Page No. | Date

Probability and Statistics - 2

Assignment 1

Suhani Kulkarni
Roll No - 42

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

1. θ is large then

$$\frac{X-\theta}{\sqrt{\theta}} \sim N(0,1)$$ - approximate confidence interval for θ.

A single value has been observed - 200

$\therefore$ 95% approximate confidence interval is as follows: 'Symmetrical

$$\therefore 0.95 = P\left(-1.96 < \frac{X-\theta}{\sqrt{\cancel{100}\theta}} < 1.96\right)$$

- The value are taken from table book

$Z_{2.5\%} = -1.96 \quad Z_{97.5\%} = 1.96$

$$\therefore 0.95 = P\left(-1.96 < \frac{200-\theta}{\sqrt{\hat{\theta}}} < 1.96\right)$$

$$\therefore 0.95 = P\left[(200-0.5) - 1.96\cdot\sqrt{\hat{\theta}} < \theta < (200+0.5) + 1.96\cdot\sqrt{\hat{\theta}}\right]$$

$\therefore$ The approximate 95% CI is

$$\left(199.5 - 1.96\cdot\sqrt{\hat{\theta}} < \theta < 200.5 + 1.96\cdot\sqrt{\theta}\right)$$

Page No.

Date

2. $X_i : i = 1, 2 \dots n$ 'n' independent & identically distributed random variables with mean 'μ' and variance 'σ^2'.

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

1) $E[\bar{X}] = E\left[\frac{\sum X_i}{n}\right] = \frac{1}{n} E[\sum X_i] = \frac{1}{n} \sum E[X_i]$ - X_i's are iid r.v

$= \frac{1}{n} \sum \mu$ μ is the mean of the X_i's iid r.v

$= \frac{1}{n} \times n\mu$

$= \mu$

$\boxed{\therefore E[\bar{X}] = \mu}$

$V[\bar{X}] = V\left[\frac{\sum X_i}{n}\right] = \frac{1}{n^2} \cdot V[\sum X_i] = \frac{1}{n^2} \sum V[X_i]$ - X_i's are iid r.v

$= \frac{1}{n^2} \sum \sigma^2$ 'σ^2' is the variance of iid r.v

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

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

$\boxed{\therefore V[\bar{X}] = \frac{\sigma^2}{n}}$

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

$= \frac{1}{n-1} E[\sum X_i^2 - n\bar{X}^2]$

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

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

Now, we know that

$V[x_i] = E[x_i^2] - E[x_i]^2$

$\therefore E[x_i^2] = V[x_i] + E[x_i]^2$

Here $V[x_i] = \sigma^2$ $E[x_i]^2 = \mu^2$

Similarly, $\bar{x} \sim N(\mu, \sigma^2/n)$

$\therefore V[\bar{x}] = E[\bar{x}_i^2] - E[\bar{x}]^2$

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

Here $V[\bar{x}] = \sigma^2/n$, $E[\bar{x}]^2 = \mu^2$

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

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

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

$= \sigma^2$

$\boxed{\therefore E(S^2) = \sigma^2}$

x_i's are from a normal population.

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

iii) Variance of a χ^2_{n-1} distribution is $2 \times$ degree of freedom
i.e $2 \times (n-1)$

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

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

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

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

Page No.
Date

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

We know that $E(s^2) = \sigma^2 = 100$

$\therefore$ minus and plus 50% value of 100 is 50 and 150.

$$\therefore P(50 < s^2 < 100) = P\left(\frac{9 \times 50}{100} < \frac{(n-1)s^2}{\sigma^2} < \frac{9 \times 100}{100}\right)$$

$$= P(4.5 < \chi^2_9 < 9) \quad \because \left(\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}\right)$$

$$= P(\chi^2_9 < 9) - P(\chi^2_9 < 4.5)$$

$$= 0.5627 - 0.1245$$

values from table book.

$$= 0.4382$$

3. $X \sim$ No. of claims.

$X \sim$ Binomial $(4, p)$.

A random sample of 180 policies was observed

$\therefore$ No. of claims	0	1	2	3	4
Frequency.	86	75	16	2	1

$\therefore$ i) MLE

Maximum Likelihood Estimate for p.

The likelihood function is given as:-

$$L(p) = \prod \cdot f_x(x)$$

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

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

$$\times \left[{}^4C_3 \cdot p^3 \cdot (1-p)\right]^2 \times \left[{}^4C_4 \cdot p^4 \cdot (1-p)^0\right]$$

Let the combination values be taken as constants.

Page No.
Date

$\therefore L(p) = p[(1-p)^4]^{86} \cdot [p\cdot(1-p)^3]^{75} \cdot [p^2(1-p)^2]^{16} \cdot [p^3(1-p)]^2 \cdot (\cancel{\ }p^1)$ × constant

$= (1-p)^{344} \cdot [p^{75} \cdot (1-p)^{125}] \cdot [p^{32} \cdot (1-p)^{32}] \cdot [p^6 \cdot (1-p)^2] \cdot p^4$ × constant

$L(p) = (1-p)^{603} \cdot p^{117}$ × constant

Taking ln on both sides,

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

Differentiating on both sides w.r.t p

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

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

Equating it to zero

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

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

$\therefore 117(1-p) = 603p$

$\therefore 117 - 117p = 603p$

$\therefore 117 = 603p + 117p$

$\therefore 720 \cancel{486} p = \cancel{47}\ 117$

$\therefore \hat{p} = \frac{117}{\cancel{486}\ 720} = \cancel{0.2407}\ 0.1625$

$\boxed{\therefore \hat{p} = \cancel{0.2407}}$ 0.1625

Good fit - H_0: The given Sample is from a binomial H_1: Not from a Binomial

ii) The probabilities with the ~~p~~ calculated value of the parameters are as follows:-

$P(X=0) = {}^4C_0 \cdot p^0 \cdot (1-p)^4 = (1-\cancel{0.2400}\ 0.1625)^4 = 0.49197$

$P(X=1) = {}^4C_1 \cdot p^1 \cdot (1-p)^3 = 4 \times (0.1625) \times (1-0.1625)^3 = 0.38183$

$P(X=2) = {}^4C_2 \cdot p^2 \cdot (1-p)^2 = 6 \times (0.1625)^2 \times (1-0.1625)^2 = 0.11113$

$P(X=3) = {}^4C_3 \cdot p^3 \cdot (1-p) = 4 \cdot (0.1625)^3 (1-0.1625) = 0.01437$

$P(X=4) = {}^4C_4\, p^4 \cdot (1-p)^0 = (0.1625)^4 = 0.0006$

∴ The expected values can be calculated with the help of these probabilities as 180 × probability

∴ The table for observed and expected values is as follows:-

No. of claims	Observed values	Expected values	~~$(O_i-E_i)^2/E_i$~~
0	86	88.55	~~0.07343~~
1	75	68.73	~~0.57199~~
2	16	20.00	~~0.8~~
3	2	2.59	~~0.1344~~
4	1	0.13	~~5.8223~~
	180	180.	

∴ We can club few rows together

No. of claims	O_i	E_i	$(O_i-E_i)^2/E_i$
0	86	88.55	0.0734
1	75	68.73	0.57199
2 +	19	22.72.	0.60908
			1.2546.

We know that

$$\chi^2 = \Sigma(O_i - E_i)^2/E_i$$

$$\text{Degree of freedom} = n - \underbrace{1}_{\text{Balancing figure}} - 1 - \underbrace{1}_{\text{no. of Parameter estimated}} = 4-1-1-1 = 1$$

∴ The 5% critical value of χ^2 is 3.841.

∴ We do not reject H_0.

∴ Model is a good fit.

Page No.
Date

Q.4.]

$$f_X(x) = \frac{c\beta^3}{(x+\beta)^4} \qquad x > 0,\ \beta > 0$$

c - constant $\qquad$ β - parameter

i) From the actuarial formulae and table book.

The p.d.f. for pareto distribution is

$$f_X(x) = \frac{\alpha\lambda^\alpha}{(x+\lambda)^{\alpha+1}} \quad , x>0, \lambda>0, \alpha>0$$

$$\therefore E[X] = \frac{\lambda}{\alpha-1} \qquad V[X] = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} \quad , (\alpha > 2)$$

By comparing the above distribution and the pareto distribution

we can see that $c = 3$ (α known)

Here $\lambda = \beta$ — which is not known.

$$\therefore E[x] = \frac{\lambda}{\alpha-1} = \frac{\beta}{3-1} = \frac{\beta}{2} \qquad V[x] = \frac{\alpha\lambda^2}{(\alpha-1)^2(\alpha-2)} = \frac{3\times\beta^2}{(3-1)^2(3-2)}$$

$$= \frac{3\beta^2}{4\times1} = \frac{3\beta^2}{4}$$

ii) Random Sample $X_1, X_2, \ldots X_n$ is taken

$$E[x] = \frac{\lambda}{\alpha-1} = \frac{\beta}{3-1} = \frac{\beta}{2}$$

$\therefore$ Equating Sample mean $\bar{X}_n$ with $E[X]$

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

$$\boxed{\therefore \hat{\beta} = 2 \cdot \bar{X}_n.}$$

To check the unbiasedness

$$\text{Bias} = E[\hat{\beta}] - \beta$$

$$E[\hat{\beta}] = E[2\bar{X}_n] = 2\cdot E[\bar{X}_n] = 2\cdot E[X] = 2\cdot\frac{\beta}{2} = \beta$$

Page No.
Date

$\text{Bias} = E[\hat{\beta}] - \beta = \beta - \beta = 0$

$\therefore$ The Method of Moment estimator is unbiased.

Consistency

The estimator is said to be consistent when $n \to \infty$ and $MSE \to 0$

$\because MSE = Var + Bias^2$

$\therefore Var[\hat{\beta}] = Var[2\bar{X}_n] = 2^2 \cdot Var[\bar{x}_n]$

$= 4 \cdot \frac{Var[x]}{n}$ - Xi's are independent

$= \frac{\cancel{4} \cdot \left(\frac{3}{\cancel{4}}\beta^2\right)}{n}$

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

$\therefore$ as $n \to \infty$, MSE will tend to 0 i.e $MSE \to 0$

$\therefore$ The estimator is consistent.

Q.5] i)

Page No.
Date

Q.6.] $H_0: p = 0.1$ $\quad H_1: p > 0.1$

Initially 12 missiles are fired.
Ho is rejected if 3 or more independent hits — in 12
and 5 or more in 24.

i) Type I error is the ~~prob~~ event of rejecting Ho when it is true.

$p = 0.1$

X — no. of hits in first step 12 misiles and Y be the no. of hits in second step 12 misiles

$\therefore P(\text{Type 1 error}) = P(X \geq 3 / p = 0.1)$
$+ [P(X=0) \times P(Y \geq 5)] + [P(X=1) \times P(Y \geq 4)]$
$+ [P(X=2) \times P(Y \geq 3)]$ ~~+ P(X=3) × P(Y≥2~~
~~+ P(X≥3) × P(~~

from the table book ~~[1 − P(X<2)] + 0.2824)~~

$= [1 - P(X<3)] + [P(X=0) \cdot [1 - P(Y<5)]] + [P(X=1) \times 1 - (P<4)]$
$+ P(X=2) \times [1 - P(X<3)]$

$= (1 - 0.8891) + [(0.2824)(1 - 0.99957)]$
$+ [(0.6590 - 0.2824)(1 - 0.9744)]$
$+ [(0.8891 - 0.6590)(1 - 0.8891)]$

$= 0.1109 + 0.001214 + 0.009641 + 0.02552$
$= 0.14726$

ii) Probability of rejecting Ho when $(p = 0.3)$
$: P(X \geq 3 / p = 0.3) + [P(X=0) \cdot P(Y \geq 5) / p = 0.3]$
$+ [P(X=1) \cdot P(Y \geq 4)] / p = 0.3]$
$+ [P(X=2) \cdot P(Y \geq 3) / p = 0.3]$

from the table book

~~= (1 − 0.2528) + [0.0138 (1 − 0.7237)]~~
~~+ (0.2744 − 0.0138~~

$= (1 - 0.2528) + [(0.0138) \times (1 - 0.7237)]$
$+ [(0.0850 - 0.0138) \times (1 - 0.4925)]$
$+ [(0.2528 - 0.0850) \times (1 - 0.2528)]$

Page No.
Date

$= 0.7472 + 0.003813 + 0.036134 + 0.12538$
$= 0.91253$

iii) Calculate P(Type II error) - $p = 0.3$

$\therefore$ P(Type II error) = P(Accepting H_0, when it is false)
= 1 - P(Rejecting H_0, when ~~than~~ $p = 0.3$)
$= 1 - 0.91253$
$= 0.087473$

iv) 10 Space Agencies developed similar missiles. 120 missiles were fired.
$\therefore n = \cancel{40}\ 120$
X - No. of hits $= 40$

By MOM
$\therefore$
$np = \cancel{E[x]}$ $\qquad np = E[x]$
$\cancel{np = \bar{x}}$ $\qquad n\,np = X$
$\cancel{\hat{p} = \bar{x}}$ $\qquad \therefore \hat{p} = \frac{x}{n} = \frac{40}{120} = 0.333$

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

$\therefore$ The Lower bound 90% interval can be calculated by

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

from the table books

$Z_{\cancel{90\%}/2} = Z_{10\%} = 1.2816$

$\therefore$ Lower bound $= \hat{p} - 1.2816 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$$= \left[0.333 - 1.2816\ \frac{0.3333(1-0.3333)}{120}\right]$$

$= 0.2782$

$\therefore$ The confidence interval 90% $(0.2782, \infty)$

Page No.
Date

v) $H_0 = 0.278$
$H_1 = p > 0.278$
Level of significance - 10%.

Test statistic with continuity correction.

$$Z_{cal} = \frac{\bar{X} - np}{\sqrt{np(1-p)}}$$

$$= \frac{(40 - 0.5) - (120 \times 0.278)}{\sqrt{120 \times 0.2782 (1 - 0.278)}}$$

$$= \frac{39.5 - 33.36}{\sqrt{24.08592}}$$

$$Z_{cal} = \cancel{1.245}\ 1.2511$$

This is a one sided test
from the table book,
~~for lower tail~~ The 10% point of the $N(0,1)$ distribution is 1.2816

i.e $Z_{tab} = 1.2816$
as $Z_{cal} < Z_{tab}$
We do not reject H_0.

vi) $n_1 = 12$ $n_2 = 15$
$\bar{X}_1 = 65$ $\bar{X}_2 = 70$ $S_1 = 54$ $S_2 = 70$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$
~~for~~ The pivotal quantity for this

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

Page No.
Date

Pooled variance

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

$$= \frac{(11 \times 54^2) + (14 \times 70^2)}{12+15-2}$$

$$= 4027.04.$$

$$\therefore \frac{(65-70)-0}{\sqrt{4027.04 \times \left(\frac{1}{12} + \frac{1}{15}\right)}} = \frac{-5}{63.459 \times 0.387298} = -0.20343$$

It is a Symmetrical CI

$\therefore t_{25, 2.5\%}$ $(-2.060, \quad 2.060)$

lower tail — upper tail

As the calculated value (test statistic) lies between these value

We donot reject H0

$\therefore \mu_1 = \mu_2$

Page No.
Date

Q.7] Let

i) $\bar{X}_{public} = \frac{\sum x_{i(public)}}{n_{public}} = \frac{24+45+29+33+20+40+26.5+25+27.5}{9}$

$= 270/9$

$= 30$

$S^2_{public} = \frac{\sum x^2_{public} - n\bar{x}^2_{public}}{n_{public} - 1} = 64.3125$

$\bar{X}_{private} = \frac{\sum x_{i(private)}}{n_{(private)}} = \frac{30+54.5+30+40+28.5+36+36.5+30.5+35.5}{9}$

$= \frac{315.5}{9}$

$= 35.0556$

$S^2_{private} = \frac{\sum x^2_{private} - n\bar{x}^2_{private}}{n_{private} - 1} = 67.4027$

ii) It is assumed the samples are from a normal population. Two-sided t test can be applied when the populations have equal variance

To test this

Let $H_0 : \sigma_1^2 = \sigma_2^2 \quad H_1 : \sigma_1^2 \neq \sigma_2^2$.

$\therefore$ Test statistic

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

$$\therefore \frac{S_1^2 \times \sigma_2^2}{\sigma_1^2 \; S_2^2} = \frac{S_1^2}{S_2^2} = \frac{64.3125}{67.4027} = 0.95415$$

From the table book,

At 5% LoS $\quad F_{8,8,97.5\%} = 4.433$

$F_{8,8,2.5\%} = 0.2256$.

Our calculated value falls in this range

$\therefore$ We do not reject Ho

ie we can assume that they have equal variances.

iii) We have to test if the treatments in private hospitals lead to higher claim size

$\therefore$ let $H_0: \mu_{private} = \mu_{public}$ $\quad H_1: \mu_{private} > \mu_{public}$.

$\therefore$ Test Statistic

$$\frac{(\bar{X}_{public} - \bar{X}_{private}) - (\mu_{public} - \mu_{private})}{S_p \cdot \sqrt{\frac{1}{n_{private}} + \frac{1}{n_{public}}}} \sim t_{n_1+n_2-2}$$

~~Pooled variance~~

$\therefore S_p^2$ (pooled variance) $= \cancel{30-35} \; \frac{(n_{pu}-1)S^2_{pu} + (n_{pr}-1)S^2_{pr}}{(n_{pu}+n_{pr}-2)}$

$$= \frac{(8 \times 64.3125) + (8 \times 67.4027)}{(9+9-2)}$$

$= 65.86 \quad \therefore \sqrt{S_p^2} = 8.1153$

$\therefore$ Value of test statistic

$$\frac{(30 - 35.0556) - (0)}{8.1153 \cdot \sqrt{\frac{1}{9} + \frac{1}{9}}} = -1.321$$

3825589

$\therefore t_{cal} = -1.321$.

Page No.
Date

$\therefore$ The level of significance is 5%

R/ We have to see if $\mu_{private} > \mu_{public}$.

$\therefore$ 5% upper tail value
of t_{16} is 1.746
i.e $t_{tab} = 1.746$

As $t_{cal} < t_{tab}$
We do not reject H_0.
$\therefore$ Claims from private hospitals is similar to claims in public hospital

Page No.
Date

Q.8] $\hat{\theta}$ is an estimator of parameter θ

i) Define unbiased estimator.

Unbiased estimator is such that the expected value of the ~~unbiased~~ estimator is equal to the true parameter value.

i.e if $E[\hat{\theta}] - \theta = 0$, the $\hat{\theta}$ is an unbiased estimator of θ.

ii) Define Bias

An estimator is said to be biased if the expected value of the estimator is not equal to the true parameter value

i.e $E[\hat{\theta}] - \theta \neq 0$.

$\therefore$ $\text{Bias} = E[\hat{\theta}] - \theta$

iii) Define MSE of this estimator $\hat{\theta}$

Mean Square Error (MSE) $= E[(g(x) - \theta)^2]$

Mean Square error is the second moment of $g(\underline{x})$ about θ.

MSE can also written by the following formula

$\text{MSE} = \text{Variance} + \text{Bias}^2$

iv) $\tilde{\theta}$ is another estimator of $\hat{\theta}$ - same parameter.

$\hat{\theta}$ has no bias but higher MSE than $\tilde{\theta}$.

$\tilde{\theta}$ has positive bias.

$\tilde{\theta}$ is a more efficient estimator of θ even if it has a positive bias. This is because lower the MSE, ~~the~~ more efficient is the estimator

So even though $\hat{\theta}$ is an unbiased estimator, it is less efficient due to high MSE.

Page No.
Date

v) ~~Th~~ Either one of the two estimators may be termed as consistent ~~wh~~ when Mean Square Error MSE $\to 0$ as $n \to \infty$

vi) The two method of estimating θ are as follows:-

a) Method of Moments:
The parameter (here θ) is estimated by equating the population moments with the sample moments and solve for the parameter.

b) Maximum Likelihood Estimator
The Likelihood function is obtained by finding the product of the probability distributions and then after taking log first derivative is taken and equated to zero to find the estimator the maximises the likelihood.

Q.9.] i) $t_K = \dfrac{N(0,1)}{\sqrt{\chi^2_K / K}}$.

$N(0,1)$ - is the standard normal random variable
χ^2_K - is the χ^2_K random variable with degrees of freedom K.
$N(0,1)$ & χ^2_K are independent   K - degrees of freedom.

ii) Mean and Variance of t_K, $K > 2$
t_K distribution
Mean = 0
Variance = $\dfrac{K}{(K-2)}$.

iii) A Sample is taken from normal population
$\bar{X} = 50$ (Sample mean) $S^2 = 48.667$ (Sample variance) $n = 10$

a) We know that
$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Page No.
Date

$\therefore$ 99% Confidence interval is as follows:-

from the table book

~~0.99 =~~ $\;\; \cancel{t_{9,\,0.5\%}} = \cancel{3.250}$

~~upper tail~~

$\alpha = 1\%$ $\quad\therefore\quad t_{9,\,0.5\%} = 3.250$ (upper tail)

$\alpha/2 = \frac{1}{2}\% = 0.5\%$

As it is a symmetrical interval

$t_{9,\,0.5\%} = -3.250$ (lower tail)

$$\therefore 0.99 = P\left(-3.250 < \frac{\overline{X} - \mu}{S/\sqrt{n}} < 3.250\right)$$

$$\therefore 0.99 = \left(\overline{X} - 3.250 \cdot \frac{S}{\sqrt{n}} < \mu < \overline{X} + 3.250 \cdot \frac{S}{\sqrt{n}}\right)$$

$$= \left(50 - 3.250 \cdot \frac{6.98}{\sqrt{10}} < \mu < 50 + 3.250 \cdot \frac{6.98}{\sqrt{10}}\right)$$

$$= (42.826 < \mu < 57.174)$$

$\therefore$ 99% C.I for the population mean $(42.83\,,\,57.17)$

b) Parameters as result from part (ii)

Mean = 0 $\quad$ Variance $= \frac{K}{(K-2)} = \frac{9}{9-2} = \frac{9}{7} \approx 1.286$

$\therefore$ taking thes values

~~99~~ C.I $= 0 \pm 3.250 \cdot \frac{\sqrt{1.286}}{\sqrt{10}}$

$\therefore$ 99% C.I is given by $(-1.165\,,\,1.165)$

Page No.
Date

Q.10] X_i's ~ No. of claims in a year
X_i's ~ Poi. (μ)

1000 policies were observed. $\mu_{rate} = 3$.

IF $X < 3100$, accept $\mu = 3$
$X \geq 3100$, reject $\mu = 3$.

i) Type 1 error can be defined as the probability of rejecting Ho, when it is true.

In the above scenario.
X is the random variable. - Total no. of claims.
~~P(Type 1 error) = P(Reject Ho when Ho true)~~
~~= P(X ≥ 3100, when X~~
$X \sim Poi(n\mu)$ i.e $X \sim Poi(3000)$
Ho: (Null hypothesis).

$\therefore$ P (Type 1 error) = P(Reject Ho when Ho true)
$= P(X > 3100, X \sim Poi(3000))$
As $n\mu = 3000$ (large)
$\therefore$ CLT can be applied
$\therefore X \sim N(3000, 3000)$
$\therefore \frac{X - n\mu}{\sqrt{n\mu}} \sim N(0,1)$

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

$= P(Z > 1.825)$
$= 1 - \phi(1.825) = 1 - 0.966$
$= 0.034$

$\therefore \phi(1.82) = 0.96562$
0.00076
0.01
$\phi(1.83) = 0.96638$
0.005

Page No.
Date

ii) Type II Error is the probability of accepting Ho when in fact it is false.

iii) Power of a test is defined as $(1-\beta)$

i.e The probability of rejecting Ho, when it is false

In terms of μ

P(reject Ho, when Ho is false) = $P(X > 3100$, when $X \sim (n\mu))$

$\therefore P(X > 3100)$

By CLT

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

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

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

iv) No. of claims (actual) observed = 2900.

~~As X < 2900~~

$n\mu = 2900$

$$\therefore \hat{\mu} = \frac{2900}{1000} = 2.9.$$

$\therefore$ Let $\hat{\mu}$ be the estimator of μ

$$\therefore \hat{\mu} \sim N(\mu, \hat{\mu}/n)$$

from the table book

$$0.99 = \left(-2.5758 < \frac{\hat{\mu}-\mu}{\sqrt{\hat{\mu}/n}} < 2.5758\right)$$

Page No.
Date

$$= \left(~~2.5758~~\ \hat{\mu} - 2.5758\sqrt{\frac{\hat{\mu}}{n}} < \mu < \hat{\mu} + 2.5758\sqrt{\frac{\hat{\mu}}{n}} \right)$$

$\therefore$ 99% CI is given by

$$\left(2.9 - 2.5758 \cdot \sqrt{\frac{2.9}{1000}} ,\ 2.9 + 2.5758\sqrt{\frac{2.9}{1000}} \right)$$

i.e

$(2.76, 3.04)$

Q.11] Random Sample of basic salaries of 12 employees is selected

$\therefore n = 12$

$\Sigma x = 213,200 \quad \Sigma x_i^2 = 4,919,860,000$

Sample mean $\bar{x} = 17,767$

Sample Standard deviation $S = 10,144$.

Salary of two ~~Sample~~ temporary employees are 11,000 & 48000 respectively

Salary of two randomly selected permanent employees is 27,500 & 31,500

$\therefore \bar{X} = \frac{\Sigma x}{n}$

~~Revised Sample mean = Σx -~~

$\therefore$ Revised Σx = Σx − Salary of temporary employees + Salary of randomly selected permanent employees

$= 213,200 - (11,000 + 48,000) + (27,500 + 31,500)$.

$= 213200 - 59,000 + 59000$

$= 213,200$

$\therefore$ Revised sample mean $= \frac{213,200}{12} = 17,766.667$

≈ 17767

Revised $\Sigma x^2 = \Sigma x^2 -$ (Sum of squares of temporary employee salaries) + (Sum of squares of permanent Employee Salaries)

$= ~~213~~\ 4,919,860,000 - (11000^2 + 48000^2)$

$= \boxed{4243360000} + (27,500^2 + 31500^2)$

$\therefore$ Revised variance $= \dfrac{\sum x_i^2 {}_{\text{revised}} - n\bar{x}^2}{n-1}$

$= \dfrac{4243360000 - 12 \times 17767^2}{12-1}$

$= 41396775.64$

$\therefore$ Revised sample standard deviation $= \sqrt{41396775.64}$

$= 6434.03$

We can observe that the sample mean remains unchanged however the sample standard deviation has decreased.

Q.12] $X_1, X_2, \ldots X_n$ a random sample from Uniform distribution over $(0, \theta)$

θ - unknown parameter (>0).