DATE
PAGE

Probability & Statistics - 2
Assignment - 1

[Roll no. - 38]

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

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

$$\hat{\theta} = \frac{\Sigma x}{n} = \frac{200}{1} = 200 \quad \text{(for poisson)}$$

since only a single value is given,
so $\overline{x} = 200$

95% CI :

$$\overline{x} \pm 1.96 \sqrt{\frac{200}{1}}$$

$$= (172.28, 227.72)$$

2. $X_i = 1, 2, \ldots, n$ $n \to$ iid

$\overline{x} = \sum_1^n x_i$ mean $= \mu$

var $= \sigma^2$

$$s^2 = \frac{1}{n-1} \sum_i^n (x_i - \overline{x})^2$$

(i) Mean :

$$E(\overline{x}) = E\left[\frac{\Sigma(x_i)}{n}\right]$$

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

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

DATE
PAGE

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

variance:

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

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

$$\boxed{V(\bar{x}) = \frac{\sigma^2}{n}}$$

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

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

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

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

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

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

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

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

DATE
PAGE

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

$= V\left(\chi^2_{n-1}\right)$

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

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

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

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

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

$\Rightarrow P(100-50 < s^2 < 100+50)$

$= P(50 < s^2 < 150)$

$= P\left(\frac{50(9)}{100} < \frac{9s^2}{\sigma^2} < \frac{150(9)}{100}\right)$

$= P(4.5 < \chi^2_9 < 13.5)$

$= P(\chi^2_9 < 13.5) - P(\chi^2_9 < 4.5)$

$= 0.8587 - 0.1245$

$= \underline{\underline{0.7342}}$

3. $X_i \sim Bin(4, p)$

PDF of binomial : ${}^nC_x p^x q^{n-x}$

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

$$L(p) = \prod_{i=1}^{n} {}^nC_x p^x (1-p)^{n-x}$$

$$= \text{constant} \times p^x (1-p)^{n-x}$$

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

$$= \text{constant} \left[(1-p)^{344} (p^{75}(1-p)^{225}) (p^{32}(1-p)^{32}) (p^6(1-p)^2) p^4 \right]$$

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

taking log on both sides:

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

differentiating both sides w.r.t. p.

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

DATE
PAGE

$\rightarrow$ put it as 0

$\Rightarrow \hat{p} = \dfrac{117}{117 + 603}$

$\hat{p} = \dfrac{117}{720} = 0.1625$

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

checking for maximum :

$\dfrac{d^2 \ln L}{dp^2} = -\dfrac{603}{(1-p)^2} - \dfrac{117}{p^2} < 0$

$\rightarrow$ Max value of $\ln L$ is at $p = 0.1625$.

(ii) Goodness of fit Test

H_0 : distbn is binomial ~~with~~ $\rightarrow Bin(4,p)$

H_1 : distbn is not binomial $\rightarrow Bin(4,p)$

using $\hat{p} = 0.1625$ from part (a), the probabilities for this binomial distbn. are:

$P(X=0) = (1-p)^4 = 0.49197$

$P(X=1) = 4p(1-p)^3 = 0.38183$

$P(X=2) = 6p^2(1-p)^2 = 0.11113$

$P(X=3) = 4p^3(1-p) = 0.01437$

$P(X=4) = p^4 = 0.00070$

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

$\hat{p} = \underline{\underline{0.1625}}$

DATE
PAGE

$P(X=0) = {}^4C_0\, p^0 (1-p)^4 = 0.4919$
$P(X=1) = {}^4C_1\, p^1 (1-p)^3 = 0.3813$
$P(X=2) = 6p^2 (1-p)^2 = 0.1111$
$P(X=3) = 4p^3 (1-p) = 0.0143$
$P(X=4) = p^4 = 0.0006$

	0	1	≥ 2
O_i	86	75	19
E_i	88.542	68.774	22.68

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

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

at 5% Los: $Z_{tab} = 3.841$

$Z_{cal} < Z_{tab}$

$\Rightarrow$ We do not reject H_0

$\Rightarrow$ No. of claims follow binomial distribution $\rightarrow$ Bin(4, p)

DATE
PAGE

4. (i) $f(x) = \dfrac{c\beta^3}{(x+\beta)^4}$: $x > 0$, $\beta > 0$

$c \rightarrow$ constant

$\beta \rightarrow$ parameter

On comparing with pareto distribution:

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

we determine : $\boxed{c = 3}$

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

$$V(X) = \boxed{3/4 \, \beta^2}$$

(ii) sample mean $= \dfrac{\sum_{i=1}^{n} x_i}{n} = \overline{X}_n$

and population mean (calculated in (i)) $= \beta/2$

according to method of moments :-

$$\overline{X}_n = \beta/2$$

$$\boxed{\hat{\beta} = 2\overline{X}_n}$$

$$\text{Bias} = E[g(X)] - \theta$$

$E[g(X)] = E[\hat{\beta}] = E[2\overline{X}_n] = 2E[\overline{X}_n] =$
$2E[X] = \cancel{2}(\beta/\cancel{2}) = \beta$

$$\text{Bias} = \beta - \beta = \underline{\underline{0}} \rightarrow \text{unbiased}$$

DATE
PAGE

Mean square error $= Var(\hat{\beta}) + (Bias(\hat{\beta}))^2$

$$Var(\hat{\beta}) = Var[2\overline{X}_n] = 4\,Var[\overline{X}_n]$$

$$= 4\,Var\left[\frac{\Sigma x_i}{n}\right] = \cancel{4 Var(\quad)}$$

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

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

Hence $MSE = \frac{3\beta^2}{n} + 0 = \frac{3\beta^2}{n}$.

(iii) $$MSE[b\overline{X}_n] = Var[b\overline{X}_n] + (Bias[b\overline{X}_n])^2$$

$$= b^2\,\frac{3}{4}\frac{\beta^2}{n} + (E[b\overline{X}_n - \beta))^2$$

$$= \frac{b^3\,3\beta^2}{4n} + \left(\beta\left(\frac{b}{2} - 1\right)\right)^2$$

$$= \frac{\beta^2}{n}\left(\frac{3}{4}b^2 + n\left(\frac{b}{2} - 1\right)^2\right)$$

differentiating the MSE wrt b :-

$$\frac{d(MSE)}{db} = \frac{\beta^2}{n}\left(\frac{3}{2}b + \cancel{2}n\left(b/2 - 1\right)\left(1/\cancel{2}\right)\right)$$

equate to 0.

$$\Rightarrow b = \frac{2n}{(n+3)} = \frac{2}{(1 + 3/n)}$$

DATE
PAGE

differentiating the MSE a second time respect to b,

$$\frac{d^2(MSE)}{db^2} = \frac{\beta^2}{n}\left(3/2 + n(1/2)\right)$$

> 0

$\rightarrow$ a minimum value of MSE is at

$$b = \frac{2}{1 + 3/n}$$

(iv)

5. (i) sample mean $= \bar{x}$

sample ~~variation~~ std. deviation $= s$

popln mean $= \frac{1}{2}(a+b)$

popln s.d $= \frac{b-a}{\sqrt{12}} = \frac{b-a}{2\sqrt{3}}$

acc. to method of moments

mean:

$$\bar{x} = \frac{1}{2}(a+b)$$

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

$$b = 2\bar{x} - a \quad -(1)$$

DATE
PAGE

std. dev.

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

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

$$a = b - 2\sqrt{3}\, s \quad — ②$$

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

from ①

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

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

(ii) using ②:

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

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

~~$2\bar{x}$ ... $-2b$~~

[using ~~②~~
$a = 2\bar{x} - b$
from ①]

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

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

(iii) Sample: 1, 2, 3, 4, 60

$$\bar{x} = 14$$

$$s = 25.74$$

Method of moments estimates using above formulae are:

DATE
PAGE

$$\hat{a} = 14 - \sqrt{3}(25.74)$$

$$= -30.583$$

$$\hat{b} = 14 + \sqrt{3}(25.74)$$

$$= 58.583$$

for $U(a, b)$ probability of sample point beings less than 'a' or greater than 'b' is zero & we have a sample value 60 which is greater than estimated b.
This shows potential weakness of method of moments.

(iv) $\underline{a = 0}$ sample : $x_1, x_2 \dots x_n$

$$f(x) = \frac{1}{b-a} = \frac{1}{b}$$ for this particular case.

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

$$L = f(x_1) f(x_2) \dots f(x_n)$$

$$L = \frac{1}{b} \times \frac{1}{b} \dots \frac{1}{b} \text{ n times}$$

$$L = \frac{1}{b^n}.$$

taking log on both sides

$$\ln L = \ln 1 - n \ln b$$

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

DATE
PAGE

differentiating wrt b

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

$\frac{dl}{db} = 0 \rightarrow b = \infty$, which is not possible.

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

$\frac{n}{b^2} > 0 \Rightarrow$ minimum likelihood estimate

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

~~6.~~ $H_0 : p = 0.1$

~~$H_1 : p > 0.1$~~

~~One - sided test~~

7. (i) sample mean $= \frac{\Sigma x_i}{n}$

Sample mean for ~~private~~ public hospital = $\boxed{30}$

sample mean for private hospital = $\boxed{35.055}$

DATE

sample variance $= \cancel{\frac{1}{n-1}\left[\sum \bar{x}_i^2 - (\sum x_i)^2\right]} = \frac{\sum (x-\bar{x})^2}{n-1}$

$$\boxed{\begin{aligned} s^2_{public} &= 64.3125 \\ s^2_{private} &= 67.4028 \end{aligned}}$$

(ii) 2-sided test.

$H_0 : \sigma_1^2 = \sigma_2^2$ vs $H_1 : \sigma_1^2 \neq \sigma_2^2$

T.S. : $\dfrac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \sim F_{n_1 + n_2 - 1}$

value of statistics : $\dfrac{64.31}{67.42} = 0.9542$

$F_{8,8}$ values at 5% levels are 0.2256 & 4.433.

We do not reject H_0.

(iii) $H_0 : \mu_{pr} = \mu_{pu}$

$H_1 : \mu_{pr} > \mu_{pu}$

$$TS = \frac{(\bar{x}_2 - \bar{x}_1) - (\mu_2 - \mu_1)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{35.055 - 30}{8.1152\sqrt{\frac{2}{9}}}$$

$TS = 1.32151$

$Z_{tab} = 1.746$

$\therefore Z_{cal} < Z_{tab}$

we do not reject H_0.

8. (i) Unbiased estimator:
Consider a random sample $X = (X_1, X_2, X_3 \ldots X_n)$ from a distribution with parameter θ and $\hat{\theta}$ as an estimator, property of unbiasedness states that:

$$E(\hat{\theta}) = \theta$$

⇒ Mean value equals the true parameter value.

(ii) 'Bias' – a measure of the difference b/w the expected value of the estimator and parameter being estimated.

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

(iii) MSE is defined by:

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

used for comparing estimators generally.

$$MSE = \text{variance} + \text{bias}^2$$

iv) $\tilde{\theta}$ with positive bias is an efficient estimator.
Biased estimator whose value does not deviate (high MSE) very far from true value would be preferable to an unbiased one with high MSE.

DATE
PAGE

v)

Estimators ~~Parameters~~ will be consistent when MSE $\to 0$ & $n \to \infty$.
(n = size of sample)

(vi) ① Method of moments estimator → equating population moments to corresponding sample moments.

~~② Maximum of likelihood estimator~~ :

② Method of maximum likelihood :

9. (i) $t_k \Rightarrow \dfrac{N(0,1)}{\sqrt{\chi^2_k / k}}$

- k → degree of freedom.
- $N(0,1)$ & χ^2_k are 2 independent random variables.

(ii) for $k > 2$

$$E(t_k) = 0$$

$$V(t_k) = \frac{k}{k-2}$$

(iii) (a) $\dfrac{\bar{x} - \mu}{s/\sqrt{n}} \sim t_{n-1}$

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

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

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

99% confidence interval for μ =

$$\left(50 - 3.25 \times \frac{\sqrt{48.667}}{\sqrt{10}},\ 50 + 3.25 \times \sqrt{\frac{48.667}{10}}\right)$$

$$= (42.83,\ 57.17)$$

(b) from (ii) part,

$$t_{n-1} \sim N\left(0, \sqrt{\frac{n-1}{n-3}}\right)$$

$$t_{n-1} \sim N\left(0, \sqrt{\frac{9}{7}}\right)$$

i.e.

$$\frac{\bar{X} - \mu}{s/\sqrt{7n/9}} \sim N(0,1)$$

DATE
PAGE

Critical values level of confidence 2.58

$$CI = \left(50 - 2.58\sqrt{\frac{48.667}{10} \times \frac{9}{7}},\ 50 + 2.58\sqrt{\frac{48.667 \times 9}{10 \times 7}}\right)$$

Confidence interval = (43.54, 56.45)

11. given: $n = 12$ $\quad \Sigma x^2 = 4{,}919{,}860{,}000$

$\Sigma x = 213{,}200$ $\quad \bar{x} = 17{,}767$

$s = 10{,}144.$

New sample mean:

$$\Sigma x_{new} = \Sigma x - 11000 - 48000 + 27500 + 31500$$

$$= 213200$$

$$\bar{x}_{new} = \frac{213200}{12}$$

$$\bar{x}_{new} = 17{,}767$$

New sample sd:

$$\Sigma x^2_{new} = 4919860000 - (11000)^2 - (48000)^2 + (27500)^2 + (31500)^2$$

$$= 4243360000$$

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

$$= \frac{1}{12-1}\left(4243360000 - 12(17767)^2\right)$$

$$= 41396775.64$$

$$s_{new} = \sqrt{s^2_{new}} = 6434.032611$$

Comments:

- Sample mean remains unchanged even after addition of salaries of two permanent employees and substraction of ~~start~~ salaries of two temporary employees.

- Sample deviation reduces indicating that outliers do not deviate much from mean.

12. (i) Cramér - Rao lower bound result holds under very general conditions. But, it does not apply in cases where the range of the distn involves parameters, such as the uniform distribution.

Therefore, CRLB does not apply in this case.

(ii) Given: $Y = \max X_i$
for $0 < y < \theta$

$$P[Y < y] = P[\max X_i < y]$$

$$= P\left[\prod_{i=1}^{n} (X_i < y)\right]$$

$$= \prod_{i=1}^{n} P(X_i < y) \quad (\because x_i \text{ are iid})$$

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

DATE
PAGE

PDF:

$$g_Y(y) = \frac{d}{dy} P(Y \leq y) = \frac{n}{\theta^n} \cdot y^{n-1}$$

for any non-negative real number k.

$$E[Y^k] = \int_0^\theta y^k \frac{n}{\theta^n} y^{n-1}\, dy \quad \left(\text{multiply \& divide } n+k\right)$$

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

$$= \frac{n}{n+k} \cdot \frac{\theta^{n+k} - 0}{\theta^n}$$

$$\boxed{E[Y^k] = \frac{n\,\theta^k}{n+k}}$$

(iii) Bias $= E[\hat{\theta}(c) - \theta]$

$= E[cY - \theta]$

$= c \cdot E(Y) - \theta$

$= c \cdot \frac{n\theta}{n+1} - \theta \qquad (k=1)$

$$\boxed{\text{Bias} = \theta\left(\frac{cn}{n+1} - 1\right)}$$

MSE $= E[(\hat{\theta}(c) - \theta)^2]$

$= E[(cY - \theta)^2]$

$= E[c^2Y^2 - 2c\theta Y + \theta^2]$

$= c^2 E(Y^2) - 2c\theta E(Y) + \theta^2$

$$\boxed{\text{MSE} = c^2 \frac{n\theta^2}{n+2} - \frac{c \cdot 2n\theta^2}{n+1} + \theta^2}$$

$(k = 2, 1)$

(iv) for $\hat{\theta}(c)$ to be unbiased :-

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

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

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

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

$$c_u n = n+1$$

$$\boxed{c_u = \frac{n+1}{n}}$$

v) To minimise $MSE[\hat{\theta}(c)]$:

$$\frac{d\, MSE[\hat{\theta}(c)]}{dc} = 0$$

$$\Rightarrow \frac{d}{dc}\left(c^2 \frac{n\theta^2}{n+2} - \frac{c2n\theta^2}{n+1} + \theta^2\right)$$

$$= 2c\frac{n\theta^2}{n+2} - \frac{2n\theta^2}{n+1} = 0$$

(equating to 0)

$$\hookrightarrow \frac{\cancel{2}c\cancel{n\theta^2}}{n+2} = \frac{\cancel{2n\theta^2}}{n+1}$$

$$\boxed{c = \frac{n+2}{n+1}}$$

now,

$$\frac{d^2}{dc^2} MSE(\hat{\theta}(c)) = \frac{d}{dc}\left[\frac{2c\, n\theta^2}{n+2} - \frac{2n\theta^2}{n+1}\right]$$

$$= \cancel{\frac{2n\theta^2}{n+2}} \quad \frac{2n\theta^2}{n+2} > 0$$

$\Rightarrow$ minimum

(vi) To minimize error in estimation, it is preferred that estimator has a lower mean square error among different competing estimators.

Here, we prefer $\hat{\theta}(c_m)$ over $\hat{\theta}(c_u)$

As n becomes large:

$$\hat{\theta}(c_u) = \frac{n+1}{n} Y \longrightarrow Y$$

$$\text{IIIy} \quad \hat{\theta}(c_m) = \frac{n+2}{n+1} Y \rightarrow Y$$

Thus, 2 estimators become one & same as $n \rightarrow \infty$.

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

(i) P(Type 1 error) = P(rejecting H_0 when H_0 is true)

let X be no. of hits in first step 12 missiles & Y

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

$= (1 - 0.8891) + (0.2824)(1 - 0.9957) + (0.6590 - 0.2824)(1 - 0.9744) + (0.8891 - 0.6590)(1 - 0.8891)$

$= 0.14727 \approx 15\%$

(ii) Probability of rejecting the null hypothesis when $p = 0.3$

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

$= 0.9125 \approx 91\%$

(iii) $P(\text{type II error})$ is the probability of accepting H_0 when H_0 is false.
$= 1 - 0.9125 = 0.0874 \approx 9\%$

(iv) 10 space agencies in aggregate fired 120 missiles & recorded 40 hits.

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

DATE
PAGE

$n = 120 \quad x = 40 \quad \hat{p} = \frac{40}{120} = \frac{1}{3}$

$= 0.3333$

$z_{10\%} = 1.2816$

lower bound of 90% upper-tailed CI for p:

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

$$= \underline{\underline{0.2782}}$$

(v) $p > 0.278$ at 10% LOS

$H_0 : p = 0.278$

vs $H_1 : p > 0.278$

One sample binomial:

$$\frac{X - np_0}{\sqrt{np_0q_0}} \sim N(0,1) \text{ with continuity correction}$$

$$\frac{39.5 - 120(0.278)}{\sqrt{120(0.278)(0.722)}} = 1.2511$$

We do not reject H_0.

(vi) Lower bound of CI implies that there is only a 10% chance of $p \le 0.278$ whereas from hypothesis test, p could be less than or equal to 0.278 with probability more than 10%.

DATE
PAGE

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

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{1/n_1 + 1/n_2}} \sim t_{n_1+n_2-2}$$

$$s^2 p = \frac{1}{12} \left(11(2916) + 14(4900) \right) = 4027.04.$$

$$\frac{65 - 70 - 0}{63.459 \sqrt{\frac{1}{12} + \frac{1}{15}}} = -0.2034$$

we do not reject H_0.

10. $X \sim Poi(3000)$ $n = 1000$
$\lambda = 3$

(i) $P(\text{Type 1 error}) = P(\text{reject } H_0 | H_0 \text{ is true})$

$= P(X > 3100 \mid X \sim Poi(3000))$

$$P(X > 3100) = P\left(Z > \frac{3100 - 300}{\sqrt{3000}}\right)$$

$= 1 - P(Z < 1.825)$
$= 3.39\%$

DATE
PAGE

(ii) Type II error $\Rightarrow$ accepting H_0 when it is false.

(iii) Power of test = P(rejecting H_0 | H_0 is false)

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

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

This value depends on value of parameter.

(iv) $\hat{\mu} \sim N\left(\mu, \dfrac{\hat{\mu}}{n}\right)$

$\Rightarrow \dfrac{\hat{\mu} - \mu}{\sqrt{\hat{\mu}/n}} \sim N(0,1)$

$= P\left(-2.5758 < \dfrac{2.9 - \mu}{\sqrt{2.9/1000}} < 2.5758\right)$

$= 0.99$

CI $\doteq$ (2.7613, 3.0387).