# Probability & Statistics - 2
## Assignment - 1

Name : Shagun Alag
Class : F.Y. B
Roll no. : 97

1) $X \sim Pois(\theta)$

~~$x = 200$~~

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

$\cancel{\theta = \hat{\theta} \pm Z_{\alpha/2}\sqrt{\frac{\hat{\theta}}{n}}}$

$\theta = 200 = \bar{x} = \sigma$, $\alpha = 5\%$

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

$C.I. = \bar{x} \pm Z_{\alpha/2}\sqrt{\frac{\bar{x}}{n}}$

$C.I. = 200 \pm 1.96\sqrt{\frac{200}{1}}$

$= 200 \pm 1.96(10\sqrt{2})$

$\therefore C.I. = (172.28, 227.72)$

2) $X_i = 1, 2, 3, \ldots, n$, mean $= \mu$, variance $= \sigma^2$

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) $\cancel{E(\bar{x}) = E\left[\frac{\sum_i^n x_i}{n}\right]} = \frac{1}{n} E\left[\sum_i^n x_i\right] = 1$

$\cancel{= \sum_i E[x_i] = \sum \mu}$

i) Mean of $\bar{x} = E[\bar{x}] = E\left[\frac{\sum_i^n x_i}{n}\right]$

$= \frac{1}{n}[E(\sum_i^n x_i)] = \frac{1}{n}[\sum E(x_i)]$

$= \frac{1}{n}(\sum \mu) = \frac{1}{n}(n\mu)$

Mean of $\bar{x} = E[\bar{x}] = \underline{\underline{\mu}}$

Variance of $\bar{x} = V\left[\frac{\sum x_i}{n}\right] = \frac{1}{n^2} V[E(\bar{x})]$

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

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

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

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

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

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

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

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

~~iii)~~ $\cancel{X_i \sim N(\theta)}$. $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

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

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

$$V[S^2] = \frac{2\sigma^4}{n-1}$$

iv) $n = 10, \sigma^2 = 100$, $\cancel{\alpha = 50\%}$, $P = 50\%$.

$$\frac{(n-1)S^2}{\sigma^2} = 0.5$$

$$\frac{9(10-1)S^2}{100} = 0.5$$

$$9 \cdot S^2 = 50$$

$$9 \cdot S^2 = \frac{50}{9}$$

$$S^2 = 5.56\%$$

$$S^2 = 0.0556\%$$

3)

| No. of claims | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| No. of policies | 86 | 75 | 16 | 2 | 1 |

$n = 4,$

$X \sim Bin(4, p)$

$L = \cancel{P(X=0)} \cdot \prod_{i=1}^{180} P(X=x)$

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

$L = [{}^4C_0 \, p^4(1-p)^0]^{86} \cdot [{}^4C_1 \, p^3(1-p)]^{75} \cdot [{}^4C_2 \, p^2(1-p)^2]^{16} \cdot [{}^4C_3 \, p(1-p)^3]^2$
$\cdot [{}^4C_4 \, p^0 (1-p)^4]^1$

$= (p^4)^{86} \cdot [4p^3 - 4p^4]^{75} \cdot [6p^2(1-2p+p^2)]^{16} \cdot [4p(1-3p+3p^2-p^3)]^2$
$\cdot (1-4p+6p^2-4p^3+p^4)^1$

$= \cancel{(p^{344}) \cdot (4p^{225} - 4p^{300}}$

$L = [{}^4C_0 \, p^0(1-p)^4]^{86} \cdot [{}^4C_1 \, p^1(1-p)^3]^{75} \cdot [{}^4C_2 \, p^2(1-p)^2]^{16}$
$\cdot [{}^4C_3 \, p^3(1-p)]^2 \cdot [{}^4C_4 \, p^4(1-p)^0]^1$

$L = [(1-p)^4]^{86} \cdot [4p(1-p)^3]^{75} \cdot [6p^2(1-p)^2]^{16} \cdot [4p^3(1-p)]^2 \cdot [p^4]$

Taking log on both sides

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

$\ln L = \ln[(1-p)^{603} \cdot p^{117} \cdot 96]$

$\ln L = \ln[(1-p)^{603}] + \log[p^{117}] + \log 96$

~~Differentiating both sides with respect to 'p'~~

$\cancel{\frac{d(\ln L)}{dp} = \frac{d}{dp}[\ln (1-p}$

$\ln L = 603 \ln(1-p) + 117 \log p + \log 96$

Differentiate wrt 'p'.

$\frac{d(\ln L)}{dp} = \frac{d}{dp}[603 \ln(1-p) + 117 \log p + \log 96]$

$$\frac{d(\ln L)}{dp} = \frac{633}{1-p}(-1) + \frac{117}{p} + 0$$

$$\frac{d(\ln L)}{dp} = \frac{633}{p-1} + \frac{117}{p}$$

$$\frac{d(\ln L)}{dp} = \frac{633p + 117p - 117}{p^2 - p}$$

Let $\frac{d(\ln L)}{dp} = 0$

$$0 = \frac{\cancel{603} \cdot 750p - 117}{p(p-1)}$$

$$750p - 117 = 0$$

$$750p = 117$$

$$p = \frac{117}{750}$$

$$\cancel{p = 0.1625}$$

$$\cancel{\hat{p} = 0.156}$$

$$\hat{p} = 0.156$$

ii) a) Goodness of fit test for $\hat{p} = 0.156$

$H_0: \chi^2_{cal} < \chi^2_{tab}$ vs, $H_1: \chi^2_{cal} > \chi^2_{tab}$

$P(x=0) = [{}^4C_0\, p^0 \cdot (1-p)^4]\, 180 = 180[{}^4\cancel{C_0}\, (1-0.156)^4]$

$= 91.34$

$P(x=1) = [{}^4C_1\, p^1 (1-p)^3]\, 180 = 180[4 \times 0.156 \times (1-0.156)^3]$

$= 67.53$

$P(x=2) = [{}^4C_2\, (0.156)^2 (1-0.156)^2]\, 180$

$= 18.72$

$P(x=3) = 180[{}^4C_3 \times (0.156)^3 \times (1-0.156)]$

$= 2.301$

$P(x=4) = 180[{}^4C_4 \times (0.156)^4 \times (1-0.156)^0]$

$= 0.110$

| x | $O_i$ | $E_i$ |
|---|---|---|
| 0 | 86 | 91.34 |
| 1 | 75 | 67.53 |
| 2 | 16 | 18.72 |
| 3 | 2 | 2.31 |
| 4 | 1 | 0.11(<5) |

| x | $O_i$ | $E_i$ |
|---|---|---|
| 0 | 86 | 91.34 |
| 1 | 75 | 67.53 |
| 2 | 16 | 18.72 |
| 3+ | 3 | 2.42 (<5) |

| x | $O_i$ | $E_i$ | $Z^2 = \frac{(O_i-E_i)^2}{E_i}$ |
|---|---|---|---|
| 0 | 86 | 91.34 | 0.312 |
| 1 | 75 | 67.53 | 0.826 |
| 2+ | 19 | 21.14 | 0.217 |
| | 180 | 180 | $\Sigma z^2 = 1.355$ |

$\nu = n-1-$ no. of clubbing $-$ no. of parameters estimated

$\nu = 5-1-2-1$

$\nu = 1$

$\chi^2_{cal} = 1.355$

$\chi^2_{tab} = \chi^2_1$ (At 5% LOS) $=$ ~~5.024~~ $3.841$

At 5% LOS, $\chi^2_{cal} < \chi^2_{tab}$, hence we do not reject $H_0$.

$\chi^2_{tab} = \chi^2_1 =$ (At 10% LoS) 2.706

At 10% LoS, $\chi^2_{cal} < \chi^2_{tab}$, hence we do not reject $H_0$.

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

i) $\int_0^\infty f(x)\,dx = 1$

$\int_0^\infty \frac{c\beta^3}{(x+\beta)^4} = 1$

$c\beta^3 \int_0^\infty \frac{1}{(x+\beta)^4} = 1$

$c\beta^3 \left[\frac{(x+\beta)^{-3}}{-3}\right]_0^\infty = 1$

$\frac{c\beta^3}{3\beta^3} = 1$

$c = 3$

$E(X) = \int_0^\infty x \cdot f(x)\,dx$

$= 3\beta^3 \int_0^\infty \frac{x}{(x+\beta)^4}\,dx$

$= 3\beta^3 \left[\frac{(x+\beta)^{-2}}{-2} - \frac{\beta(x+\beta)^{-3}}{-3}\right]_0^\infty$

$= 3\beta^3 \left[0 - 0 - \left(\frac{\beta^{-2}}{-2} + \frac{\beta^{-2}}{3}\right)\right]$

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

$= 3\beta^3 \left(\frac{-2+3}{6\beta^2}\right)$

$E(x) = \frac{\beta}{2}$

$$E(x^2) = \int_0^\infty x^2 f(x)\,dx$$

$$= 3\beta^3 \int_0^\infty \frac{x^2}{(x+\beta)^4}\,dx$$

$$= \cancel{3\beta^3}\Big[ \cancel{3\beta^2 \int_0^\beta x^2 (x+\beta)^{-4}\,dx}$$

$$\cancel{= 3\beta^2}$$

ii) $\beta = x_1, x_2, x_3, \dots, n$.

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

$$E(\hat{\beta}) = E(2\bar{x}_n) = \frac{2n}{n} \times \frac{\beta}{2} = \beta$$

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

$$= b^2 \frac{3}{n} \times \frac{\beta^2}{n} + (E(b\bar{x}_n) - \beta)^2$$

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

iv) $\hat{\beta} = 2\bar{x}_n$ is an unbiased estimator fn (ii). $b\bar{x}_n$ $\dot{\omega} = \frac{2}{1+3/n}$

is a biased estimator. It is consistent.

5) $U(a, b)$, $\hat{a} = \bar{x} - \sqrt{3}\,s$, $\bar{x}$ = sample mean, $s$ = sample st. deviation.

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

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

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

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

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

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

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

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

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

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

$s = \frac{1}{\sqrt{12}}$

$\bar{x} = \frac{1}{2}(\bar{x} - \sqrt{3}\, s + b)$

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

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

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

iii) Sample $= 1, 2, 3, 4, 5$

$\bar{x} = \frac{15}{5} = 3.2$

$Var = \frac{(5-1)^2}{12} = \frac{25}{12} = \frac{4}{3}$

st. deviation $= \sqrt{\frac{4}{3}} = 1.54$ $\sqrt{\frac{25}{12}} = \frac{5}{2\sqrt{3}}$

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

$= 3 - (\sqrt{3})\left(\sqrt{\frac{4}{3}}\right) = 3 - 2 = 1$

$\hat{b} = \bar{x} + \sqrt{3}\, s = 3 + \sqrt{3}\left(\frac{\sqrt{4}}{\sqrt{3}}\right) = 3 + 2 = 5$

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

$= 3.2 - \sqrt{3}\left(\frac{5}{2\sqrt{3}}\right)$

$= 0.7$

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

$= 3.2 + \sqrt{3}\left(\frac{5}{2\sqrt{3}}\right)$

$\hat{b} = 5.7$

Here, $\hat{a} = 0.7 \neq a = 1$

& $\hat{b} = 5.7 \neq b = 6$

Hence, the Method of Moments is not completely reliable here.

6) $H_0 : \cancel{p = 0.1}$ vs $\cancel{H_1 : p > 0.1}$

$P(X \geq 3)$

(v) $a = 0,\ x_i = x_1, x_2, \ldots, x_n$

$X \sim U(0, b)$

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

$L = \prod_{i=1}^{n} f(x_i) = \frac{1}{b^n} = b^{-n}$

Taking log on both sides

$\ln L = -n \ln b$

Differentiate both wrt 'b'.

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

$\frac{-n}{b} = 0$

$b = \infty$.

6) $H_0 : p = 0.1$
$H_1 : p > 0.1$

i) $P(\text{Type I error}) = P[\text{Reject } H_0 \mid H_0 \text{ is True}]$
$= P[X \geq 3 \mid p = 0.1] + P[X = 0 \mid p = 0.1] \cdot P[Y \geq 5 \mid p = 0.1]$
$+ P[X = 1 \mid p = 0.1] \cdot P[Y \geq 4 \mid p = 0.1] + P[X = 2 \mid p = 0.1]$
$\cdot P[Y \geq 3 \mid p = 0.1]$
$= \cancel{0.085237 + (0.282433)}$
$= 0.14727$

$P(\text{Type I error}) = 15\%$.

ii) $P(X \geq 3 \mid P = 0.3) + P(X = 0 \mid P = 0.3) = 0.91253 = 91\%$.

iii) $P(\text{Type II error}) = 1 - 0.91253 = 0.08747 = 9\%$.

iv) $X \sim Bin(120, 1/3)$, $\hat{p} = \frac{40}{120} = \frac{1}{3}$, $\alpha = 10\%$

$$P^{CI}_{p} = \hat{p} \pm z_\alpha \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$= \frac{1}{3} - 1.2186 \sqrt{\frac{1/3\,(1-1/3)}{120}}$$

$$= \frac{1}{3} - 1.2186 \sqrt{\frac{1/3\,(2/3)}{120}}$$

$$= \frac{1}{3} - 1.2186 \sqrt{\frac{2}{9 \times 120}}$$

$$= 0.281$$

Lower bound $= 0.281$

v) $H_0: p = 0.278$ , $H_1: p > 0.278$

$\alpha = 10\%$

~~T.S. $\frac{(x \pm 1/2)/n - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$~~

~~$\frac{x - np_0}{\sqrt{np_0 q_0}} \sim N(0,1) = 1.2$~~

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

$x = 40 - 1/2$ (continuity correction)

$x = 39.5$

$$\frac{39.5 - 120(0.278)}{\sqrt{120(0.278)(1-0.278)}} = \frac{6.14}{4.908} = 1.2511$$

$Z_{cal} < Z_{tab}$.

Hence, we donot reject $H_0$.

7)

| City | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Public | 24 | 45 | 29 | 33 | 20 | 40 | 26.5 | 25 | 27.5 |
| Private | 30 | 54.5 | 30 | 40 | 28.5 | 36 | 30.5 | 30.5 | 35.5 |

i) ~~Average~~ Public -

$$\bar{x}_1 = \frac{\Sigma x_1}{n} = \frac{270}{9} = 30$$

$$S_1^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n-1} = \frac{\Sigma x_1^2 - n\bar{x}_1^2}{n-1} = \frac{8614.5 - 9(30)^2}{8}$$

$$S_1^2 = \frac{8614.5 - 8100}{8} = \frac{514.5}{8} = 64.3125$$

Private -

$$\bar{X}_2 = \frac{\sum \bar{X}_2}{n} = \frac{315.5}{9} = 35.056$$

$$S_2^2 = \frac{\sum (x_2 - \bar{X}_2)^2}{n-1} = \frac{\sum x_2^2 - n\bar{x}_2^2}{n-1} = \frac{11599.25 - 9(35.056)^2}{8}$$

$$S_2^2 = \frac{11599.25 - 11060.31}{8} = 67.3675$$

ii) The primary condition for testing equal means is $\sigma_1^2 = \sigma_2^2$.

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

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

$$F_{cal} = \frac{S_1^2}{S_2^2} = \frac{64.3125}{67.3675} = 0.9547$$

$F_{tab} = F_{8,8}$ (At. 5% LOS) $= ~~4.899~~$
$= (0.2154, 4.2322)$

So, $F_{tab} < F_{cal} < F_{tab}$, so, we do not reject $H_0$.
Hence, the primary condition for testing equal means is true.

iii) $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 < \mu_2$.

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

$$S_p^2 = \frac{S_1^2(n_1-1) + S_2^2(n_2-1)}{(n_1+n_2-2)} = \frac{64.3125(8) + 67.3675(8)}{16}$$

$= 65.84$

$S_p = \sqrt{65.84} = 8.114$

$$\frac{(35.056-30)-(0)}{8.114\sqrt{\frac{1}{9}+\frac{1}{9}}} = t_{cal}$$

$$t_{cal} = \frac{5.056}{3.8248}$$

$t_{cal} = 1.322$

$t_{tab} = t_{16,5\%} = 1.746$

$t_{cal} < t_{tab}$.

Hence, we do not reject $H_0$.

So, we conclude that $\mu_1 \approx \mu_2$.

8) i) If we have a random sample $\underline{x} = (x_1, x_2, x_3, \ldots, x_n)$ from a distribution with an unknown parameter $\theta$ and $g(\underline{x})$ is an estimator of $\theta$, it seems desirable that $E[g(\underline{x})] = \theta$. This is the property of unbiasedness. So, $\hat{\theta} = \theta$

ii) If an estimator is biased, its bias is given by $E[g(\underline{x})] - \theta$, i.e., it is a measure of the difference between the expected value of the estimator and the parameter being tested.

iii) The mean square error of an estimator $\hat{\theta}$, for $\theta$ is defined by: $MSE(\hat{\theta}) = E[(\hat{\theta}-\theta)^2]$.

iv) Since, $MSE(\hat{\theta}) > MSE(\tilde{\theta})$, $\tilde{\theta}$ is more efficient because an estimator with a lower MSE is said to be more efficient.

v) Either of the two estimators would be termed as consistent when their MSE $\rightarrow 0$ as $n \rightarrow \infty$.

vi) Any two methods of estimating $\theta$ are :

a) Method of Moments – Its basic principle is to equate population moments to corresponding sample moments and solve for the parameters.

b) Maximum ~~likelihood~~ Likelihood Estimation – It is widely regarded as the best general method of finding estimators.

9) i) $t_k$ can be defined as $\dfrac{N(0,1)}{\sqrt{\chi^2_k / k}}$, where, $k$ denotes the degrees of freedom and $N(0,1)$ and $\chi^2_k$ ~~are~~ variables are independent.

ii) For $k > 2$,

Mean of $t_k = 0$

Variance of $t_k = \dfrac{k}{k-2}$

iii) $n = 10$, $\bar{x} = 50$, $s^2 = 48.667$, $s = 6.976$

a) $\alpha = 1\%$,

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

$$CI = \bar{x} \pm (t_{\alpha/2, n-1}) \cdot s/\sqrt{n}$$

$$= 50 \pm (t_{0.5\%, 9})(6.976/\sqrt{10})$$

$$= 50 \pm (3.250)(2.206)$$

$$CI = (42.8313, 57.1695)$$

Scanned with CamScanner

~~f) $\mu = 0$, $\sigma^2 = k/(k-2) = 9/(9-2) = 1.286$~~

b) $X \sim N(50, 48.667)$

$$CI = \bar{x} \pm Z_{\alpha/2} \, \sigma/\sqrt{n}$$
$$= 50 \pm (Z_{0.57})(6.976/\sqrt{10})$$
$$= 50 \pm (2.5758)(2.206)$$
$$CI = (44.318, 55.682)$$

10) $n = 1000$, $X \sim Poi(3)$, $X \sim N(3000, 3000)$

i) Type I error = $P[H_0 \text{ is rejected} \mid H_0 \text{ is true}]$.

$$= P[X > 3100 \mid \lambda = 3].$$
$$= P\left[Z > \frac{x - n\lambda}{\sqrt{n\lambda}}\right]$$
$$= P\left[Z > \frac{3100 - 3000}{\sqrt{3000}}\right]$$
$$= P[Z > 1.8257]$$
$$= 1 - P[Z < 1.8257]$$
$$= 3.39\%$$

ii) Type II error = $P[\text{do not reject } H_0 \mid H_0 \text{ is false}]$

$$= P[X < 3100 \mid \lambda = n\lambda]$$
$$= P\left[Z < \frac{3100 - n\lambda}{n\lambda}\right]$$

iii) Power = $P[\text{reject } H_0 \mid H_0 \text{ is false}]$

$$= P[X > 3100 \mid \lambda = n\lambda]$$
$$= 1 - P\left[\frac{3100 - n\lambda}{\sqrt{n\lambda}}\right]$$

iv) $n = 2900$, $\alpha = 1\%$

$CI = \lambda \pm Z_{\alpha/2} \sqrt{\lambda / n}$

$= 3 \pm (2.5758)\sqrt{3/2900}$

$=$

$CI = (2.9172, 3.0828)$

11) $n = 12$, $\Sigma x = 213,200$, $\Sigma x^2 = 4,919,860,000$

$\bar{x} = 17,767$, $s = 10,144$

Tem. emp. $= 11,000 + 48,000 = 59,000$

Perm. emp. $= 27,500 + 31,500 = 59,000$

New $\Sigma x = 213,200 - 59,000 + 59,000 = 213,200$

There will be no change in the mean.

New $\Sigma x^2 = 4,919,860,000 - (11,000^2) - (48,000)^2 + (27,500^2) + (31,500^2)$

$= 4,919,860,000 - 121,000000 - 2,304,000,000 + 756,250,000 + 992,250,000$

New $\Sigma x^2 = 4,243,360000$

New sample ~~mean~~ variance $= \dfrac{4,243,360,000 - 12(17,767)^2}{11}$

$= 41,396,775.64$

New sample st. deviation $= \sqrt{41,396,775.64}$

$= 6434.033$

The new mean is same but st. dev. is lower.

12) $X_1, X_2, \ldots X_n$ , $X \sim U(0, \theta)$ , $a=0$ , $b=\theta$.

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

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

Taking log on both sides

$$\ln(L) = \frac{n}{\theta}$$

Differentiating both sides wrt '$\theta$'

$$\frac{d(\ln L)}{d\theta} = \frac{-n}{\theta} = 0$$

So, $\theta$ = infinity.

$$CRLB = \frac{1}{E\{[\frac{\partial}{\partial\theta} \log L(\theta, \underline{x})]^2\}}$$

If $\theta$ is infinite, we will ~~not~~ not be able to compute CRLB. Hence, it does not apply here.

ii) $P[\text{Max } x_i < y] = \prod_{i=1}^{n} [\int_0^y dx/\theta] = \cancel{(y/\theta)^n}\ \frac{n}{\theta^n} \cdot y^{n-1}$

$$E[y^k] = \int y^k \cdot \frac{n}{\theta^n} \cdot y^{n-1}\, dy$$

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

$$= \frac{n}{n+k} \left( \frac{\theta^{n+k} - 0}{\theta^n} \right)$$

$$= \frac{n\theta^k}{n+k}$$

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

$$= E[cy - \theta]$$

$$= c\,E[y] - \theta$$

$$= c \cdot \frac{n\theta}{n+1} - \theta$$

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

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

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

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

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

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

$k = 2,\ k = 1$

$$= c^2\left(\frac{n\theta^k}{n+k}\right) - 2c\theta\left(\frac{n\theta^k}{n+k}\right) + \theta^2$$

$$= c^2\left(\frac{n\theta^2}{n+2}\right) - 2c\theta\left(\frac{n\theta}{n+1}\right) + \theta^2$$

$$MSE = c^2\left(\frac{n\theta^2}{n+2}\right) - 2c\left(\frac{n\theta^2}{n+1}\right) + \theta^2.$$

iv) For unbiased estimator, $\text{Bias}(\hat{\theta}(c)) = 0$

$$E(\hat{\theta}c) - \theta = 0$$

$$E(cy) - \theta = 0$$

$$c\,E(y) - \theta = 0$$

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

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

$$\frac{cn}{n+1} - 1 = 0$$

$$c_n = 1(n+1)$$

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

v) To minimize MSE $\frac{d}{dc} MSE = 0$

$$\frac{d}{dc}\left[\frac{c^2 n\theta^2}{n+2} - \frac{2cn\theta^2}{n+1} + \theta^2\right] = 0$$

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

$$2n\theta^2\left(\frac{c}{n+2} - \frac{1}{n+1}\right) = 0$$

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

$$\therefore c/n+2 = \frac{1}{n+1}$$

$$\therefore c_m = \frac{n+2}{n+1}$$