NAME : SALONI TAYSHETE
ROLL NO. : 76

papergrid

Date: /(1)/

ASSIGNMENT -1

Probability & Statistics 2

Q1. $X \sim Poi(\theta)$

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

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

$X \sim Poi(200)$

$$95\%CI = \mu \pm 1.96\sqrt{\theta}$$
$$= 200 \pm 1.96\sqrt{200}$$
$$= 200 + 1.96\sqrt{200} \quad OR \quad 200 - 1.96\sqrt{200}$$
$$= 227.7186 \quad OR \quad 172.2814$$

$\therefore$ The interval is $(172.2814, 227.7186)$

Q2. $\bar{X} = \sum_{i=1}^{x} \frac{X_i}{n}$ = sample mean

$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})$ = sample variance

(i) $E(\bar{X}) = E\left(\frac{\Sigma X_i}{n}\right) = \frac{1}{n} E(\Sigma X_i) = \frac{1}{\cancel{n}} * \cancel{n}\mu = \boxed{\mu}$

$V(\bar{X}) = V\left(\frac{\Sigma X_i}{n}\right) = \frac{1}{n^2} V[E(\bar{X})] = \frac{\cancel{n}\sigma^2}{n^{\cancel{2}}} = \boxed{\frac{\sigma^2}{n}}$

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

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

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

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

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

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

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

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

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

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

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

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

$$\frac{9 \times s^2}{100} = 0.5$$

$$s^2 = \frac{5}{1\cancel{0}} \times \frac{10\cancel{0}}{9}$$

$$s^2 = \frac{50}{9}$$

$$\boxed{\therefore s = 2.357\%}$$

Q6.(ii) $P(X \geq 3 / P = 0.3)$
$+ P(X = 0 / P = 0.3) . P(Y \geq 5 / P = 0.3)$
$+ P(X = 1 / P = 0.3) . P(Y \geq 4 / P = 0.3)$
$+ P(X = 2 / P = 0.3) . P(Y \geq 3 / P = 0.3)$

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

(iii) P(type II error) is the probability of accepting Ho when Ho is false.

$\therefore$ P(type II error) $= 1 - 0.91253$
$= 0.08747 \approx \boxed{9\% \ (p = 0.3)}$

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

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

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

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

$\therefore$ At CI 90% $p \in (0.3333, 0.2782)$

Q 8 (i) Unbiased estimator:

$Eg(X) = \theta$

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

(ii) If $E(\theta) \neq \theta$, 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) Method of moments, Maximum likelihood estimator

Q3. Let no. of claims be a random variable X.
$X \sim Bin(4, p)$

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

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

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

$\times [{}^4C_3 \cdot p^3 (1-p)^1]^2 \times [{}^4C_4 \cdot p^4]^1$

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

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

$\frac{dL}{dp} = 0 + \frac{117}{p} - \frac{603}{1-p}$

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

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

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

$117 - 117p = 603p$

$\therefore 720p = 117$

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

(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

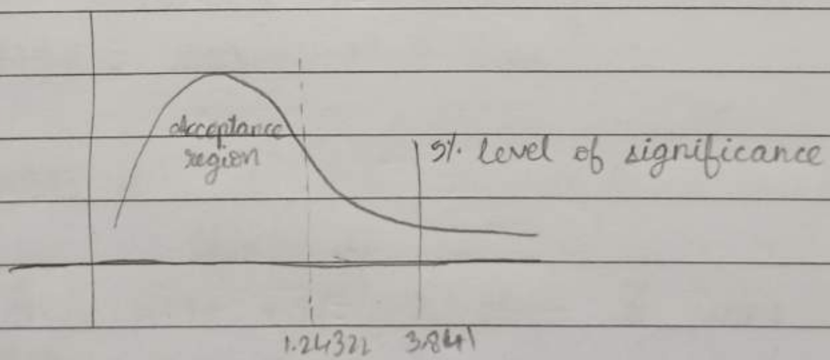
x	0	1	2	3	4
O_i	86	75	16	2	1
E_i	88.542	68.724	19.998	2.574	0.108

$p = 0.1625$

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

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

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



We have insufficient evidence to reject H_0 at 5% level of significance, therefore, no. of claims conformed to the binomial model.

Q4. $f(x) = \frac{C\beta^3}{(x+\beta)^4}$; $x > 0$, $\beta > 0$

(i) Comparing the above $f(x)$ to PDF of Pareto distribution (2 parameter version)
For pareto,

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

Thus β corresponds to λ

$\therefore$ $c = 3$ i.e. $\alpha = 3$

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

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

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

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

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

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

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

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

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

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

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

$= 0$ $\therefore$ unbiased.

$\Rightarrow MSE = V(\hat{\beta}) + (Bias)^2 = V(\hat{\beta}) = V(2\bar{X})$

$= 4V(\bar{X})$

$= 4V\left(\frac{\sum X_i}{n}\right) = \frac{4\sum V(X_i)}{n^2} = \frac{\cancel{4}}{n^{\cancel{2}}} \times \frac{3\beta^2 \cancel{n}}{\cancel{4}}$

$\therefore MSE = \frac{3\beta^2}{n}$

As n increases, MSE tends to 0
Hence, $\hat{\beta}$ is constant.

Q5.(i) $X \sim U(a,b)$

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

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

$$b = 2\bar{X} - a \quad \dots ①$$

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

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

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

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

$$2\sqrt{3}s = 2\bar{X} - a - a \quad \text{[from ①]}$$

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

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

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

(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 observations : 1, 2, 8, 50, 150

$$\bar{X} = 42.2$$

$$s = 63.57$$

$$\hat{a} = \bar{X} - \sqrt{3}s = 42.2 - \sqrt{3}(63.57) = -67.906$$

$\hat{b} = \bar{x} + \sqrt{3}\, s = 42.2 + \sqrt{3}\,(63.57) = 152.3064$

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

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

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

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

$\log L = -n \log b$

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

$\frac{dL}{db} = 0$, gives $b = \infty$ which is not possible

Also,

$\frac{d^2 l}{db^2} = \frac{n}{b^2}$

$\frac{n}{b^2} > 0$, which we have a minimum likelihood estimate

Thus, we say, $\hat{b} = \max(X_i)$