Probability & Statistics II

Assignment 1

Q1 $X \sim \text{Poisson}(\theta)$

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

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

$$\hat{\lambda} = \frac{X_i}{n} = 200 \rightarrow \text{given}$$

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

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

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

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

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

$$= (172.3, 227.7)$$

Q2 $X_i : i = 1, 2 \ldots n$

$Mean = \mu$

$Varian = \sigma^2$

$\bar{X} = \sum_1^n X_i / n$

$S^2 = \frac{1}{n-1} \sum_i^n (X_i - \bar{X})^2$

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

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

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

$= \mu$

$\therefore$ Hence shown

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

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

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

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

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

$\therefore$ Hence shown

ii) Show $E(S^2) = \sigma^2$

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

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

Lets conside $V(\bar{x})$

$$V(X) = E(X_i^2) - (EX)^2$$

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

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

$$E(\bar{x}^2) = \mu^2 + \frac{\sigma^2}{n} \to eq^n (3)$$

Compute eq^n (2) & 3 into eq^n (1)

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

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

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

$$E(S)^2 = \sigma^2$$

∴ Hence shown

iii) $X_i \sim (n-1)S^2/\sigma^2$

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

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

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

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

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

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

$\therefore$ hence shown

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

$P(0.5\sigma^2 < S^2 < 1.5\sigma^2) = P(S^2 < 1.5\sigma^2) - P(S^2 < 0.5\sigma^2)$

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

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

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

$$= 0.8587$$

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

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

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

$$= 0.1243$$

$P(50 < S^2 < 150) = 0.8587 - 0.1243$

$$= 0.7342$$

Q3 180 policies
Bin $n = 4$ & p

0	1	2	3	4
86	75	16	2	1

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

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

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

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

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

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

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

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

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

$117 - 117p = 603p$

$720p = 117$

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

$\hat{p} = 0.1625$

ii) Goodness of fit ~~te~~ test

H_0 : no. of claims follow binomial distributio
H_1 : ~~H_0 is~~ no. of claims do not follow binomial model

n	O_i	E_i
0	86	88.5
1	75	68.7
2	16	19.9
3	2	2.8
4	1	0.108

$p = 0.162$

$P(X=0) = 0.49$

$P(X=1) = 0.38$

$P(X=2) = 0.11$

$P(X=3) = 0.014$

$P(X=4) = 0.0066$

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

$T_{cal} = 1.24322 \sim \chi^2_1$

$T_{(tab)} = 3.841$

$\therefore$ We Do not reject H_0

Q4 $f(x) = \frac{c\beta^2}{(x+\beta)^4} : x > 0, \beta > 0$

i) P.D.f for pareto distributio $= \frac{\alpha \lambda^{\alpha}}{(\lambda + x)^{\alpha+1}}$

$f(x) =$ pdf of pareto distributi.

$\therefore C = 3$ (as C correspond to α here)

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

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

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

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

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

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

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

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

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

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

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

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

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

$= \beta - \beta$

$= 0$

$\therefore$ Unbiased

$M.SE = V(\hat{\beta}) + (\text{Bias}^2)$

$= V(\hat{\beta}) + 0$

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

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

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

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

$M.SE = \frac{12\beta^2}{n}$

$\therefore$ as n increase, M.S.E leans toward zero, $\hat{\beta}$ is constant

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

$$E(\hat{\beta}) = E(b\bar{x})$$
$$= bE\left(\frac{\sum x_i}{n}\right)$$
$$= \frac{b}{n}\sum E(x_1)$$
$$= \frac{b}{\cancel{n}} \times \frac{\beta}{2} \times \cancel{n}$$
$$= \frac{b\beta}{2}$$

$$V(\hat{\beta}) = V(b\bar{x})$$
$$= b^2 V\left(\frac{\sum x_i}{n}\right)$$
$$= \frac{b^2}{n^2}\sum V(x_i)$$
$$= \frac{b^2}{n^{\cancel{2}}} \times \frac{3}{4}\beta^2 \times \cancel{n}$$
$$= \frac{3}{4}\frac{b^2\beta^2}{n}$$

$$MSE = V(\hat{\beta}) + (Bias)^2$$
$$= \frac{3b^2\beta^2}{4n} + \left(\frac{b\beta}{2} - \beta\right)^2$$
$$= \frac{3b^2\beta^2 + nb^2\beta^2 + 4n\beta - 4n\beta b)}{4n}$$

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

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

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

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

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

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

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

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

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

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

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

Thus, $b=\frac{2}{(1+3/n)}$ reduces the M.S.E

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

$$= \frac{\cancel{8}\, n}{(1+3n)} \times \frac{\beta}{2}$$

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

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

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

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

MSE = $\cancel{E}$

Q5 i) $U(a,b)$ $\hat{a} = \bar{x} - \sqrt{3}$

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\hat{b} = \bar{x} + \sqrt{3}s$$

iii) Sample · 1, 28, 50, 150

$\bar{x} = 42.2$

$s = 63.57$

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

$= -67.906$

$\hat{b} = \bar{x} + \sqrt{3}s$

$= 152.306$

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

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

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

$\log L = -n \log b$

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

$\frac{dl}{db} = 0 \quad \therefore b = \infty$ which is not possible

Also

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

$\frac{n}{b^2} >$ this has M.L.E

$\hat{b} = \text{Max}(X_i)$

Q6 $H_0 : p = 0.1$
$H_1 : p > 0.1$

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

$= (1 - 0.8891) + 0.2824(1 - 0.885) + (0.658 - 0.2824)(1 - 0.9744) + (0.889 - 0.6590)(1 - 0.8881)$
$= 0.1472 \approx 15\%$

ii) P of rejecting H_0 when $p = 0.3$

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

$= (1 - 0.2528) + 0.0138(1 - 0.7237) + (0.0860)(1 - 0.4925) + (0.2528 - 0.085)(1 - 0.2525)$
$= 0.9125 \approx 91$

iii) Probability of type II error

$= 1 - 0.91253$

$= 0.08747$

$\approx 9\%$

iv) $n = 120,\ x = 40$

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

$$\hat{p} = \frac{40}{120} = \frac{1}{3} = 0.33$$

$z_{10\%} = 1.2816$

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

$$= 0.33\cancel{3} - 1.2816\sqrt{\frac{0.33(1-0.33)}{120}}$$

$= 0.28$

(

Page No.
Date: / /

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

$\frac{39.5 - 120(0.27)}{\sqrt{120 \times 0.278 \times 0.722}}$

$= 1.251$

$\therefore$ Do not reject H_0 at $10\% \cdot 2 - 0 \cdot 5$

vi) The confidence interval implies that there is only a 10% change in test ~~cat~~ carried out

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

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

~~$S_p^2 = \frac{1}{28}(11(2046) + 14(4800)) = 1402.64$~~

$T \cdot S = \frac{65 - 70}{6.5 - 43 \sqrt{\frac{1}{1} + \frac{1}{15}}} = -0.02034$

Page No.
Date: / /

$$TS = \frac{65-70}{63.458\sqrt{\frac{1}{12}+\frac{1}{15}}}$$

$= -0.2034$

There $\pm 2.060 (= t_{25, 0.025})$

$\therefore$ Do not Reject H_0 at 5% level

Q7 i) Public

$\bar{X}_1 = 30$

$S_1^2 = 64.3125$

Private

$\bar{X}_2 = 35.055$

$S_2 = 67.4027$

ii) Populations with equal varial

$H_0: \sigma_1 = \sigma_2$

$H_1: \sigma_1 \neq \sigma_2$

$$T.S = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

$$= \frac{64.31}{67.43} = 0.9542$$

At 5% LoS condition is satisfied

$\therefore$ Do not reject H_0

iii) $H_0 : \mu_B = \mu_{B}$
$H_1 : \mu_B > \mu_{Br}$

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

$$= \frac{35.0955 - 30}{8.1152\sqrt{\frac{2}{9}}}$$

$T.S = |-3215|$

Tab val $= t_{16}\ 5\%$
$= 1.746$

$\because T_{Scal} < T.S_{tab}$

$\therefore$ Reject H_0

Q8 i) An unbiased estimator of a parameter is an estimator whose expecte value is equal to the parameter

ii) Bias is anything that leads to a systematic difference between the true parameters of a population.

iii) $MSE(\hat{\theta}) = Var(\hat{\theta}) + Bias$
∴ measures the average square of errors

iv) Estimator having less MSE is more consistent

(v) Both will be considered as consistent when MSE minimizes

(vi) 1) Maximum likelihood Estimate
2) method of moment

Q9 i) $t_K \rightarrow \frac{N(0,1)}{\sqrt{\chi^2_K / K}}$

K : degree of freedom
2 random variables $N(0,1)$ & χ^2_K are independent

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

Page No.
Date: / /

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

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

$P\left(\bar{X} - t_{9,0.5}\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{9,0.5\%}\frac{s}{\sqrt{n}}\right)$
$= 0.\cancel{88}$

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

99% CI for μ $\left(30 - 3.25\frac{\sqrt{48.69}}{\sqrt{10}}, 30+3.25\frac{\sqrt{48.66}}{\sqrt{10}}\right)$

$= (42.83, 37.17)$

b) Error (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
$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim N\left(0, \sqrt{\frac{9}{7}}\right)$

$$\frac{\bar{X}-\lambda}{s/\sqrt{n_8}} \sim N(0,1)$$

Critic val level is 2-5

$$C \cdot I = \left(50 \cdot 2 - 3 \cdot 38\sqrt{\frac{48.667 \times 9}{10 \quad 7}},\ 50+2 \cdot 88\sqrt{\frac{48.667 \times 9}{10}}\right)$$

$$C \cdot I = (43 \cdot 34, 56 \cdot 45)$$

Q10 $n = 1000$

$\lambda = 3$

$X \sim Poi(n\lambda)$

$X \sim Poi(3000)$

i) $P(\text{Type I error}) = P(X > 3100 \mid X \sim Poi(3000))$

Using normal approx.

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

$$= 1 - P(Z < 1 \cdot 825)$$

$$= 3 \cdot 39\%$$

ii) Type II error = event of accepting hypothesis when it's false

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

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

Value of power depends on the value of parameter

iv) $P\left(-2.5758 < \frac{2.8 - \mu}{\sqrt{2.8/100}} < 2.5758\right) = 0.99$

$CI = (2.7613, 3.0387)$

Q11 Sample mean

$\Sigma X' = \Sigma X - 11000 - 48,000 + 27,500 + 31,500$
$= 213,200$

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

$= 17,767$

Sample Standard Deviati

$\Sigma x^{2'} = 4,918,860,000 - (11,600)^2 - (48,000)^2 + (27,300)^2 + (31,500)^2$

$= 4,243,360,000$

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

$$= \frac{1}{11}\left(4,243,360,000 - 12(17,767)^2\right)$$

$= 4,13,96,773.64$

$s' = 6434.033$

$\therefore$ Sample means remains the same with revised data and Standard deviation reduces.

Q12 i) The CRB ~~results~~ holds true under the condition where the variance of any such estimator is at least as high as the inverse of Fisher information.

In this case, the uniform distribution due to the ~~discont~~ discontinuity, to the derivative in the formula does not make sense.

ii) $P[X < y] = P(\max X_i < y)$
$= P\left[\bigcap_{i=1}^{n} (X_i < y)\right]$

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

$= \prod_{i=1}^{n} \left(\int_0^y \frac{dx}{\theta} \right)$

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

Pdf of y will be $= \frac{n}{\theta^n} \cdot y^{n-1}$

for $0 < y < \theta$

for any non-negative value

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

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

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

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

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

iv) $$MSE\{\hat{\theta}(c)\} = E\{(\hat{\theta}(c) - \theta)^2\}$$
$$= E\{(cY - \theta)^2\}$$
$$= E\{c^2Y^2 - 2c\theta Y + \theta^2\}$$

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

Page No.

Date: / /

iv) $(\hat{\theta}_C)$ to be unbiased estimator of θ

$$\text{Bias}(\hat{\theta}(C_4)) = 0$$

$$\left(\frac{C_4 n}{n+1} - 1\right)\theta = 0$$

$$\Rightarrow C_4 = \frac{n+1}{n}$$