Vidya Vikas Education Trust
Universal College of Engineering

Page No. ____ Date: ____

Probability & Statistics — Vishwar Shah.

Q1. $\dfrac{\hat{\lambda} - \lambda}{\sqrt{\hat{\lambda}/n}} \sim N(0,1)$

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

$\hat{\lambda} = 200$

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

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

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

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

$= 0.95$

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

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

$= (172.28141,\ 227.7185)$

Q2. $X_i;\ i = 1, 2, \dots, n$ are iid.

Mean $= \mu$

Variance $= \sigma^2$

$\bar{X} = \dfrac{\sum X_i}{n}$

$S^2 = \dfrac{1}{n-1}\left(\sum x_i - n\bar{x}^2\right)$

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

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

$V(\bar{x}) = V\left(\frac{\Sigma x_i}{n}\right)$

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

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

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

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

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

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

Now,

$V(x) = E(x_i^2) - [E(x)]^2$

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

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

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

$= \sigma^2.$

Q3. Let no. of claims be a random variable X.

$X \sim Bin(4, p)$

(i.) $L =$

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

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

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

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

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

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

$\hat{p} = 0.1625$

(ii.) Goodness of fit test.

H_0 : no. of claims confirms to the binomial model

H_1 : 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

Vidya Vikas Education Trust
Universal College of Engineering
Page No.
Date:

$p = 0.1625$

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

$P(X=1) = 0.3818$

$P(X=2) = 0.1111$

$P(X=3) = 0.0143$

$P(X=4) = 0.0006$

O_i	86	75	19
E_i	88.542	68.724	22.68

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

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

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

(i) Comparing the above $f(x)$ to PDF of dist (2 parameter)

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

Thus, β corresponds to λ.

=> Thus, $C = 3$ (as C corresponds to α here)

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

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

Vidya Vikas Education Trust
Universal College of Engineering

Page No.
Date:

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

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

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

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

$\beta = 2\bar{x}$

$$\begin{aligned} \text{Bias} &= E(\beta) - \beta \\ &= E(2\bar{x}) - \beta \\ &= \frac{2}{n} \sum E(x_i) - \beta \\ &= \frac{2}{n} \times \frac{n\beta}{2} - \beta \end{aligned}$$

Bias $= 0$, hence unbiased.

$$\begin{aligned} MSE &= V(\hat{\beta}) + (\text{Bias})^2 \\ &= V(\hat{\beta}) \\ &= V(2\bar{x}) \\ &= 4V\left(\frac{\sum x_i}{n}\right) \\ &= \frac{4}{n^2} \sum V(x_i) \\ &= \frac{3\beta^2}{n} \end{aligned}$$

Vidya Vikas Education Trust
Universal College of Engineering
Page No.
Date:

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

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

$= b$

$bE\left(\frac{\sum x_i}{n}\right)$

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

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

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

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

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

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

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

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

$MSE = V(\hat{\beta}_2) + (Bias)^2$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

thus, ~~ere~~ we have a negative bias & hence it undederistimate the true parameter β.

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

$= \beta^2\left(\frac{4n^2 + 12n - 4n^2}{n+3}\right)$

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

As $n \to \infty$, MSE doesnot tend to 0

Hence, it is inconsitent.

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

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

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

$a = 2\overline{X} - b \quad \text{--- } ①$

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

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

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

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

$2\sqrt{3}S = 2\overline{X} - a - a \quad \{b = 2\overline{X} - a\}$

$2\sqrt{3}S = 2\overline{X} - 2a$

$\sqrt{3}S = \overline{X} - a$

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

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

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

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

$\overline{X} = 42.2$

$S = 63.57$

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

$= -67.906$

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

$= 152.3064$

Hence we can conclude that value of $\hat{b}$ is approve 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.

likelihood estimate.

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

$$L = \prod_{i=j} f(x_j)$$

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

$$\log L = -n \log b$$

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

$\frac{dl}{db} = 0$, 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, $\hat{b} = \max(x_i)$

Q7. (i) Public

$\bar{X}_1 = 30$

$S_1^2 = 64.3125$

Private

$\bar{X}_2 = 35.0555$

$S_2^2 = 67.4027$

(ii) 2 sided t-test can be applied in case the samples come from population with equal variancy.

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

Test statistics is $\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{n_1 - 1, n_2 - 1}$

value of statistics $\frac{64.31}{67.42} = 0.9542$

(iii) ~~$H_0 : \mu_{ga} = \mu_{gu}$~~

$H_1 : \mu_{ga} > \mu_{gu}$

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

TS. = 1.32151

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

= 1.746

Since TScal < tabulated value,

We have ~~s~~ insufficient evidence to reject H_0.

Q8 (i) Unbiased estimate.

→ An unbiased estimate is an estimator where expect value is equal to the parameter.

→ Estimate is said to be unbiased if bias is equal to zero for all values of parameter θ.

(ii) biased

→ A biased estimate is one which delivers an estimate which is consistently different from the parameter to be estimated.

(iii) The mean square error of an estimate $\hat{\theta}$ is defined by, $MSE(\hat{\theta}) = var(\theta) + Bias^2$

(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) (i) Method of moments

(ii) Maximum likelihood estimate.

10. $n = 100$ $\qquad X \sim Poi(n\lambda)$
$\lambda = 3$ $\qquad X \sim Poi(3000)$

(i) P(Type I error) = P(reject H_0 / H_0 is true)
$= P(X > 3100 \mid X \sim Poi(3000))$

Using normal error
$$P(X > 3100) = P\left(Z > \frac{3100 - 3000}{\sqrt{3000}}\right)$$
$$= 1 - P(Z < 1.825)$$
$$= 3.39\%$$

(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)
$= P[(X > 3100 \mid X \sim Poi(n\lambda) - N(n\lambda, n\lambda)]$
$$= 1 - P\left(Z < \frac{3100 - n\lambda}{\sqrt{n\lambda}}\right)$$

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

(iv) $\hat{\mu}$ is estimator of μ, $\hat{\mu}$ follows $N(\mu, \frac{\hat{\mu}}{n})$
Hence, $\frac{\hat{\mu} - \mu}{\sqrt{\hat{\mu}/n}} \sim N(0,1)$

Q11. $n = 12$
$\Sigma x = 213,200$
$\Sigma x^2 = 4,919,860,000$
$\bar{x} = 17,767$
$S = 10,144$

SKA
ITS MY

Vidya Vikas Education Trust
Universal College of Engineering

Page No.
Date:

New Sample mean

$\Sigma X' = \Sigma X - 11000 - 48000 + 27500 + 31500$

$= 2,13,200$

$\bar{X}' = \frac{213200}{12} = 17,767$

New sample standard deviation

$\Sigma x^{2'} = 4919860000 - (11,000)^2 - (48000)^2$

$+ (27500)^2 + (31500)^2$

$= 4,243,360,000.$

1.2 $X_1, X_2, X_3 \dots X_n \sim u(0, \theta)$, $a = 0, b = 0$

(i) $E(Y) = \frac{\theta}{2}$ $V(Y) = \frac{\theta^2}{12}$

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

$\ln(L) = a/\theta$

$\frac{dn}{a} = \frac{-n}{\theta}$ $\theta =$ infinitie.