Prob. & Stats Assignment.
Gargi Tendulkar
Sec. B Roll No. 77.

Q.1 $X \sim Poi(\theta)$

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

$$P[X_1 < \theta < X_2] = 0.95$$

$X \sim Poi(200)$

$$95\% \text{ CI} = \mu \pm 1.96\sqrt{\sigma}$$
$$= 200 \pm 1.96\sqrt{200}.$$

Hence, the interval is $(172.2814, 227.7186)$.

Q2 $\cancel{\bar{x} = \sum_{i=1}^{x} (x_i - \bar{x}) = \text{sample variance}}$

$\bar{x} =$

Q2 $\bar{x} = \sum_{i=1}^{x} x_i$ = sample mean.

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

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

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

(ii) $V(\bar{x}) = V\left(\frac{\Sigma x_i}{n}\right) = \frac{1}{n^2} V[E(\bar{x})] = \frac{n\sigma^2}{n^2} = \boxed{\frac{\sigma^2}{n}}$

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

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

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

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

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

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

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

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

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

$\therefore$ ~~(n-1) var (s~~

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

(iv) $n=10$; $\sigma^2 = 100$.

$\therefore$

$\frac{(n-1)s^2}{\sigma^2} = 0.5 \Rightarrow \frac{9\times s^2}{100} = 0.5$

$\therefore s^2 = \frac{500}{9} = \boxed{5.56\%}$

Q3 (i) $X \sim Bin(4, p)$

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

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

$$L = \left[\text{const.}\ \cancel{p^0} \cdot (1-p)^4\right]^{86} \cdot \left[\text{const.}\ p^1 (1-p)^3\right]^{75} \cdot \left[\text{const.}\ p^2 (1-p)^2\right]^{16} \cdot \left[\text{const}\ p^3 (1-p)^1\right]^{2} \cdot \left[\text{const.}\ p^4\, \cancel{(1-p)}\right]^{1}$$

$$\ln L = \left[(86\times 4)\ln(1-p)\right] \cancel{\left[(75\times3)\ln p + \ln(\right.}$$
$$\left[(75\times1)\ln p + (75\times3)\ln(1-p)\right] \cdot$$
$$\left[(16\times2)\ln p + (16\times2)\ln(1-p)\right] \cdot$$
$$\left[(2\times3)\ln p + (2\times1)\ln(1-p)\right] \cdot$$
$$\left[4\ln p\right]$$

$$\underline{\ln L} = \left[344\ln(1-p)\right] + \left[75\ln p + 225\ln(1-p)\right] + \left[4\ln p\right] + \left[32\ln p + 32\ln(1-p)\right] + \left[6\ln p + 2\ln(1-p)\right]$$

$\ln L = 633 \ln(1-p) + 117 \ln p.$

$$\frac{d \ln L}{dp} = \frac{633}{1-p} + \frac{117}{p}$$

$$0 = \frac{633p + 117 - 117p}{p(1}$$

$$0 = \frac{633}{1-p} + \frac{117}{p}$$

$$\therefore \frac{117}{p} = -\frac{633}{1-p}$$

$$\frac{1-p}{p} = \frac{-633}{117}$$

$$\therefore \frac{1}{p} - 1 = \frac{-633}{177}$$

$$\frac{1}{p} =$$

$117 \cancel{p} = 750p$

$$\therefore p = \frac{750p}{117}$$

$$\boxed{p = 0.156}$$

(ii) $p(x=0) = {}^4C_0\, p^0 (1-p)^4$
$= 91.34$

$P(x=1) = {}^4C_1\, p^1 (1-p)^3$
$= 67.53$

$P(x=2) = {}^4C_2\, p^2 (1-p)^2$
$= 18.72$

$P(x=3) = {}^4C_3\, p^3 (1-p)^1 = 2.31$

$P(x=4) = {}^4C_4\, p^4 (1-p)^0 = 0.11$

* If not reject on 5%, then go on 10% *

* If reject on 5% then make it more lineant, go on 2.5% *

$f(x)$	E_i	O_i	$(O_i-E_i)^2/E_i$
$P(x=0)$	91.84	86	0.3123
$P(x=1)$	67.53	75	0.8263
$P(x=2)$	18.72	16	0.3982
$P(x=3)$	2.31	2	0.0416
$P(x=4)$	0.11	1	7.2009
			8.7793

$$\boxed{\chi^2_{cal} = 8.7793;\ \chi^2_{tab} = 7.815}$$

$\because$ at 5%, $\chi^2_{cal} > \chi^2_{tab}$ $\therefore$ we reject.

at 2.5%, $\chi^2_{cal} < \chi^2_{tab}$ $\therefore$ do not reject.

Q4 $f(x) = \dfrac{c\beta^3}{(x+\beta)^4}$; $x>0$, $\beta>0$

$$\int_0^\infty f(x)\left[\frac{(x+\beta)^{-3}}{-3}\right]_0^\infty = \int f(x) = 1.$$

$$c\beta^3 \int_0^\infty \frac{1}{\ldots} = 1$$

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

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

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

$\therefore \boxed{c=3}$

$$\frac{\beta^3-(1+\beta)^3}{\beta^3(1+\beta)^3}$$

$$\therefore E(x) = \int_0^\infty x f(x)\,dx.$$

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

Date
Page

$$3B^3 \int_0^\infty \frac{x+B}{(x+B)^4} - \frac{B}{(x+B)^4} \, dx$$

$$= 3B^3 \left[\frac{(x+B)^{-2}}{-2} - \frac{B(x+B)^{-3}}{-3} \right]_0^\infty$$

$$= 3B^3 \left[0 - 0 - \left[\left(\frac{B^{-2}}{-2} \right) - \left(\frac{B^{-2}}{-3} \right) \right] \right]$$

$$= 3B^3 \left[\frac{B^{-2}}{2} - \frac{B^{-2}}{3} \right]$$

$$= \frac{B^3}{2} \left[B^{-2} \right] = \frac{B}{2}$$

$$\therefore E(X^2) = \int_0^\infty x^2 f(x) \, dx$$

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

Using Pareto Distribution:

$\frac{3B^2}{5} \Rightarrow V(X)$ where, $\alpha = 3$; $\lambda = B$

(ii) $B = x_1, x_2 \dots x_n$

$E(x) = \bar{x}$; $B/2 = \bar{x}$ $\therefore \boxed{\hat{B} = 2\bar{X}_n}$

$$E(\hat{B}) = E(2\bar{X}_n) = 2E\left(\frac{\Sigma x_i}{n}\right)$$

$$= \frac{2}{n} \Sigma [E(x_i)] = \frac{2}{n} \cdot n \cdot \frac{B}{2} = \boxed{B}$$

unbiased

Q5. $\cancel{\hat{a} = \bar{x} - \sqrt{3}s}$

(i) $X \sim U(a+b)$

$$\bar{x} = \frac{1}{2}(a+b)$$

$$s = \frac{1}{\sqrt{12}}(b-a).$$

$$\therefore \sqrt{12} + a = b$$

$$\bar{x} = \frac{1}{2}\left[a + \sqrt{12}s + a\right]$$

$$\bar{x} = \frac{1}{\cancel{2}}\left[\cancel{2}a + \cancel{2}\sqrt{3}s\right] \quad \therefore \boxed{\hat{a} = \bar{x} - \sqrt{3}s}$$

(ii) $\bar{x} = \frac{1}{2}(a+b)$

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

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

$$\bar{x} = \frac{1}{2}\left[b - \sqrt{12}s + b\right]$$

$$= \frac{1}{\cancel{2}}\left[\cancel{2}b - \cancel{2}\sqrt{3}s\right] \Rightarrow \boxed{\hat{b} = (\bar{x} + \sqrt{3}s)}$$

(iii) Sample → 1, 2, 3, 4, 5, 6.

$\bar{x} = 12$; $s = 21.27$.

$\hat{a} = 12 - \sqrt{3}(21.27) = -24.84$

$\hat{b} = 12 + \sqrt{3}(21.27) = 48.84.$

for $U(a,b)$ the prob. of a sample pt. being less than 'a' or greater than 'b' is zero. and we have a sample value so that

it ~~is~~ greater than our estimate 'b'.
It highlights a potential webinar of mom.

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

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

$$L = f(x_1) \cdot f(x_2) \cdots f(x_n)$$

$$= \frac{1}{b} \cdot \frac{1}{b} \cdots \frac{1}{b} = \frac{1}{b^n}$$

$$l = -n\log b.$$

$$\frac{dl}{db} = \frac{-n}{b} = 0 \; ; \quad b \to \infty$$

$$\frac{d^2 l}{db^2} = \frac{+n}{b^2} > 0 \; ; \quad \text{minimum at } b \to \infty.$$

so, using common sense,
$L(\theta)$ is maximum when θ is minimum $b > m$

$$\boxed{\hat{b} = \text{Max } x_i}$$

0.5074

Q6(i) P(type 1 error) = P(Reject H_0 / H_0 is True)
X = no. of hits in 1st step of 12 missles.
Y = no. of hits in 2nd step of 12 missles.

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

$$= (1 - 0.8891) + (0.2824)(1 - 0.9957)$$
$$+ (0.6590 - 0.2824)(1 - 0.9994)$$
$$+ (0.8891 - 0.6590)(1 - 0.8891) = 0.1472$$

classmate

Date ______
Page ______

Q6 (ii) $P(X \geq 3 \mid P = 0.3)$ ~~$P(X=0 \mid P=0.3)$~~

$+ P(X=0 \mid p=0.3) \cdot P(Y \geq 5, P=0.3)$

$+ P(X=1 \mid P=0.3) \cdot P(Y \geq 4 \mid P=0.3)$

$+ P(X=2 \mid P=0.3) \cdot P(Y \geq 3 \mid P=0.3)$

$= \boxed{0.91253 \approx 91\%}$

(iii) P(~~tp~~ type II error) is the prop. of accepting Ho when Ho is false.

$= 1 - 0.91253$

$= 0.08747 \approx \boxed{9\% (P = 0.3)}$

(iv) 10 space agencies fire 120 missiles and record 40 hits.

$X \sim Bin(120, 1/3)$; $\hat{p} = \frac{40}{120} = \frac{1}{3}$

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

$P\left(-1.2816 < \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < 1.2816\right)$

$\boxed{\text{At C.I. } 90\% \; p \in (0.3333, 0.2782)}$

$\approx \boxed{15\%}$

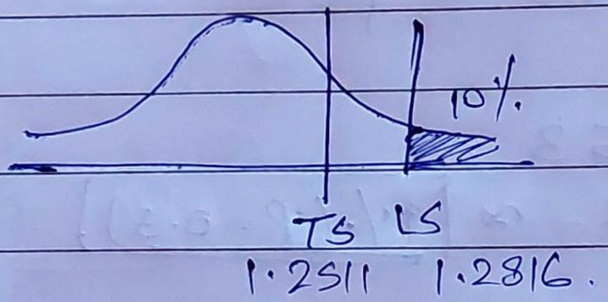
(v) $p > 0.278$ at 10% L.S.

$H_0 \Rightarrow \quad p = 0.278$

$H_1 \Rightarrow \quad p > 0.278$

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

$$\frac{39.5 - 120 \times 0.278}{\sqrt{120 \times 0.278 \times 0.722}} = \boxed{1.2511}$$



At T.S < 1.2816.

we have insufficient evidence to reject H_0 at 10%.

(vi) Lower Bound of CI implies that there is only 10% chances 'p' ≤ 0.2782, whereas, from H.T. 'p' could be less than equal to 0.2780 with prob more than 10%. (approx. 10.5% 1.2511)

The ~~minor~~ disconnect b/w CI and HI at same level is because of:

1) Applying continuity correction in H.T.
2) Use of ~~sample~~ sample pop. $\hat{p}$ to estimate pop. var.

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

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

$$S_p^2 = \frac{1}{25}(11 \times 2916 + 14 \times 4900) = 4027.04.$$

$\Rightarrow$ ~~PQ~~ $= \Rightarrow$ Reject H_0 at

classmate

Date
Page

Q7(i) private:

sample variance (S) = 8.2099.

sample mean. $(\bar{x})$ = 35.06

Public.

Sample variance (S) = ~~8~~ 8.0195

Sample mean $(\bar{x})$ = 30.

(ii) $H_0 \Rightarrow \bar{x_1} = \bar{x_2}$

$H_1 \Rightarrow$ ~~#~~ $\bar{x_1} \neq \bar{x_2}$

$$= \frac{30 - 35.06}{Sp. \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Where, $sp = \frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}$.

$$Sp = \frac{8(64.3124) + 8(67.4024)}{16} = 65.8574.$$

$$\therefore H_1 = \frac{30 - 35.06}{65.8574 \sqrt{\frac{1}{65.3124} + \frac{1}{68.4024}}}$$

H_1 = 1.3226.

Z_{tab} = (-2.624, 2.624).

∴ Do not reject H_0.

5%

(iii) Private hospitals will have less claims.

Q8.

(i) Unbiased estimator:

~~If we have a random sample~~ $\hat{\theta}$

$Eg(x) = \theta$

$Eg(x) = \mu \rightarrow g(E)$ is unbiased estimator.

(ii) If $E(\hat{\theta}) \neq \theta$ the ~~the~~ estimator is biased

(iii) MSE = Variance + Bias^2.

(iv) $MSE \propto \frac{1}{\text{efficiency}}$.

$\therefore \hat{\theta}$ is more efficient than $\tilde{\theta}$ since the MSE of $\hat{\theta}$ is less than $\tilde{\theta}$

(v) When the MSE is more towards 0 and the n increases, then the estimator would be more efficient.

(vi) methods of moments.
Maximum likelihood Estimator.

Q9 (i) $t_k \sim \dfrac{z}{\sqrt{\chi^2_k/k}}$

$z \sim N(0,1)$.

χ^2_k = Chi square dist with degree of freedom k.

k = degree of freedom.

(ii) $\mu(t_k) = 0$ $\quad var(t_k) = \dfrac{k}{k-2}$

(iii) $n = 10$

$\mu = 50$

$s^2 = 48.667$.

$t_{0.05,9} = 3.25$.

(a) $\mu = 50 \pm 3.25 \sqrt{\dfrac{48.667}{10}} = (42.8303, 57.1699)$

(b) $CI = 50 \pm 2.5758 \sqrt{\dfrac{48.667}{10}} = (44.3176, 55.6823)$

classmate

Date

Page

Q10 $n = 1000$ $\quad X \sim Poi(3)$ $\quad X \sim N(3000, 3000)$

(i) Type 1 Error = P(Ho being rejected when it is T)

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

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

$= P[Z > 1.8257)$.

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

$= \boxed{3.39\%}$

(ii) Type 2 Error = P(Ho not being rejected when it is false)

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

$$= P\left[X < \frac{3100 - n\lambda}{\sqrt{n\lambda}}\right]$$

(iii) Power = P[reject Ho | Ho is false]

$= P[X > 3100 \mid \lambda = n\lambda]$

$$= 1 - P\left[\frac{3100 - n\lambda}{\sqrt{n\lambda}}\right]$$

(iv) $n = 2900$.

$$CI = 3 \pm 2.5758 \sqrt{\frac{3}{2900}}$$

$$\boxed{CI = (2.9172, 3.0828)}$$

Q11 $\Sigma x = 213.200$

$\Sigma x^2 = 4919860000$

sample mean $= 17767$

sample s.d $= 10144$

$n = 12.$

$\Sigma x' = (213.200) - 11000 - 48000 = 154200.$

$\Sigma x'' = 154000 + 27500 + 31000 = \underline{213200}$

$\boxed{\bar{x}' = 17767}$

$\Sigma x^{2'} = 4919860000 - 11000^2 - 48000^2$

$= 2494860000$

$\Sigma x^{2''} = 2494860000 + 27800^2 + 31500^2$

$= 4243360000$

$$s = \frac{1}{\sqrt{11}} \sqrt{4243360000 - 12(17761)^2}$$

$\boxed{s = 6434.0326}$

Q12.

(i) The CRLB result holds under very general ~~cond~~ conditions except where range of the distribution involves the parameter, such as unit distribution in this case.
This is due to a discontinuity so the derivative in the formula doesn't make any sense.

(ii) ~~$Y = Max_i \, x_i$~~

$$P[Y < y] = P[Max_i < y]$$
$$= P\left\{\bigcap_{i=1}^{n} (x_i < y)\right\}$$
$$= \prod_{i=1}^{n} P(x_i < y)$$
$$= \prod_{i=1}^{n}\left[\int_0^y \frac{dx}{\theta}\right] = \left(\frac{y}{\theta}\right)^n$$

Thus, PDF of Y will be $g_Y(y) = \frac{d}{dy} P[Y = y]$

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

Now for $k \geq 0$

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

classmate

Date

Page

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

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

$$= E[cY - \theta]$$

$$= cE(Y) - \theta$$

$$= c \cdot \frac{n\theta}{n+1} - \theta.$$

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

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

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