Date

Probability and Statistics Assignment

Ques 1

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

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

$$\hat{\lambda} = 200$$

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

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

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

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

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

$$\left(200-1.96\sqrt{200},\ 200+1.96\sqrt{200}\right)$$

$$(172.28141,\ 227.7185)$$

Date

Ques⇒ x_i: $i = 1, 2, \ldots, n$ are n

Mean = μ variance = σ^2

$\bar{x} = \frac{\Sigma x_i}{n}$ (Given) , $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

1. $E[\bar{x}] = E\left[\frac{\Sigma x_i}{n}\right]$

$= \frac{1}{n} E[\Sigma x_i]$

$= \frac{1}{n} E[x_1 + x_2 + x_3 + \cdots + x_n]$

$= \frac{1}{n} [E[x_1] + E[x_2] + \cdots + E[x_n]]$

$= \frac{1}{\cancel{n}} \times \cancel{n}\mu = \mu$

2 $V[\bar{x}] = V\left[\frac{\Sigma x_i}{n}\right] = \frac{1}{n^2} V[\Sigma x_i]$

$= \frac{1}{n^2} [V[x_1 + x_2 + \cdots + x_n]]$

[∵ Independ

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

$= \sigma^2/n$

Date

(ii) To show $E[s^2] = \sigma^2$

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

$$= \frac{1}{n-1} E\left[\sum x_i^2 - n\bar{x}^2\right]$$

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

$$= \frac{1}{n-1}\left[E[x_1^2 + x_2^2 + \dots + x_n^2] - nE[\bar{x}^2]\right]$$

since	since
$E[x^2] = V(x) - [E[x]]^2$	$E[\bar{x}^2] = V(\bar{x}) - [E[\bar{x}]]^2$
$E[x^2] = \sigma^2 - \mu^2$	$= \sigma^2/n - \mu^2$

$$= \frac{1}{n-1}\left[E[x_1^2] + E[x_2^2] + \dots + E[x_n^2] - nE[\bar{x}^2]\right]$$

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

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

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

Date

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

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

$\cancel{E[s^2] = E s^2}$

$$V[s^2] = V[s^2]$$

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

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

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

$$= \frac{2\sigma^4}{n-1}$$

(iv)

$$P[-0.5\sigma^2 < s^2 < 0.5\sigma^2]$$

$$P[-0.5(100) < s^2 < 0.5(100)]$$

$$P[-50 < s^2 < 50]$$

$$P\left[-\frac{50(n-1)}{\sigma^2} < \frac{s^2(n-1)}{\sigma^2} < \frac{50(n-1)}{\sigma^2}\right]$$

$$P\left[-\frac{\cancel{50}\times 9}{\cancel{100}_2} < \chi^2_9 < \frac{\cancel{50}\times 9}{\cancel{100}_2}\right]$$

$$P[-4.5 < \chi^2_9 < 4.5]$$

$$= P[\chi^2_9 < 4.5] - P[\chi^2_9 \leq -4.5]$$

$$= 0.1245$$

Date

Ques 3

$X \sim \text{Binom}(4, p)$

$\underline{x} = 180$

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

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

$$= (1-p)^{86n+75n-75+16n-32+2n-6+n-4} \times p^{75+32+6+4} \times \text{constant}$$

$$L = (1-p)^{180n-117} p^{117} \cdot \text{constant}$$

taking log on both sides

$$\log L = (180n-117)\log(1-p) + 117\log p + \log \text{const}$$

diff w.r.t p

$$\frac{dl}{dp} = -\frac{(180n-117)}{1-p} + \frac{117}{p} = 0$$

$$\frac{117}{p} = \frac{180n-117}{1-p}$$

$$117 - \cancel{117p} = 180np - \cancel{117p}$$

$$\frac{117}{180n} = p$$

$$\frac{117}{180 \times 4} = \hat{p}$$

$$\hat{p} = 0.1625$$

#

Date

Goodness of fit test

HO : No. of claims conforms to the binomial model

H1 : No. of claims does not conforms to the binomial model.

	0	1	2	3	4
Oi	86	75	16	2	1
Ei	88.542	68.724	19.998	2.574	0.108

$$P(X=0) = {}^4C_0\, p^0 (1-0.1625)^4 = 0.4919$$

$$P(X=1) = {}^4C_1 (0.1625)(1-0.1625)^3 = 0.3818$$

$$P(X=2) = {}^4C_2 (0.1625)^2(1-0.1625)^2 = 0.1111$$

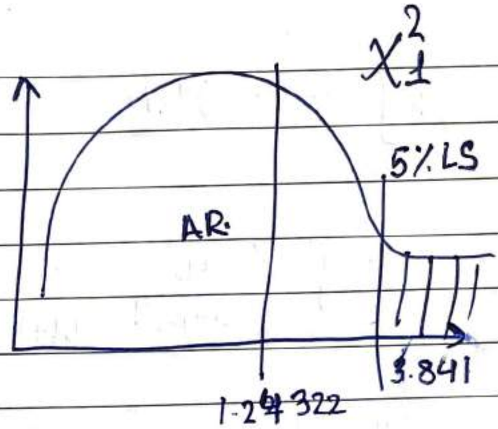
$$P(X=3) = {}^4C_3 (0.1625)^3(1-0.1625)^1 = 0.0143$$

$$P(X=4) = {}^4C_4 (0.1625)^4(1-0.1625)^0 = 0.0006$$

Oi	86	75	19
Ei	88.542	68.724	22.68

$$T.S = \frac{\Sigma (Oi-Ei)^2}{Ei} \sim \chi^2_{5-1-1-2}$$

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



We have insufficient evidence to reject Ho at LS 5%.
Therefore, No. of claims conforms to the binomial model.

Ques 4 =>

$$f(x) = \frac{CB^3}{(x+B)^4} \; : \; x>0 \, , \, B>0$$

C is a constant B is a parameter

$$\int_0^\infty \frac{CB^3}{(x+B)^4} dx = 1$$

$$CB^3 \int_0^\infty (x+B)^{-4} dx = 1$$

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

$$\frac{CB^3}{-3} \left[0 - (B)^{-3} \right] = 1$$

$$\frac{C\cancel{B^3}}{3} \times \frac{1}{\cancel{B^3}} = 1$$

$$\boxed{C=3}$$

Date

$$E[x] = 3\beta^3 \int_0^\infty x(x+\beta)^{-4}\,dx$$

$$= 3\beta^3 \int_0^\infty \left(\frac{x+\beta}{(x+\beta)^4} - \frac{\beta}{(x+\beta)^4} \right) dx$$

$$= 3\beta \int_0^\infty \left(\frac{1}{(x+\beta)^3} - \frac{\beta}{(x+\beta)^4} \right) dx$$

$$= 3\beta \int_0^\infty (x+\beta)^{-3}\,dx - 3\beta^2 \int_0^\infty (x+\beta)^{-4}\,dx$$

$$f(x) = \frac{3\beta^3}{(\beta+x)^{3+1}} \qquad , \beta > 0$$

This follows pareto distribution

$$\alpha = 3 \qquad \lambda = \beta$$

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

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

#

Date

ii Method of moments

$\underline{X} = x_1, x_2, x_3, x_4, \ldots, x_n$

Sample mean $= \frac{\Sigma x_i}{n} = \overline{x}$

Distribution mean $= B/2$

mom states that,

sample mean = distribution mean

$\overline{X} = B/2$

$$\boxed{2\overline{x} = \hat{B}}$$

Unbaisedness,

$$E[\hat{\theta}] = \theta$$

$$E[\hat{B}] = E[2\overline{x}]$$

$$= 2E[\overline{x}]$$

$$= 2E\left[\frac{\Sigma x_i}{n}\right]$$

$$= \frac{2}{n} E[x_1 + x_2 + x_3 + \cdots + x_n]$$

$$= \frac{2}{n}\left[E[x_1] + E[x_2] + \cdots + E[x_n]\right]$$

$$= 2/n \left[B/2 + B/2 + \cdots B/2\right]$$

$$= \cancel{2}/\cancel{n} \times \cancel{n} B/\cancel{2} = B$$

Yes it is unbaised

For consistency

$$M.S.E(\hat{B}) = Var(\hat{B}) + Bias^2(\hat{B})$$

$$= Var(2\overline{x}) + 0$$

$$= 4 Var\left(\frac{\Sigma x_i}{n}\right)$$

$$= \frac{4}{n^2} Var(x_1 + x_2 + \cdots + x_n)$$

$$= \frac{4}{n^2}\left[Var(x_1) + Var(x_2) + \cdots + Var(x_n)\right]$$

$$= \cancel{4}/n^{\cancel{2}} \times \frac{3 \cancel{n} B^2}{\cancel{4}}$$

$$= \frac{3B^2}{n} \text{ as } n \to \infty \text{ MSE minimizes}$$

So it is consistent

Date

(iii) $MSE(b\bar{X}_n) = b^2 Var(\bar{X}_n) + Bias^2(b\bar{X}_n)$

$$= b^2 \, var\left(\frac{\Sigma x_i}{n}\right) + \left(Bias\,(b\bar{X}_n)\right)^2$$

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

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

$$MSE(b\bar{X}_n) = \frac{B^2}{4}\left(\frac{3b^2}{n} + b^2 + 4 - 4b\right)$$

differentiation with respect to b

$$\frac{d}{db} MSE(b\bar{X}_n) = \frac{B^2}{4}\left(\frac{6b}{n} + 2b - 4\right) = 0$$

$$= \frac{6b}{n} + 2b - 4 = 0$$

$$\frac{3b}{n} + b - 2 = 0$$

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

$$b\left(\frac{3}{n} + 1\right) = 2$$

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

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

$n \to \infty$
MSE $\to$

As in part (ii) $n \to \infty$ MSE becomes 0 but this is not applied in this because as $n \to \infty$ MSE is not defined.

Date

Ques 5 ⇒ $X \sim U(a,b)$

Sample mean = $\bar{x}$
Sample stan dev = S

$E[X] = \frac{a+b}{2}$ $\qquad V(X) = \frac{(a-b)^2}{12}$

Sample mean = Population mean

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

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

sample var = Population variance

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

$$(12S)^{1/2} = a-b$$

$$a+b = 2\bar{x}$$
$$-(a-b) = 2\sqrt{3S^2}$$

$$2a = 2\bar{x} - 2S\sqrt{3}$$
$$\boxed{\hat{a} = \bar{x} - \sqrt{3}S}$$

$$b = 2\bar{x} + a$$
$$\hat{b} = 2\bar{x} + \hat{a}$$
$$\hat{b} = 2\bar{x} - \bar{x} + \sqrt{3}S$$
$$\boxed{\hat{b} = \bar{x} + \sqrt{3}S}$$

#

Date

(iii) Sample of 5 observations.

1, 2, 8, 50, 150

$\frac{\sum x_i}{n} = 42.2$ $\quad \sigma = 63.57$

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

$= 42.2 - \sqrt{3}(63.57$ $\quad \cancel{= 152.30}$

$= -67.906$

$\hat{b} = 152.3064$

$E[X] = \frac{\hat{a}+\hat{b}}{2} = 42.2$

$V(X)$

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

(iv)

$X \sim U(0, b)$

$0 < x < b$

$f(x) = \frac{1}{b}$

$L = \frac{1}{b} \times \frac{1}{b} \times \frac{1}{b} \times - \frac{1}{b}$

$L = \frac{1}{b^n}$

$L = b^{-n}$

$l = -n \log b$

$\frac{dl}{db} = \frac{-n}{b} = 0$

#

$\theta < x < \theta$ $\quad x \le \max|x_i|$

Date

$$\hat{\theta} = \max(x_i)$$

Ques6 ⇒

$$H_0 : p = 0.1$$
$$H_1 : p > 0.1$$

I. P(Type I error) = P(Rejecting H_0 / H_0 is true)

X be the no. of hits in first 12 missile and Y be no. of hits in step 12 missiles.

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

$$= (1-0.8891) + 0.2824\,(1-0.9957) + (0.6590-0.2824)(1-0.9744) + (0.8891-0.659)(1-0.889)$$

$$= 0.14727 \approx 15\%$$

II) Prob. of rejecting null hypothesis when $p=0.3$

$$= P(X \geq 3 \mid p=0.3) + P(X=0 \mid p=0.3)\,P(Y \geq 5 \mid p=0.3) + P(X=1 \mid p=0.3)\,P(Y \geq 4 \mid p=0.3) + P(X=2 \mid p=0.3)\,P(Y \geq 3 \mid p=0.3)$$

$$= (1-0.2528) + 0.0138\,(1-0.7237) + (0.0850-0.0138)(1-0.4925) + (0.2528-0.085)(1-0.2528)$$

$$= 0.9125 \approx 91\%$$

(iii) Prob. of type II error

Accepting H_0 when H_0 is false

$= 1-0.91253 = 0.08747 \approx 9\%$ when $(p=0.3)$

Date

(iv) 10 Space agencies in aggregate fired 120 missiles and recorded 40 hits

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

$n = 120 \quad x = 40 \qquad \hat{p} = \frac{40}{120} = 1/3 = 0.3333$

$Z_{10\%} = 1.2816$

Lower bound of 90% right tailed confidence interval for p is

$$\hat{p} - 1.2816\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = 0.3333 - 1.2816\sqrt{\frac{0.333(1-0.333)}{120}} = 0.2782$$

(v) $p > 0.278$ at 10% LOS

$H0 :- p = 0.278$ vs $H_1\ p > 0.278$

One sample Binomial model

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

$$\frac{39.5 - 120(0.278)}{\sqrt{120(0.278)(0.722)}} = 1.2511$$

we dont have sufficient evidence to reject H0 at 10% LS.

vi Lower bound of confidence interval implies that there is only a 10% chance of $p \le 0.2782$ whereas from the hypothesis test, 'p' could be less than or equal to 0.2780 with prob more than 10%.

#

Date

(vii) $H_0: \mu_1 = \mu_2$ vs $\mu_1 \neq \mu_2$

$$P = \theta \qquad \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

Given,

$\bar{X}_1 = 65$; $\bar{X}_2 = 70$, $S_1 = 54$, $S_2 = 70$, $n_1 = 12$, $n_2 = 15$

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

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

There is $\pm$ 2.060 $(= t_{25;2.5\%})$.

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

Ques 7 ⇒

(i)

PUBLIC		PRIVATE	
Sample mean	= 30	Sample mean	= 35.0555
Sample variance	= 64.3125	Sample variance	= 67.4027

(ii) The primary condition show that the data should be reasonably random. There should be little skewness and no outliers in the data for each sample.

(iii)

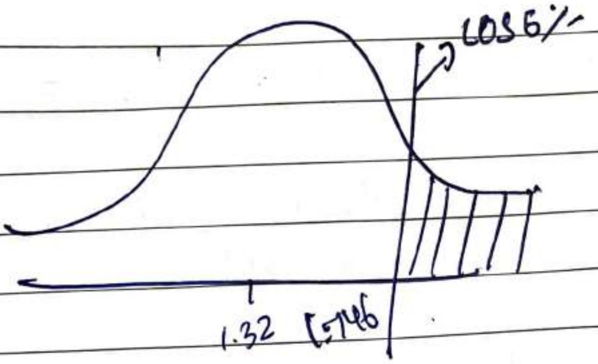
$H_0: \mu_{pu} = \mu_{pr}$

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

$$T.S = \frac{(\bar{X}_{pr} - \bar{X}_{pu}) - (\mu_{pr} - \mu_{pu})}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{16}$$

$$= \frac{35.0555 - 30}{8.1152\sqrt{2/9}} = 1.3215121 \#$$

Date



Hence we have insufficient evidence to reject H0.

Ques 8 =>

1 Unbiased estimator

=> An unbiased estimator is an estimator whose expected value is equal to the parameter

=> Estimator is said to be unbiased if its bias is equal to zero for all values of parameter θ.

2 Biased

=> A biased estimator is one which deliver an estimate which is consistently different from the parameter to be estimated.

3 The mean sq. error of an estimator $\hat{\theta}$ is defined by

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

iv The estimator having less MSE is more consistent

v They both would considered be consistent when MSE minimizes

such that when $n \to \infty$ $MSE \to 0$

vi method of moments.

maximum likelihood estimate

Ques9 ⇒ Date

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

ii) Mean and variance t_K for $K>2$ are 0
Variance $= \frac{K}{K-2}$

(iii) a) We know that for a sample from a normal population
$$\frac{\bar{x}-\mu}{S/\sqrt{n}} \sim t_{n-1}$$

For given interval $t_9 = 3.25$
confidence interval for μ is this

$$\left(50 - 3.25\sqrt{\frac{48.667}{10}}, 50+3.25\sqrt{\frac{48.667}{10}}\right)$$

i.e $(42.83, 57.17)$

b) From (ii)
$$t_{n-1} \sim N\left(0, \sqrt{\frac{n-1}{n-3}}\right) \sim N\left(0, \sqrt{9/7}\right)$$
i.e.
$$\frac{\bar{x}-\mu}{S/\sqrt{n}} \sim N(0, \sqrt{9/7})$$

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

Critical values level of confidence 2.58
confidence interval

$$\left(50 - 2.58\sqrt{\frac{48.667}{10} \times 9/7}, 50+2.58\sqrt{\frac{48.667}{10} \times 9/7}\right)$$
$(43.54, 56.45)$

#

Ques 10 ⇒

Date

$$n = 1000 \quad , \lambda = 3$$

$$X \sim Poi(n\lambda)$$
$$X \sim Poi(3000)$$

(i) $P(\text{Type I error}) = P(\text{Reject } H_0 / H_0 \text{ is true})$
$= P(X > 3100 / X \sim Poi(3000))$

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 $= P[\text{Rejecting } H_0 / H_0 \text{ is false})$

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

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

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

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

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

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

#

Ques11

$n = 12$ emp.

$\Sigma x = 213200$

$\Sigma x^2 = 4919860000$

Sample mean $= 17767$

Sample std dev $= 10144$

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

$= 213200$

Revised mean $= \frac{213200}{12} = 17767$

New samp S.D.

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

$$= 4243360000$$

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

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

$$= 41396775.64$$

$$s' = 6434.03261$$

The sample mean remains the same after the inclusion of the 2 permanent employees as the 2 values removed are outliers and the 2 values include were around sample mean. The SD reduces as the 2 values now include do not deviate very much from the sample mean cause variance to reduce and hence reduction of sample S.D.

Date

Ques 12

(i) The CRLB results hold under very general conditions except where the range of distribution involves the parameter, such as the uniform dist. in this case.

This is due to discontinuity derivate formula does not make sense.

ii we have

$$Y = \max_i X_i$$

since each X_i lies b/w 0 and θ, the support for Y will be 0 and θ

For $0 < y < \theta$

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

PDF will be

$$= \frac{n}{\theta^n} y^{n-1} \quad \text{for } 0<y<\theta$$

$$E[Y^K] = \int_0^\theta y^K \frac{n}{\theta^n} \cdot y^{n-1} \, dy$$
$$= \frac{n}{n+K} \int_0^\theta \frac{n+k}{\theta^n} \cdot y^{n+k-1} \, dy$$
$$= \frac{n}{n+K} \cdot \frac{\theta^{n+K} - 0}{\theta^n} = \frac{n\theta^K}{n+K}$$

#

Date

(iii) Bias of the estimator $\hat{\theta}(c)$

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

Mean square error.

$$MSE[\hat{\theta}(c)] = E[\hat{\theta}(c) - \theta]^2$$
$$= E[c^2Y^2 - 2c\theta Y + \theta^2]$$
$$= c^2 \cdot E[Y^2] - 2c\theta E[Y] + \theta^2$$
$$= \frac{c^2 n\theta^2}{n+2} - \frac{c\,2n\theta^2}{n+1} + \theta^2$$

(iv) For $\hat{\theta}(c)$ to be an unbias estimator θ we need

$$\text{Bias}[\hat{\theta}(c_u)] = 0$$

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

V In order to minimize $MSE[\hat{\theta}(c)]$ we need

$$0 = \frac{d}{dc} MSE[\hat{\theta}(c)] = \frac{d}{dc}\left[\frac{c^2 n\theta^2}{n+2} - \frac{c\,2n\theta^2}{n+1} + \theta^2\right]$$
$$= 2c\frac{n\theta^2}{n+2} - \frac{2n\theta^2}{n+1}$$

Thus

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

Date

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