1. Consider a random sample $\{X_1, \ldots, X_n\}$ of size n>1 with an unknown mean $\mu \in (-\infty, \infty)$ and unknown variance $\sigma^2 \in (0, \infty)$. Show that S_n^2 has a smaller MSE than S_{n-1}^2 .

$$
\begin{aligned}
& MSE(S_{n-1}^2) - MSE(S_n^2) \\
&= Var(S_{n-1}^2) + Bias(S_{n-1}^2) - Var(\frac{n-1}{n} S_n^2) - Bias(\frac{n-1}{n} S_n^2) \\
&= \frac{1}{n}(\mu_4 - \frac{n-3}{n-1}\sigma^4) + 0 - (\frac{n-1}{n})^2 \frac{1}{n}(\mu_4 - \frac{n-3}{n-1}\sigma^4) - (\frac{\sigma^2}{n})^2 \\
&= \frac{1}{n}\mu_4 - \frac{n-3}{n(n-1)}\sigma^4 - \frac{(n-1)^2}{n^3}\mu_4 + \frac{(n-1)(n-3)}{n^3}\sigma^4 - \frac{\sigma^4}{n^2} \\
&= \frac{n^2 - n^2 + 2n - 1}{n^3}\mu_4 + \frac{-n^2(n-3) + (n-1)^2(n-3) - n(n-1)}{n^3(n-1)}\sigma^4 \\
&= \frac{2n-1}{n^3}\mu_4 + \frac{n(8-3n)-3}{(n-1)n^3}\sigma^4
\end{aligned}
$$

where for big n,

$$
\begin{aligned}
\lim_{n\to\infty} \frac{MSE(S_{n-1}^2) - MSE(S_n^2)}{n^2} &= -3\sigma^4 + 2\mu_4 = 2\mu_4 - 3\sigma^4 \cdot \kappa \cdot \kappa^{-1} \\
&= 2\mu_4 - 3\sigma^4 \frac{\mu_4}{\sigma^4} \cdot \kappa^{-1} \\
&= 2\mu_4 - 3\mu_4 \cdot \kappa^{-1} \\
&= \mu_4(2 - \frac{3}{\kappa})
\end{aligned}
$$

which is only greater then 0 when $\kappa > 1.5$, while the lower bound of κ is just 1.

$\therefore S_n^2$ has a smaller MSE than S_{n-1}^2 for large n if $\kappa > 1.5$

(I tested on Bernoulli distribution with p =0.5, which should have $\kappa = 1$, which indeed show the MSE of S_n^2 to be bigger than S_{n-1}^2 for n>5)

2. Consider a random sample $\{X_1, \ldots, X_n\}$ of size n>2 from $U[0, \theta]$, where θ is positive and finite. We found that X_n is the MLE of θ, and it is easy to see that $2\bar{X}$ is the MME of θ. In the sense of MSE, which one is better ? Please justify your answer.

$$f_{X_{(n)}} = nf_X(x)F_X(x)^{n-1} = \frac{nx^{n-1}}{\theta^n}$$

$$E(X_{(n)}) = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n} dx$$

$$= \frac{n}{\theta^n}[\frac{x^{n+1}}{n+1}]_0^\theta$$

$$= \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1}$$

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

$$Var(X_{(n)}) = \int_0^\theta (x - \frac{n}{n+1}\theta^2)\frac{nx^{n-1}}{\theta^n}dx$$

$$= [\frac{nx^{n+2}}{(n+2)\theta^n}]_0^\theta - \frac{2n}{n+1}\theta[\frac{nx^{n+1}}{(n+1)\theta^n}]_0^\theta + (\theta \cdot \frac{n}{n+1})^2[\frac{x^n}{\theta^n}]_0^\theta$$

$$= \frac{n\theta^2}{n+2} - \frac{2n\theta \cdot n\theta}{(n+1)^2} + \frac{n^2\theta^2}{(n+1)^2}$$

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

$$E(2\bar{X}) = 2\frac{\theta}{2} = \theta$$

$$Var(2\bar{X}) = 4Var(\bar{X}) = \frac{4}{12n}\theta^2 = \frac{\theta^2}{3n}$$

$$MSE(X_{(n)}) = \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} + (\frac{\theta}{n+1})^2$$
$$= \frac{2\theta^2}{(n+2)(n+1)} \in O(\frac{1}{n^2})$$
$$MSE(2\bar{X}) = Var(2\bar{X}) = \frac{\theta^2}{3n} \in O(\frac{1}{n})$$

Since MSE of $X_{(n)}$ is smaller for big n, it is a better estimator

3. Consider a random sample $\{X_1, \ldots, X_n\}$ of size n>1from a uniform distribution on an interval $[\mu - \sqrt{3}\sigma, \mu + \sqrt{3}\sigma]$, where $\mu \in (-\infty, \infty)$ and $\sigma \in (0, \infty)$. Find

   a. the MMEs of μ and σ, and

$$E(X) = \mu$$
$$\therefore \hat{\mu} = \frac{\sum_i^n x_i}{n}$$

$$Var(X) = \frac{1}{12}(\mu + \sqrt{3}\sigma - \mu + \sqrt{3}\sigma)^2 = \frac{\sigma^2}{4}$$
$$\therefore \hat{\sigma} = \sqrt{4S_n(X)}$$

b. the MLEs of μ and σ

$$L = \prod_{i=1}^{n} f_X(x_i|\theta)$$
$$= \prod_{i=1}^{n} \frac{1}{2\sqrt{3}\sigma} \text{ if } \forall i \in n, x_i \in [\mu - \sqrt{3}, \mu + \sqrt{3}], 0 \text{ otherwise}$$
$$\frac{\partial}{\partial\sigma} \log L = -n \log(2\sqrt{3}\sigma)$$

which is monotonic decreasing when σ increases $\implies$ the likelihood is maximized when $\forall i \in n, x_i \in [\mu - \sqrt{3}, \mu + \sqrt{3}]$ and σ is minimized $\iff$ the range is minimized

$\implies X_{(1)} = \mu - \sqrt{3}, X_{(n)} = \mu + \sqrt{3}$

$$\hat{\mu} = \frac{X_{(1)} + X_{(n)}}{2}$$
$$\hat{\sigma} = \frac{X_{(n)} - X_{(1)}}{2\sqrt{3}}$$

4. Consider a random sample $\{X_1, \ldots, X_n\}$ from a distribution with a pdf defined by

$$f(X|\theta) = \begin{cases} \theta x^{\theta-1} & \text{, for } x \in (0,1) \\ 0 & \text{, otherwise} \end{cases}$$

where $0 < \theta < \infty$. Let $g(\theta) = \frac{1}{\theta}$

a) Find the MLE of $g(\theta)$

$$L = \prod_{i=1}^{n} \theta x_i^{\theta-1}$$
$$\log L = n \log \theta + \sum_{i=1}^{n} (\theta - 1) \log x_i$$
$$\frac{\partial L}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i$$

Set $\frac{\partial L}{\partial \theta}$ to 0:

$$\frac{n}{\theta} = -\sum_{i=1}^{n} \log x_i$$
$$\frac{1}{\theta} = -\frac{\sum_{i=1}^{n} \log x_i}{n}$$

$$MLE(g(\theta)) = \frac{-\sum_{i=1}^{n} \log x_i}{n} = \overline{-\log X}$$

b) Is the MLE an unbiased estimator of $g(\theta)$

$$
\begin{aligned}
E(-\overline{\log X}) &= E(-\log X) \\
&= \int_0^1 -\log x \cdot \theta x^{\theta-1} dx \\
&= -\theta \int_0^1 \log x \cdot d\frac{x^\theta}{\theta} \\
&= -\theta([\frac{x^\theta}{\theta}\log x]_0^1 - \int_0^1 \frac{x^\theta}{\theta} d\log x) \\
&= -\theta(0 - \lim_{x \to 0} \frac{x^\theta}{\theta}\log x - \int_0^1 \frac{x^{\theta-1}}{\theta} dx) \\
&= -\theta(-[\frac{x^\theta}{\theta^2}]_0^1) = \frac{1}{\theta} = g(\theta)
\end{aligned}
$$

$\therefore$ the MLE is unbiased

Given that the regularity conditions hold, then

c) Find the C-R inequality for $g(\theta)$

$$
\begin{aligned}
\frac{\partial^2}{\partial \theta^2} L &= -\frac{n}{\theta^2} \\
-E(-\frac{n}{\theta^2}) &= \frac{n}{\theta^2} \\
Var(g(\theta)) &\geq \frac{(\frac{d}{d\theta}\theta^{-1})^2}{\frac{n}{\theta^2}} \\
&= \frac{\theta^2(-\theta^{-2})^2}{n} \\
&= \frac{1}{n\theta^2}
\end{aligned}
$$

$\therefore$ the lower bound $Var(\frac{1}{\theta}) = \frac{1}{n \cdot \theta^2}$

d) Show that the MLE found in (a) is the UMVUE of $g(\theta)$.

$$Var(\overline{-\log X}) = \frac{1}{n} Var(-\log X)$$

$$Var(-\log X) = \int_0^1 (-\log x - \frac{1}{\theta})^2 \theta x^{\theta-1} dx$$

$$= \int_0^1 (\log x^2 + \frac{2\log x}{\theta} + \frac{1}{\theta^2}) \theta x^{\theta-1} dx$$

$$= \int_0^1 \log^2 x \cdot \theta x^{\theta-1} dx + \int_0^1 2\log x \cdot x^{\theta-1} dx + \int_0^1 \frac{1}{\theta} x^{\theta-1} dx$$

$$= \int_0^1 \log^2 x dx^\theta + \int_0^1 2\log x \cdot x^{\theta-1} dx + \int_0^1 \frac{1}{\theta} x^{\theta-1} dx$$

$$= 0 - \lim_{x\to 0} x^\theta \log^2(x) - \int_0^1 2\log x \cdot x^{\theta-1} dx + \int_0^1 2\log x \cdot x^{\theta-1} dx + \int_0^1 \frac{1}{\theta} x^{\theta-1} dx$$

$$= \int_0^1 \frac{1}{\theta} x^{\theta-1} dx$$

$$= \frac{1}{\theta} [\frac{x^\theta}{\theta}]_0^1$$

$$= \frac{1}{\theta^2}$$

$$Var(\overline{-\log X}) = \frac{1}{n\theta^2}$$

which is indeed the lower bound given by CR inequality.

Hence, the MLE is indeed the UMBUE of $g(\theta)$