Name Ronit shah R. No. 67 Class / Div. Date

Subject Probability and statistics-2 Topic Assignment - 1 Page No.

Assignment - 1

1) $x \sim poi(\theta)$

$\frac{x - \theta}{\sqrt{\theta}} \sim N(0,1)$

Given : 95% symmetrical confidence interval, $\lambda = 200$

$P[x_1 < \theta < x_2] = 0.95$

$x \sim poi(\lambda)$

$\therefore x \sim poi(200)$

95% confidence interval $= \mu \pm 1.96\sqrt{\sigma}$

we know that, $\lambda = \mu = \sigma$

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

$= 200 + 1.96\sqrt{200}$; $200 - 1.96\sqrt{200}$

95% CI $= 172.28$; 227.71

$\therefore$ The confidence interval is 172.28 and 227.71

2) sample mean $(\bar{x}) = \sum_i^n x_i$; sample variance $(s^2) = \frac{1}{n-1} \sum_i^n (x_i - \bar{x})^2$

i) $E(\bar{x}) = E\left(\frac{\sum x_i}{n}\right) = \frac{1}{n} \sum E(x_i) = \frac{n\mu}{n} = \mu$

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

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

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

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

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

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

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

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

$= \sigma^2$

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

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

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

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

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

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

2) iv) Given : $n = 10$; $\sigma^2 = 100$

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

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

s^2 will fall between $\pm$ 50% $E(s)$

$\therefore P(s^2 \leq 50)$

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

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

$= 0.1245$

$P(s^2 \leq 150)$

$$= P\left[\frac{(n-1)s^2}{\sigma^2} \leq \frac{150 \times 9}{100}\right]$$

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

$= 0.8587$

$\therefore P(50 \leq s^2 \leq 150) = 0.8587 - 0.1245$

$= 0.7342$

3)

No of claims	No of Policy	O_i	$(O_i - E_i)^2/E_i$
0	86	$^{86}/_{180} = 0.4778$	0.000404
1	75	0.4167	0.00319
2	16	0.0889	0.004436
3	2	0.0111	0.000756
4	1	0.0055	0.03308
	180		0.04187

i) $L(p) = \left({}^4c_0 \cdot p^0 \cdot (1-p)^4\right)^{86} \left({}^4c_1 \cdot p^1 \cdot (1-p)^3\right)^{75} \left({}^4c_2 \cdot p^2 \cdot (1-p)^2\right)^{16}$

$\left({}^4c_3 \cdot p^3 \cdot (1-p)^1\right)^2 \cdot \left({}^4c_4 \cdot p^4 \cdot (1-p)^0\right)^1$

$= \left(1 \cdot 1 \cdot (1-p)^4\right)^{86} \left(4 \cdot p \cdot (1-p)^3\right)^{75} \left(6 \cdot p^2 \cdot (1-p)^{20}\right)^{16} \cdot$

$\left(4 \cdot p^3 \cdot (1-p)^{\cancel{2}}\right)^2 \cdot \left(\cancel{\,} 1 \cdot p^4\right)^1$

$= \left[(1-p)^{344}\right] \cdot \left[4^{75} \cdot p^{75} \cdot (1-p)^{225}\right] \cdot \left[6^{16} \cdot p^{32} \cdot (1-p)^{320}\right]$

$\left[4^2 \cdot p^6 \cdot (1-p)^2\right] \cdot \left[p^4\right]$

$= 4^{75} \cdot 6^{16} \cdot 4^2 \cdot p^{117} \cdot (1-p)^{603}$

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

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

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

$117 - 117p - 603p = 0$

$720p = 117$

$p = 0.1625$

3) ii) $P(x=0) = {}^4c_0 . p^0 . (1-p)^4 = 91.34$

$P(x=1) = {}^4c_1 . p^1 . (1-p)^4 = 67.53$

$P(x=2) = {}^4c_2 . p^2 . (1-p)^4 = 18.72$

$P(x=3) = {}^4c_3 . p^3 . (1-p)^4 = 2.41$

$P(x=4) = {}^4c_4 . p^4 . (1-p)^4 = 0.12$

~~Probability~~	x	O_i	E_i	$(O_i - E_i)^2 / E_i$
~~P(x=0)~~	0	86	91.34	0.3123
~~P(x=1)~~	1	75	67.53	0.8263
~~P(x=2)~~	2	19 ~~16~~	21.25 ~~18.72~~	0.2382 ~~0.3952~~
~~P(x=3)~~	~~3~~	~~3~~	~~2.41~~	~~0.06915~~ 1.3768
	~~4~~	~~0~~	~~0.12~~	

$\chi^2_{cal} = 1.3768$

$\chi^2_{Tab} = 0.75$

$\therefore \chi^2_{cal} > \chi^2_{Tab}$

Thus, we will Reject H_0 at 5%.

4) ii) $\beta = x_1, x_2, x_3, \ldots \ldots x_n$

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

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

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

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

$= \beta$

$\therefore$ The Results are unbiased

4) iii) $MSE = Var(b\bar{x}_n) + Bias(b\bar{x}_n)^2$

$= b^2 \times 3/4 \times \beta^2/2 + \left[E(b\bar{x}_n) - \beta \right]^2$

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

iv) As we saw ~~earliear~~ earlier, $\hat{\beta} = 2\bar{x}_n$ is an unbiased estimator.

5) i) $x \sim U(a+b)$

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

$s = \frac{1}{\sqrt{12}}(b-a)$

$\sqrt{12}\, s + a = b$

$s(\sqrt{12}) + a = b$

$\bar{x} = \frac{1}{2}(s(\sqrt{12}) + a + a)$

$= \frac{1}{2}(2a + \sqrt{4 \times 3}\, s)$

$= \frac{1}{2}(2a + 2\sqrt{3}s)$

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

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

5) ii) $\bar{Y} = \frac{1}{2}(a+b)$

$S = \frac{1}{\sqrt{12}}(b-a)$

$a = b - \sqrt{12}\,S$

$\bar{Y} = \frac{1}{2}\left[b - \sqrt{12}\,S + b\right]$

$\bar{Y} = \frac{1}{2}\left[2b - 2\sqrt{3}\,S\right]$

$\therefore \bar{Y} = b - \sqrt{3}\,S$

$\therefore \hat{b} = \bar{Y} + \sqrt{3}\,S$

6) i) $P(\text{Type 1 error}) = P[\text{Reject } H_0 \mid H_0 \text{ is True}]$

$= P[x \geq 3 \mid p = 0.1] + P[x = 0 \mid p = 0.1]$

$+ P[x \geq 4 \mid p = 0.1] + P[x = 2 \mid p = 0.1]$

$+ P[x \geq 5 \mid p = 0.1] + P[x = 1 \mid p = 0.1]$

$= 0.14727$

$P(\text{Type 1 Error}) \approx 15\%$

ii) $P(x \geq 3 \mid p = 0.3) + P(x = 0 \mid p = 0.3) = 0.9125 = 91.25\%$

iii) $P(\text{Type 2 error}) = 1 - 0.9125 = 0.0875$

7) i) Public :

$\bar{x} = 30$

$s^2 = 64.31$

Private :

$\bar{x} = 35.06$

$s^2 = 67.402$

ii) $H_0 : \bar{x}_1 = \bar{x}_2$

$H_1 : \bar{x}_1 \neq \bar{x}_2$

$$\frac{(30 - 35.06)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \doteq \frac{(30 - 35.06)}{\sqrt{8.115}\ \sqrt{0.222}}$$

$$S_p^2 = \frac{8(64.31) + 8(67.402)}{16} - 65.857$$

$\therefore z = (-2.624, 2.624)$

$\therefore$ Do not Reject H_0

iii) Private Hospital treatments does result in higher claim size.

8) i) An unbiased estimator of a Parameter is an estimator whose expected value is equal to the Parameter.

ii) Bias of an estimator is the difference between this estimator's expected value and the true value of the Parameter being estimated.

iii) Mean square Error is a measure of how close a fitted line is to data points.

8) i) unbiased Estimator = $E(\hat{\theta}) = \theta$

ii) $Bias = E(\hat{\theta}) - \theta$

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

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

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

$= Var(\hat{\theta}) + bias^2(\hat{\theta})$

$\therefore MSE(\hat{\theta}) = Var(\hat{\theta}) + bias^2(\hat{\theta})$

9) i) $z \sim N(0,1)$

$$t_k \sim \frac{z}{\sqrt{\left(\frac{\chi^2_k}{k}\right)}}$$

χ^2_k = Chi-square Distribution with k degrees of freedom

ii) $E(t_k) = 0$

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

iii) $n = 10$; $\bar{x} = 50$; $s^2 = 48.667$

Population mean at 99% confidence interval are:

a) $CI = 50 \pm 3.25\sqrt{\frac{48.667}{10}} = (42.81, 57.17)$

b) $CI = 50 \pm 2.5758\sqrt{\frac{48.667}{10}} = (44.32, 55.68)$

10) i) Type 1 Error $= P[H_0 \text{ reject} \mid H_0 \text{ is true}]$

$$= P[x > 3100 \mid x = 3]$$

$$= P\left[z > \frac{3100 - 3000}{\sqrt{3000}}\right]$$

$$= 1 - P[z < 1.8257]$$

Type 1 Error = 3.89%

10) (v) Confidence interval for 99% confidence interval:

$CI = \cancel{3\pm}\ 3 \pm 2.5758 \sqrt{3/2900}$

~~CI is (2.9~~ $CI = 2.9172$; $CI = 3.0828$

11) Given:

$\Sigma x = 213200$; $\Sigma x^2 = 4919860000$

Sample mean = 17767 ; Sample standard Deviation = 10144

New $\Sigma x = 213200 - 11000 - 48000 + 27500 + 31500$

$= 213200$

~~New $\bar{x}$ = 2132~~

New $\bar{x} = \dfrac{213200}{12} = 17767$

New $\Sigma x^2 = 4919860000 - (11000)^2 - (48000)^2 + (27500)^2 + (31500)^2$

$\Sigma x^2 = 4243360000$

New $\bar{x}^2 = 315666289$

$s = \sqrt{\dfrac{1}{n-1} \cdot \Sigma x^2 - n(\bar{x})^2}$

$s = \sqrt{\dfrac{1}{12-1} \cdot 4243360000 - 12(315666289)}$

$s = 6434.03$

Thus, As we can observe, change in salary values of employees, the standard Deviation has changed. It is less than the sample mean.