frac{y}{\theta}\right)^n$$

$\therefore$ the probability density function of Y will be,

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

Now, for any non-negative real no. K,

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

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

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

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

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

$$\text{Bias}(\hat{\theta}(c)) = E(\hat{\theta}(c) - \theta)$$
$$= E[cY - \theta]$$
$$= cE(Y) - \theta$$
$$= \frac{cn\theta}{n+1} - \theta \quad [\because \text{(ii) with } K=1]$$
$$= \left(\frac{cn}{n+1} - 1\right)\theta$$

MSE of $\hat{\theta}(c)$ is given as →

$$MSE[\hat{\theta}(c)] = E[(\hat{\theta}(c) - \theta)^2]$$
$$= E[(cY - \theta)^2]$$
$$= E[c^2Y^2 - 2c\theta Y + \theta^2]$$
$$= c^2E(Y^2) - 2c\theta E(Y) + \theta^2$$
$$= c^2\frac{n\theta^2}{n+2} - \frac{2c\,2n\theta^2}{n+1} + \theta^2 \quad [\because \text{(ii) with } K=1 \,\&\, 2]$$

(iv) For $\hat{\theta}(c)$ to be an unbiased estimator of θ, we need →

$$\text{Bias}(\hat{\theta}(c)) = 0$$
$$\left(\frac{cn}{n+1} - 1\right)\cdot\theta = 0$$
$$\Rightarrow c = \frac{n+1}{n}$$

Page No.
Date

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

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

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

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

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

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

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

∴ Minimum

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

As n becomes large, $\hat{\theta}(c_u) = \frac{n+1}{n} Y \to y$

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

∴ the 2 estimator becomes one & same as $n \to \infty$.