

Probability and Statistics

Assignment

$$\textcircled{1} \quad X \sim \text{Poi}(\theta) \Rightarrow \cancel{X \sim N(\theta, \theta)}$$

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

$$P(X \leq 200) = 0.025$$

$$P(X < 200.5) = 0.025$$

$$P\left(Z < \frac{200.5 - \theta}{\sqrt{\theta}}\right) = 0.025$$

$$-1.96 = \frac{200.5 - \theta}{\sqrt{\theta}}$$

$$\Rightarrow \theta - 1.96\sqrt{\theta} - 200.5 = 0$$

$$\Rightarrow \underline{\theta = 230.24}$$

$$P(X \geq 200) = 0.025$$

$$P(X > 199.5) = 0.025$$

$$P\left(Z > \frac{199.5 - \theta}{\sqrt{\theta}}\right) = 0.025$$

$$1.96 = \frac{199.5 - \theta}{\sqrt{\theta}}$$

$$\Rightarrow \theta + 1.96\sqrt{\theta} - 199.5 = 0$$

$$\underline{\theta = 173.67}$$

$\therefore$ CI for θ at 95% $\Rightarrow (173.67, 230.64)$

$$\textcircled{2} \quad \bar{X} = \sum_{i=1}^n X_i \left(\frac{1}{n}\right)$$

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

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

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

$$= \frac{1}{n} \sum E(X_i)$$

$$= \frac{1}{n^2} (\sum \text{Var}(X_i))$$

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

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

$$= \mu$$

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

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

$$\begin{aligned}
 \text{ii. } E(S^2) &= E\left(\frac{1}{n-1} \sum (X_i - \bar{X})^2\right) \dots \left(\sum (X_i - \bar{X})^2 = \sum (X_i^2 + \bar{X}^2 - 2X_i\bar{X})\right) \\
 &= E\left(\frac{1}{n-1} (\sum X_i^2 - n\bar{X}^2)\right) \\
 &= \frac{1}{n-1} (\sum E(X_i^2) - nE(\bar{X}^2)) \\
 &= \frac{1}{n-1} (n\sigma^2 + n\mu^2 - n\mu^2 - \sigma^2) \\
 &= \frac{1}{n-1} (\sigma^2)(n-1) = \sigma^2
 \end{aligned}$$

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

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

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

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

$$\text{iv. } n=10 \quad \sigma^2=100 \quad E(S^2) = \sigma^2 = 100$$

$$\Rightarrow P[50 \leq S^2 \leq 100] = P\left[\frac{9(50)}{100} \leq \chi^2_9 \leq \frac{100(9)}{100}\right]$$

$$\begin{aligned}
 &= P[4.5 \leq \chi^2_9 \leq 13.5] \Rightarrow P(\chi^2_9 \leq 13.5) - P(\chi^2_9 \leq 4.5) \\
 &\Rightarrow 0.8587 - 0.1245 \\
 &\Rightarrow 0.7342
 \end{aligned}$$

③ $n=4$

i.
$$L(p) = \left[{}^4C_0 p^0 (1-p)^4 \right]^{86} \times \left[{}^4C_1 p^1 (1-p)^3 \right]^{75} \times \left[{}^4C_2 p^2 (1-p)^2 \right]^{16} \\ \times \left[{}^4C_3 p^3 (1-p) \right]^2 \times \left[{}^4C_4 p^4 (1-p)^0 \right]^1$$

$$= \text{const.} \times (1-p)^{344} \times p^{75} \times (1-p)^{225} \times p^{32} \times (1-p)^{32} \\ \times p^6 \times (1-p)^2 \times p^4$$

$$= \text{const.} \times (1-p)^{603} \times p^{117}$$

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

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

$$603p = 117 - 117p$$

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

ii.

X	0	1	2	3	4
O _i	86	75	16	2	1
E _i	88.55	68.729	20.003	2.587	0.1255

$\Rightarrow$

X	0	1	≥ 2
O _i	86	75	19
E _i	88.55	68.729	22.7155

$$\sum \frac{(O_i - E_i)^2}{E_i} = \chi^2_{df} \quad df = 3 - 1 - 1 = 1$$

$$\chi^2_1 = 1.2533, \quad \chi^2_1 \text{ at } 5\% (1.0.5) = 3.841$$

As $\chi^2_{cal} < \chi^2_{tab}$, Binomial model is not a good fit.

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

$$i. \Rightarrow \int_0^{\infty} f(x) dx = 1 \Rightarrow \int_0^{\infty} \frac{c\beta^3}{(x+\beta)^4} dx = 1$$

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

$$1 = \frac{c\beta^3}{-3} \left[0 - \frac{1}{\beta^3} \right] = \frac{c}{3}$$

$$\Rightarrow \boxed{c=3}$$

$\therefore f(x) = \frac{3\beta^3}{(x+\beta)^4}$ is pareto distribution with $\alpha=3$.

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

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

$$ii. \quad \bar{x} = \mu$$

$$\Rightarrow \frac{\sum x_i}{n} = \frac{\beta}{2} \Rightarrow \boxed{\hat{\beta} = 2\bar{x}_n}$$

$$\begin{aligned} E(\hat{\beta}) - \beta &= E(2\bar{x}_n) - \beta \quad \therefore \hat{\beta} \text{ is unbiased estimator of } \beta \\ &= 2E(\bar{x}_n) - \beta \\ &= 2\left(\frac{\beta}{2}\right) - \beta = 0 \end{aligned}$$

$$MSE(\hat{\beta}) = \text{Var}(\hat{\beta}) + [\text{Bias}(\hat{\beta})]^2$$

$$\dots [\text{Bias}(\hat{\beta}) = 0]$$

$$\text{Var}(\hat{\beta}) = \text{Var}(2\bar{x}_n)$$

$$= 4 \text{Var}(\bar{x}_n) = 4 \left(\text{Var}\left(\frac{\sum x}{n}\right) \right)$$

$$= \frac{4}{n^2} \sum \text{Var}(x) = \frac{4}{n^2} (n) \left(\frac{3}{4} \beta^2 \right) = \frac{3\beta^2}{n}$$

$$MSE(\hat{\beta}) = \frac{3\beta^2}{n} + 0 = \frac{3\beta^2}{n}$$

As $n \rightarrow \infty$, $MSE \rightarrow 0$, $\therefore$ is consistent.

$$\begin{aligned} \text{iii. Bias}[b\bar{x}_n] &= E[b\bar{x}_n] - \beta \\ &= b E(\bar{x}_n) - \beta \\ &= b\left(\frac{\beta}{2}\right) - \beta = \beta\left(\frac{b}{2} - 1\right) \end{aligned}$$

$$\begin{aligned} \text{Var}(b\bar{x}_n) &= b^2 \text{Var}(\bar{x}_n) \\ &= \frac{b^2}{n^2} \left[n \left(\frac{3\beta^2}{4} \right) \right] = \frac{3b^2\beta^2}{4n} \end{aligned}$$

$$MSE(b\bar{x}_n) = \frac{3b^2\beta^2}{4n} + \beta^2 \left(\frac{b}{2} - 1 \right)^2 \Rightarrow \frac{6b\beta^3}{4n} + \beta^2 \left(\frac{b}{2} - 1 \right) \left(\frac{1}{2} \right)$$

$$\frac{d}{db} MSE(b\bar{x}_n) = \frac{6b\beta^3}{4n} - \beta^2 \left(\frac{b}{2} - 1 \right) \left(\frac{1}{2} \right)$$

$$0 = \frac{3b\beta^3}{2n} - \beta^2 \left(\frac{b}{2} - 1 \right)$$

$$0 = \frac{3b}{2n} - \frac{b}{2} - 1 \Rightarrow 1 = b \left[\frac{3}{2n} - \frac{1}{2} \right] \Rightarrow b = \frac{1}{\left(\frac{3}{2n} - \frac{1}{2} \right)}$$

$$0 = \frac{6b\beta}{4n} + \beta \left(\frac{b-1}{2} \right)$$

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

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

iv. We know that $\hat{\beta} = 2\bar{x}_n$ is an unbiased estimator of β as seen in (ii).

$\Rightarrow b\bar{x}_n$ is negatively biased for $b \leq 2$, $n > 1$ when $b = \frac{2}{(1+3/n)}$. Also, $n \rightarrow \infty$, $b \rightarrow 2$. $\therefore$ estimator is consistent.

⑤ $X \sim U(a, b) \Rightarrow E(X) = \bar{x} = \frac{a+b}{2} = \frac{b+a}{2}$

i. $2\bar{x} = b+a \quad \text{--- (1)}$

$$\text{Var}(X) = \frac{(b-a)^2}{12} \Rightarrow 2\sqrt{3}s = b-a$$

$$b = 2\sqrt{3}s + a \quad \text{--- (2)}$$

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

Substituting (2) in (1)

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

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

ii. Substituting (3) in (1)

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

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

iii. Let sample of five observations be 1, 2, 3, 4, 50

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

$$s = \sqrt{291.272}$$

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

$$= 12 - z\sqrt{3}(21.272)$$

$$\hat{a} = -24.844$$

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

$$= 12 + \sqrt{3}(21.272)$$

$$\hat{b} = 48.84$$

$\therefore$ Potential weakness of MOM to estimate a and b for $U(a, b)$ is invalid value, as we can see that there's a value (50) in sample greater than $\hat{b}$ which contradicts $U(a, b)$.

iv. $f(x) = \frac{1}{b} \Rightarrow L = \prod_{i=1}^n \frac{1}{b} = \frac{1}{b^n}$

$$0 \leq x_1 < x_2 \dots x_n \leq b$$

$\therefore$ For L to be maximum, b^n should be minimum, for b to be minimum,

$$\boxed{b = x_{\max}}$$

$\therefore$ MLE is $x_{\max}$

⑥ i. $H_0: p = 0.1$

$H_a: p > 0.1$

i. $P[\text{Type I}] = P[X \geq 3] + P[X=2]P[X' \geq 3] + P[X=1]P[X' \geq 4] + P[X=0]P[X' \geq 5]$

$$= (1 - 0.8891) + (0.2824)(0.0043) + (0.6590 - 0.2824)(1 - 0.9744)$$

$$+ (0.8891 - 0.6590)(1 - 0.8891) = 0.14727 \approx 15\%$$

$$p = 0.3$$

$$\text{ii. } = P[X \geq 3] + P[X=2]P[X' \geq 3] + P[X=1]P[X \geq 4] + P[X=0]P[X \geq 5]$$

$$= 0.7472 + (0.0138)(0.2763) + (0.085 - 0.0138)(0.5075) + (0.2528 - 0.085)(0.7472)$$

$$= 0.91253 = \underline{91.253\%}$$

$$\text{iii. } P[\text{Type II error}] = P[\text{Accept } H_0 \mid H_0 \text{ is false}]$$

$$= 1 - 0.91253 = 0.08747$$

$$= \underline{8.747\%}$$

$$\text{iv. } \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}\hat{q}}{n}}} \sim N(0,1)$$

$$\hat{p} = \frac{40}{120} = \frac{1}{3} \quad \hat{q} = \frac{2}{3}$$

$$Z_{10\%} = 1.2816 \Rightarrow \hat{p} - 1.2816 \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= \frac{1}{3} - 1.2816 \sqrt{\frac{(\frac{1}{3})(\frac{2}{3})}{120}}$$

$$= 0.2782$$

$$\text{v. } H_0: p = 0.278 \quad H_a: p > 0.278$$

$$Z = \frac{\bar{x} - np}{\sqrt{npq}} = \frac{39.5 - 120(0.278)}{\sqrt{120(0.278)(1-0.278)}} = 1.2511$$

$Z_{\text{tab}} = 1.2816$, As $Z_{\text{cal}} < Z_{\text{tab}}$, we have insufficient evidence to reject H_0 at 10%.

vi. The C.I states that there is 10% chance of $p \leq 0.2782$.
The hypothesis test states that p could be less than or equal to 0.278 with probability more than 10% (approximately 10.5%).

vii. $H_0: \mu_1 = \mu_2$ $H_a: \mu_1 \neq \mu_2$

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

$$\bar{x}_1 = 65 \quad \bar{x}_2 = 70 \quad S_1^2 = 54 \quad S_2^2 = 70 \quad n_1 = 12 \quad n_2 = 15$$

$$S_p^2 = 4027.04 \Rightarrow S_p = 63.459$$

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

$$t_{\text{cal}} = -0.2034 \quad t_{\text{tab}} = \pm 2.060$$

As t_{cal} lies between t_{tab} , we have insufficient evidence to reject H_0 at 5%.

⑦ i. $\bar{x} = 30$
(Public)

$$S_x^2 = 64.3125$$

$$\bar{y} = 35.055$$

$$S_y^2 = 67.4027$$

ii. The condition that needs to be true for testing equal mean is the population variances of both samples are equal.

we can look at the sample variances of both data being close enough, to assume that their populations variances are same.

iii. $H_0: \mu_y = \mu_x$ $H_a: \mu_y > \mu_x$

$$Z_c = \frac{\bar{y} - \bar{x} - (\mu_y - \mu_x)}{S_p \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \quad \dots \quad \left(S_p = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 1}} \right)$$

$$S_p = 8.1153$$

$$Z_c = \frac{35.055 - 30}{8.1153 \sqrt{\frac{2}{9}}} = 1.321$$

$$Z_{tab} = 1.746$$

∴ As $Z_{cal} < Z_{tab}$ i.e. $\left(t_{0.05, 16} \right)$ we have insufficient evidence to reject H_0 .

(8) i. Parameter = 0 Estimator = $\hat{\theta}$

Unbiased estimator $\Rightarrow E(\hat{\theta}) = 0$

ii. Bias($\hat{\theta}$) = $E(\hat{\theta}) - 0$

iii. $MSE(\hat{\theta}) = Var(\hat{\theta}) + [Bias(\hat{\theta})]^2 = E(\hat{\theta} - 0)^2$

iv. $\text{Bias}(\tilde{\theta}) = E(\tilde{\theta}) - \theta = +ve$

$$\therefore E(\tilde{\theta}) > \theta$$

$MSE(\hat{\theta}) > MSE(\tilde{\theta})$: $\tilde{\theta}$ is better, efficient parameter compared to $\hat{\theta}$ even though $\tilde{\theta}$ has got +ve bias. For a parameter to be efficient, MSE must be low.

v. If by increasing the sample size, i.e. 'n'.

$MSE \rightarrow 0$, If this is true, the parameter estimator can be termed as consistent.

vi. Two methods of estimating $\theta \Rightarrow$

- Method of Moments (MOM)
- Maximum Likelihood Estimate (MLE).

① i. $t_k = \frac{\bar{x} - \mu}{S/\sqrt{n}}$, $\bar{x}$ = Sample Mean, k = degree of freedom
 μ = Population Mean, χ^2_k and $N(0,1)$
 $= \frac{N(0,1)}{\sqrt{\chi^2_k/k}}$, S = Sample Variance, are independent
 n = Sample Size

ii. $E(t_k) = 0$, $k > 1$

$$\text{Var}(t_k) = \frac{k}{k-2}, \quad k > 2$$

iii. $n=10$, $\bar{x} = 50$, $S^2 = 48.667$

a. $t_k = \frac{\bar{x} - \mu}{S/\sqrt{n}} = \frac{50 - \mu}{\sqrt{48.667/10}}$

$$-3.25 \leq \frac{50 - \mu}{\sqrt{48.667/10}} \leq 3.25$$

$$50 - 3.25 \sqrt{\frac{48.667}{10}} \leq \mu \leq 50 + 3.25 \sqrt{\frac{48.667}{10}}$$

$\therefore$ Confidence Interval for population mean
at 99% CI $\Rightarrow (42.8303, 57.1696)$

$$b. \quad X \sim N\left(0, \frac{9}{7}\right) \Rightarrow \bar{X} \sim N\left(0, \frac{9}{70}\right)$$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \Rightarrow \frac{50 - \mu}{\sqrt{9/70}}$$

$$\Rightarrow \frac{\bar{X} - 0}{\sqrt{9/70}} \sim N(0, 1)$$

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

$$\Rightarrow \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim N\left(0, \frac{9}{7}\right)$$

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

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

$$\Rightarrow \frac{\bar{X} - \mu}{\sqrt{9S^2/7n}} \sim N(0, 1)$$

$$-2.58 < \frac{\bar{x} - \mu}{\sqrt{98^2/n}} < 2.58$$

$$\Rightarrow 50 \pm 2.58 \sqrt{\frac{9(48.667)}{7(10)}} = \text{Confidence Interval at 99\%}$$

$$\Rightarrow (43.54, 56.45)$$

$$\textcircled{b} \quad X \sim \text{Poi}(\lambda) \Rightarrow \Sigma X \sim \text{Poi}(n\lambda) \quad n=1000$$

$$\Rightarrow \frac{\Sigma X - n\lambda}{\sqrt{n\lambda}} \approx N(0,1)$$

$$H_0: \lambda = 3$$

$$H_a: \lambda \neq 3$$

$$i. \quad P[\text{type I}] = P[\text{Rejecting } H_0 \mid H_0 \text{ is true}]$$

$$= P[\Sigma X > 3100 \mid \lambda = 3]$$

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

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

$$= P[Z > 1.8257] = 0.03394$$

$$= 3.394\%$$

$$ii. \quad P[\text{type 2}] = P[\text{Accepting } H_0 \mid H_0 \text{ is false}]$$

$$= P[\Sigma X < 3100 \mid \lambda \neq 3]$$

$$= P\left[Z \leq \frac{3100 - n\lambda}{\sqrt{n\lambda}} \mid \lambda \neq 3\right]$$

$$= P\left[Z \leq \frac{3100 - 1000(\lambda)}{\sqrt{1000(\lambda)}} \mid \lambda \neq 3\right]$$

$$\begin{aligned}
 \text{iii. Power of test} &= 1 - P[\text{type II}] \\
 &= 1 - P\left[z < \frac{3100 - 1000\lambda}{\sqrt{1000\lambda}} \mid \lambda \neq 3\right] \\
 &= P\left[z > \frac{3100 - 1000\lambda}{\sqrt{1000\lambda}} \mid \lambda \neq 3\right]
 \end{aligned}$$

$$\text{iv. } \hat{\lambda} \pm 2.5758 \sqrt{\frac{\hat{\lambda}}{n}} \quad \dots \quad \hat{\lambda} = \frac{x}{n} = \frac{2900}{1000} = 2.9$$

$$\Rightarrow 2.9 \pm 2.5758 \sqrt{\frac{2.9}{1000}}$$

$$\Rightarrow 2.9 \pm 0.1387 \Rightarrow 99\% \text{ CI} = (2.7613, 3.0387)$$

ii) Initial data:

$$n = 12$$

$$\sum x = 213200 \quad \bar{x} = 17767$$

$$\sum x^2 = 4919860000 \quad s = 10144$$

After replacing two temporary employees salary:

$$n = 12$$

$$\begin{aligned}
 \sum x &= 213200 - 11000 - 48000 + 27500 + 31500 \\
 &= 213200
 \end{aligned}$$

$$\bar{x} = \frac{213200}{12} = \underline{17767}$$

$$\begin{aligned}
 \sum x^2 &= 4919860000 - 11000^2 - 48000^2 + 27500^2 + 31500^2 \\
 &= 4233360000
 \end{aligned}$$

$$s = \sqrt{\frac{\sum x^2 - 12(\bar{x})^2}{n-1}} = 6363$$
 sample standard deviation changes as there's less difference between two new values.

② $X_i \sim U(0, \theta)$

$$f(x) = \frac{1}{\theta - 0}, \quad \theta > 0$$

i. $L(\theta) = \begin{cases} \frac{1}{\theta^n}, & \theta > 0 \\ 0, & \text{otherwise} \end{cases}$

$0 \leq x_1, x_2, \dots, x_n \leq \theta$, To maximise $\frac{1}{\theta^n}$, θ^n needs to be minimized.

$\theta \geq X_{\max} \therefore \underline{L = X_{\max}}$

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

$$\frac{\partial^2 \ln L(\theta)}{\partial \theta^2} = \frac{n}{\theta^2} > 0$$

As double derivative of $L(\theta)$ is always greater than 0, CRLB for the variance for the unbiased estimator of θ is not possible.

$\hat{\theta}(c) = cY, \quad Y = X_{\max}$

ii.
$$\begin{aligned} P[Y \leq y] &= P[X_{\max} \leq y] \\ &= P[X_1, X_2, \dots, X_n \leq y] \\ &= P[X_1 \leq y] P[X_2 \leq y] \dots P[X_n \leq y] \\ &= P[X_i \leq y]^n \\ &= \left(\int_0^y \frac{1}{\theta} dx \right)^n = \left(\left[\frac{x}{\theta} \right]_0^y \right)^n = \left(\frac{y}{\theta} \right)^n = F_Y(y) \end{aligned}$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \left(\frac{y}{\theta} \right)^n = \frac{n y^{n-1}}{\theta^n}, \quad 0 < y < \theta$$

$$\begin{aligned}
 E(Y^k) &= \int_0^{\theta} y^k \frac{n}{\theta^n} y^{n-1} dy \\
 &= \frac{n}{\theta^n} \int_0^{\theta} y^{n+k-1} dy \\
 &= \frac{n}{\theta^n} \left[\frac{y^{n+k}}{n+k} \right]_0^{\theta} = \frac{n}{\theta^n(n+k)} \left[\theta^{n+k} \right] \\
 E[Y^k] &= \frac{n\theta^k}{n+k}, \quad k > 0
 \end{aligned}$$

$$\begin{aligned}
 \text{iii. Bias}[\hat{\theta}(c)] &= E(\hat{\theta}(c)) - \theta \\
 &= E(Y) - \theta \\
 &= c \left[\frac{n\theta^n}{n+1} \right] - \theta = \frac{n\theta^n - n\theta^n}{n+1} \\
 &= \frac{cn\theta^n}{n+1} - \theta
 \end{aligned}$$

$$\text{Bias}[\hat{\theta}(c)] = \theta \left[\frac{cn}{n+1} - 1 \right]$$

$$\begin{aligned}
 \text{Var}(\hat{\theta}(c)) &= \text{Var}(cY) = E(Y^2) - [E(Y)]^2 \\
 &= \left(\frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} \right) c^2
 \end{aligned}$$

$$\text{MSE}(\hat{\theta}(c)) = \text{Var}(\hat{\theta}(c)) + (\text{Bias}[\hat{\theta}(c)])^2$$

$$\begin{aligned}
 &= c^2 \left(\frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} \right) + \theta^2 \left[\frac{cn}{n+1} - 1 \right]^2 \\
 &= \frac{c^2 n \theta^2}{n+2} - \frac{c^2 n^2 \theta^2}{(n+1)^2} + \frac{c^2 n^2 \theta^2}{(n+1)^2} + \theta^2 - \frac{2cn\theta^2}{n+1} = \frac{c^2 n \theta^2}{n+2} - \frac{2cn\theta^2}{n+1} + \theta^2
 \end{aligned}$$

iv. $\text{Bias}[\hat{\theta}(c)] = \theta \left[\frac{cn}{n+1} - 1 \right] = 0$

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

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

~~$$cn = n+1 \Rightarrow c(n-1) = 1$$~~

~~$$c = \frac{1}{n-1}$$~~

v. $\text{MSE}[\hat{\theta}(c)] = \frac{c^2 n \theta^2}{n+2} - \frac{2cn\theta^2}{n+1} + \theta^2 =$

$$\frac{d}{dc} \text{MSE}(\hat{\theta}(c)) = \frac{2cn\theta^2}{n+2} - \frac{2n\theta^2}{n-1} = 0$$

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

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

vi. We will prefer $\hat{\theta}(c_m)$ for estimating θ , as low MSE is preferred over high MSE.

As $n \rightarrow \infty$, $\text{MSE}[\hat{\theta}(c_m)] \rightarrow 0$