Krishangi Trivedi 82.

Q-1. $x \sim Poisson(\theta)$

$x \sim N\left(\theta, \frac{\theta}{n}\right)$

95% CI for poisson approximation

$$= \hat{\theta} \pm Z_{2.5\%} \frac{\sqrt{\theta}}{\sqrt{n}}$$

If a single value of 200 is calculated then $n=1$, $\hat{\theta} = \frac{\Sigma x}{n} = \frac{200}{1} = 200$

$\therefore$ 95% CI $= 200 \pm 1.96 \times \sqrt{200}$

$(172.281414 ; 227.718585)$

Q-2. $X \sim N(\mu, \sigma^2)$

$X_1, X_2, X_3 - - - X_n$ are iid variables.

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{E(\mu + \mu + \mu \cdots \mu_n)}{n}$$

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

Q-2 (i) $X_1, X_2, X_3 \ldots X_n$ are iid variables

$X \sim N(\mu, \sigma^2)$

$$E(\bar{X}) = E\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right)$$

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

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

$$V(\bar{X}) = V\left(\frac{\sum x_i}{n}\right) = \frac{V\left(\sum_{i=1}^{n} x_i\right)}{n^2}$$

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

$$V(\bar{X}) = \frac{\sum(\sigma^2)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\therefore \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

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

$$S^2 = \frac{1}{n-1}\left(\sum_{i}^{n} x_i^2 - n\bar{x}^2\right).$$

$E(S^2)$

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

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

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

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

$$\therefore E(x^2) = V(x) + [E(x)]^2$$

$$\therefore E(x^2) = \sigma^2 + \mu^2 \quad \text{--- (1)}$$

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

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

$$E(\bar{x}^2) = \frac{\sigma^2}{n} + \mu^2 \quad \text{--- (2)}$$

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

... from (1) & (2)

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

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

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

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

Q.2.(iii) $V\left(\frac{(n-1)S^2}{\sigma^2}\right) = V\left(\chi^2_{(n-1)}\right)$

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

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

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

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

$P(-0.5 < S^2 < 0.5)$

$\therefore P\left(\frac{-0.5\times(10-1)}{100} < \frac{(10-1)S^2}{100} < \frac{0.5\times(10-1)}{100}\right)$

$\therefore P(-0.045 < \chi^2_9 < 0.045)$

$P(\chi^2_9 < 0.045)$

$1 - P(\chi^2_9 < -0.045)$

$P\left((-0.5\times100)+100 < S^2 < (50\%\times100)+100\right)$

$\therefore P(50 < S^2 < 150)$

$\therefore P\left(\frac{50\times(10-1)}{100} < \frac{(n-1)S^2}{100} < \frac{150\times9}{100}\right)$

$\therefore P(4.5 < \cancel{\chi^2_9}\, y < 13.5) = P(y<13.5) - P(y<4.5)$

$= 0.7342$

Date ________
Page ________

Q.3. $n = 4,\ p$

* Let $x$ be the number of classes

$x_i \sim Bin(4, p)$

$$L = \prod_{i=0}^{86} P(x=0) \times \prod_{86}^{161} P(x=1) \times \prod_{161}^{177} P(x=2) \times \prod_{177}^{179} P(x=3) \times P(x=4)$$

$$= \left[1 \times p^0 (1-p)^4\right]^{86} \times \left[4 \times p^1 \times (1-p)^3\right]^{75} \times \left[6 \times p^2 (1-p)^2\right]^{16} \times \left[4 \times p^3 (1-p)\right]^2 \times \left[1 \times p^4 (1-p)^0\right]$$

$$L = (1-p)^{344} \times 4^{75} \times p^{75} \times (1-p)^{225} \times 6^{16} \times p^{32} (1-p)^{32} \times 16 \times p^6 (1-p)^2 \times p^4$$

$$\ln(L) = 6^{16} + 32\ln(p) + 32\ln(1-p) + 16 + 6\ln(p) + 2\ln(1-p) + 4\ln(p) + 344\ln(1-p) + 4^{75} + 75\ln(p) + 225\ln(1-p)$$

$$\frac{d(\ln L)}{dp} = \frac{32}{p} + \frac{(-32)}{(1-p)} + \frac{6}{p} + \frac{(-2)}{1-p} + \frac{4}{p} + \frac{(-344)}{(1-p)} + \frac{75}{p} + \frac{(-225)}{(1-p)}$$

$$\frac{d(\ln L)}{dp} = \frac{117}{p} + \frac{(-603)}{(1-p)}$$

equating this to 0

$$\frac{\cancel{117}}{\cancel{p}} + \frac{117(1-p) - 603(p)}{p+p^2} = 0$$

$$117 - 117p - 603p = 0$$

$$\cancel{7\,4\,86\,p = 117}$$

$$p = \cancel{0.2407407407} \quad 0.1625$$

~~double~~ $$\frac{d^2(\ln L)}{dp^2} = -\frac{117}{p^2} - \frac{603}{(1-p)^2}$$

$$\therefore \frac{d^2 \ln(L)}{dp^2} < 0$$

$\therefore$ this is the maximum likelihood.

Q. 3.(ii)

| $X_i$ | $O_i$ | $E_i$ |
|---|---|---|
| 0 | 86 | $180(P(x=0)) = 88.54730$ |
| 1 | 75 | $180(P(x=1)) = 68.72904492$ |
| 2 | 16 | $180\,P(x=2) = 20.003$ |
| 3 | 2 | $180(P(x=3)) = 2.587482$ |
| 4 | 1 | $180\,P(X=4) = 0.125512207$. |
| | $\boxed{186}$ | |

$P(x=0) = 0.491940728$

$P(x=1) = 0.381828027$

$P(x=2) = 0.11113$

$P(x=3) = 0.014374902$.

$P(X=4) = 0.0069727$

$\cancel{\dfrac{(O_i - E_i)^2}{E_i}}$ We will combine the last three rows.

| $x_i$ | $O_i$ | $E_i$ |
|---|---|---|
| 0 | 86 | 88.5473 |
| 1 | 75 | 68.729045 |
| 2+ | 19 | ~~20~~ 22.717 |

$\dfrac{(O_i - E_i)^2}{E_i}$ $\qquad \Sigma\left(\dfrac{(O_i - E_i)^2}{E_i}\right) = 1.25824203$

0.073701881 degrees of freedom = $3-2-1 = 1$

0.572172618 $\qquad \chi^2_{1,5\%} = 3.841$

0.6084794

$\therefore \chi^2_{cal} < \chi^2_{tab}$ $\therefore$ we do not reject $H_0$.

Q.4. $f(x) \quad \dfrac{c\beta^{3}}{(x+\beta)^4}$ ; $x>0$, $\beta>0$

comparing this with PDF of pareto distribution given on pg 14 of Formulae & tables book.

$$f(x) = \frac{\alpha \lambda^2}{(\lambda + x)^{\alpha+1}} \; ; \quad x>0$$

we get ; $c = \alpha = 3$ ; $\beta = \lambda$

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

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

(ii) $E(X) = \bar{x} = \frac{\beta}{\alpha - 1} = \frac{\beta}{2}$

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

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

$$\text{Bias} = E(\hat{\beta}) - \beta$$
$$= E(2\bar{x}_n) - \beta$$
$$= 2(\bar{x}_n) - \beta$$
$$= 2\frac{\beta}{2} - \beta = 0$$

Bias is 0 $\therefore$ $\hat{\beta}$ is an unbiased estimator

$$\text{MSE} = V(\hat{\beta}]$$
$$= V[2\bar{x}_n]$$
$$= 2 V(\bar{x}_n)$$
$$= \cancel{\frac{2 \times 3\hat{\beta}^2}{4}}$$

$$= 4 \times V\left(\frac{\Sigma x}{n}\right)$$

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

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

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

$$= \frac{3}{n}\hat{\beta}^2$$

as $n \to \infty$, MSE $\to 0$ $\therefore$ it is an unbiased estimator

Page No.
Date

5. i) p.d.f of a continuous uniform distribution $U(a,b)$

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

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

(ii) Let s be the standard deviation

$$s = \frac{(b-a)}{\sqrt{12}} = \frac{(b-a)}{2\sqrt{3}}$$

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

$$\therefore b = 2\sqrt{3}\,s - a \quad --(i)$$

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

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

$$a = 2\bar{x} - 2\sqrt{3}s - a \quad \text{(from (i))}$$

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

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

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

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

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

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

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

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

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

(iii) Sample : 10, 11, 12, 13, 14.

$$\bar{x} = \frac{60}{5} = 12$$

$$s = 1.5811388300$$

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

$$= 9.261387212$$

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

$$= 14.73861277$$

Date ______
Page ______

(iv) $U(a, b)$, $a = 0$.

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

$$= \left[\frac{1}{b-a}\right]^n$$

$$L = \left[\frac{1}{b}\right]^n$$

$$\ln L = n \ln \frac{1}{b}(b^{-1})$$

$$\ln L = -n \ln(b)$$

$$\frac{d \ln(L)}{db} = \frac{-n}{b}$$

$$\frac{d^2 \ln(L)}{db^2} = \frac{n}{b^2}$$

$$\because \frac{d^2 \ln(L)}{db^2} > 0$$

maximum likelihood estimate of $b$ is not possible

Q.6.

Type 1 error is the probability of rejecting Ho when Ho is true.

Let $X_1$ be the number of hits in the 1st step & $X_2$ be in 2nd step.

~~Q.i) Type 1 error = $P(X_1 \geq 3) P(X_2 = 2) + P(X_1 \geq 4) P(X_2 = 1) + P(X_1 \geq 5) \cdot P(X_2 = 0)$~~

~~$= P(X_1 \geq 3) P(X_2 = 2) +$~~

~~$P(X_1 \geq 3)(P(X_2 = 2)) + P(X_1 \geq 4) P(X_2 = 1) + P(X_1 \geq 5) \cdot P(X_2 = 0)$~~

$X_1 \sim Bin(12, 0.1)$ , $X_2 \sim Bin(24, 0.1)$

$\therefore \Sigma x_1 \sim N(1.2, 1.08)$

$\Sigma x_2 \sim N(1.2, 1.08)$

Type 1 error $= P(\Sigma X_1 \geq 3) P(\Sigma X_2 \geq 2)$

~~$= P\left(\frac{\Sigma X - 1.2}{\sqrt{1.08}}\right)$~~

$= P(\Sigma x_1 > 2.5) P(\Sigma x_2 > 1.5)$

$= P\left(\frac{\Sigma X_1 - 1.2}{\sqrt{1.08}} > \frac{2.5 - 1.2}{\sqrt{1.08}}\right) P\left(\frac{\Sigma X_2 - 1.2}{\sqrt{1.08}} > \frac{1.5 - 1.2}{\sqrt{1.08}}\right)$

$= P(Z_1 > 1.250926) P(Z_2 > 0.28867513)$

$= 1 - P(Z_1 < 1.250926) \times 1 - P(Z_2 < 0.288675)$

$= (1 - 0.89435) \times (1 - 0.61401)$

$= 0.10565 \times 0.38591$

$= 0.40771392$

(ii) $P(X \geq 3)\, P(X > 2)$ / $p = 0.3$

$= \therefore \Sigma X_1 \sim N(3.6,\ 2.52)$

$\Sigma X_2 \sim N(3.2,\ 2.52)$

$P(X > 2.5)\, P(X > 1.5)$

$P\left(\frac{X_1 - 3.6}{\sqrt{2.52}} > \frac{2.5 - 3.6}{\sqrt{2.52}}\right) P\left(\frac{X_2 - 3.6}{\sqrt{2.25}} > \frac{1.5 - 3.6}{\sqrt{2.52}}\right)$

$P(Z_1 > -0.692984834)\, P(Z_2 > -1.322875656)$

$P(Z_1 < 0.6929)\, P(Z_2 < 1.323)$

$= 0.75490 \times 0.90658$

$= 0.68437242$

(iii) Type II error ~~$P(X_2 \geq 3)$~~

$P(X_1 = 3)\, P(X_2 = 1) + P(X_1 = 4)\, P(X_2 = 0)$

${}^{12}C_3 \times (0.3)^2 (0.7)^8 \times {}^{12}C_1 \times (0.3)^1 (0.7)^{11} +$

${}^{12}C_4 (0.3)^4 (0.7)^8 \times$

${}^{12}C_0 (0.3)^0 (0.7)^{12}$

$= 0.017062778 + 0.003197271$

$= 0.020262049$

Page

$\not{t}$ (iv) $\Sigma x_i = 40$

$n = 120$

$\Sigma x_i \sim N(np, npq)$

$\cancel{CI}\ \hat{p}$

$\therefore \hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$

$\cancel{p = \hat{p} \pm Z_{10\%} \sqrt{\frac{p(1-p)}{n}}}$

$\cancel{p = \frac{\Sigma x_i}{n} \pm Z_{10\%} \sqrt{\frac{\hat{p}\hat{q}}{n}}}$

$\cancel{= \frac{40}{120} \pm}$

$p = \hat{p} + Z_{10\%} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$= 0.333 + 1.2816 \sqrt{\frac{0.333(1-0.333)}{120}}$

$p = 0.38848 1283$

$(0, 0.388481283)$

Date

Page

(v) $H_0 : p = 0.278$

$H_1 : p > 0.278$

$\alpha = 10\%$

$\therefore CI = 90\%$

$\hat{p} = 0.3333$

test statistic $= \dfrac{(0.33333 - 0.278)}{\sqrt{\dfrac{0.333 \cancel{- 0.278} (1-0.333)}{120}}}$

test stat $= 1.285753011$

$Z_{10\%} = 1.2816$

$Z_{cal} > Z_{tab} \therefore$ reject $H_0$.

(vi) $p$ falls between $(0.279, 0.368481287)$

(vii)

Q.7.

Let $x_1, x_2, x_3 \ldots x_9$

Be the claim sizes of Public sector

and let $y_1, y_2, y_3 \ldots y_9$

Be the avg claim size of the private sector.

$$\bar{X} = \frac{\Sigma x}{n} = \cancel{2\ldots} \; 30.$$

$$\bar{y} = \frac{\Sigma y}{9} = 35.0556$$

$S_x^2 = 64.3125$

$S_y^2 = 67.4028$

$$E(\bar{x}) = \frac{\Sigma E(x_i)}{9} = 30$$

$$V(\bar{x}) = \frac{\Sigma V(\bar{x})}{9^2} = \frac{64.3125}{81}$$

$$= 0.79398148$$

$E(\bar{y}) = 35.0556$

$V(\bar{y}) = 0.832133$

8. (i) unbiased estimator is one whose mean value is equal to its true value.

Suppose we have a random sample
$X = (x_1, x_2, \ldots, x_n)$
from a distribution with parameter $\theta$ &
$g(\underline{X})$ is an estimator of $\theta$

~~then~~ if $E[g(\underline{X})] = \theta$ then $g(\underline{X})$ is an unbiased estimator.

(ii) If the estimator is biased then the bias is given as $E[g(\underline{X})] - \theta$
it is the measure of the difference between the expected value of the ~~bias~~ estimator & the parameter being estimated.

(iii) $MSE = E\left((g(X) - \theta)^2\right)$.

MSE is the second moment of an estimator $(g(\underline{X}))$ about $\theta$ & an estimator with lower MSE is said to be more sufficient.

Page

(iv) The estimator with lower MSE is more efficient:

$$MSE = Variance + bias^2$$

hence if MSE of $\tilde{\theta}$ is lower even if it has positive bias it means that the variance of $\tilde{\theta}$ is very low which makes it more efficient.

Whereas $\hat{\theta}$ is unbiased but ~~has~~ has larger MSE and therefore has very large variance.

(v) $\tilde{\theta}$ is more would be termed as 'consistent'.

(vi) Method of Moments: in this method we equate the population moments with the corresponding sample moments and solve for the parameter.

Maximum Likelihood Estimate (MLE):
Here the ~~pdf~~ product of the probability density functions or the ~~dis~~ densities is calculated for each sample value. Then the derivative of the log likelihood function is taken & equated to 0 to obtain the MLE parameter.

Page No.
Date

Q.9.

(i) $t_k = \frac{N(0,1)}{\sqrt{\frac{\chi^2_k}{k}}}$

where $\frac{x-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

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

$\therefore k = (n-1)$

$\therefore t_k = \frac{\frac{x-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{n-1\ s^2}{\sigma^2(n-1)}}}$

$= \frac{\frac{x-\mu}{\sigma/\sqrt{n}}}{s/\sigma}$

$= \frac{x-\mu}{s/\sqrt{n}}$

where $x$ is the sample mean
$s$ is the sample variance
$\mu$ is the population mean
$n$ is the sample size

ii) for $k > 2$

E ~~$t_k$~~

mean of $t_k$ is 0

and variance of $t_k$ is $\frac{k}{(k-2)}$

$t_k \sim N\left(0, \frac{k}{k-2}\right)$

(iii) $n = 10$

$\bar{x} = 50$ $\qquad \dfrac{\bar{x} - \mu}{s/\sqrt{n}} \sim t_9$

$s^2 = 48.667$

(a) 99% CI $\quad t_{9, 1\%}$

2.821)

$$\frac{50 - \mu}{\sqrt{48.667/10}} = 2.8121$$

$\mu = 56.2232748$

(b) $t_p \sim N\left(0, \dfrac{9}{9-2}\right)$

$t_9 \sim N\left(0, \dfrac{9}{7}\right)$

$CI = \bar{x} \pm \cancel{t_{9,1\%}}\, z_{0.5\%} \sqrt{\dfrac{9}{7}}$

$= \cancel{50 \pm 2.812 \times}$

$\cancel{= 50 \pm 2.812 \sqrt{9/7}}$

$[\cancel{46.2114847, 53.1885622}]$

$= 50 \pm 2.5758 \times \sqrt{\dfrac{9}{7}}$

$47.0795, 52.92$

classmate
Date
Page

(10) Let $X \sim Poi(3)$

X is the number of claims in a year. ~~per policy~~.

$n = 1000$

$H_0 : \lambda = 3$

$H_1 : \lambda \neq 3$

$P(X > 3100 / \lambda = 3)$

$\sum x_i \overset{\sim}{\sim} N(n\lambda, n\lambda)$

$\sum x_i \overset{\sim}{\sim} N(3000, 3000)$ ... $\because \lambda = 3$

~~$P\left(\frac{\sum X - 3100}{\sqrt{3000}}\right)$~~

$P\left(\frac{\sum x_i - 3000}{\sqrt{3000}} > \frac{3100 - 3000}{3000}\right)$

$P(Z > 1.826)$

$\Rightarrow 1 - P(Z < 1.826)$

$(1 - 0.96633)$

$= 0.0336200$

(ii) $p = P(X < 31000 / \lambda \neq 3)$

(iv) $\hat{\lambda} \sim N\left(\lambda, \frac{\lambda}{n}\right)$ $\qquad \bar{x} = \frac{2900}{1000} = 2.9$

$$99\%\ CI = \cancel{2.9} \pm \hat{\lambda} \pm Z_{0.5\%} \sqrt{\frac{\lambda}{n}}$$

$$= 2.9 \pm Z_{0.5\%} \cdot \sqrt{\frac{2.9}{1000}}$$

$$= (2.761237,\ 3.038711075)$$

11. $\Sigma x = 213200$

$\Sigma x^2 = 4,919,860,000$

$\bar{x} = \cancel{176767}$ $17767$

$s = 10144$

$s^2 = 102900736$

revised sample mean

$$= \frac{213200 - 11000 - 48000 + 27800 + 31400}{12}$$

$$= 17764.$$

revised

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

$$= 2494860000 + (27500)^2 + (31500)^2$$

$$= 2124326000$$

revised sample varia

$$s^2 = \frac{\Sigma x^2 - n\bar{x}^2}{n-1} = \frac{2124326000 - 12(17767)^2}{11}$$

$s^2 = 21396775.64$

$\therefore s = 6434.03261$