3. Let $X_1, \ldots, X_n$ are independently uniformly distributed on $(\theta, 2\theta)$.

   (a) **(2 marks)** Find the estimators from the method of moment and the method of maximum likelihood.

   (b) **(2 mark)** Find the expectation and variance of the estimator from the method of moments.

   (c) **(4 marks)** Find the expectation and variance of the estimator from the method of maximum likelihood.

   (d) **(2 mark)** Hence or otherwise, construct two unbiased estimators of θ based on the two estimators in part (a).

   (e) **(2 marks)** Compare the variances of the two unbiased estimates in (d) and comment briefly.

Solutions:

(a) For MME:

$$\begin{aligned} M_1' &= \widetilde{E(X)} \\ \frac{1}{n}\sum_{i=1}^{n} x_i &= \frac{\widetilde{\theta + 2\theta}}{2} \\ \tilde{\theta} &= \frac{2}{3}\bar{X} \end{aligned}$$

For MLE:

$$\begin{aligned} f_{X_i}(x_i) &= \frac{1}{\theta} I_{(\theta \le x_i \le 2\theta)} \\ f_{\mathbf{X}}(\mathbf{x}) &= \frac{1}{\theta^n} I_{(\theta \le x_{(1)} \le x_{(n)} \le 2\theta)} = \frac{1}{\theta^n} I_{(\frac{x_{(n)}}{2} \le \theta \le x_{(1)})} \end{aligned}$$

Therefore, from the graph of $f_{\mathbf{X}}(\mathbf{x})$, we can see that MLE of θ is $\hat{\theta} = \frac{X_{(n)}}{2}$.

(b)

$$\begin{aligned} E(\bar{X}) &= \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{3}{2}\theta \\ Var(\bar{X}) &= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\frac{\theta^2}{12} = \frac{\theta^2}{12n} \\ E(\tilde{\theta}) &= \frac{2}{3}E(\bar{X}) = \theta \\ Var(\tilde{\theta}) &= \frac{4}{9}Var(\bar{X}) = \frac{\theta^2}{27n} \end{aligned}$$

(c)

$$
\begin{aligned}
f_{X_i}(t) &= \frac{1}{\theta} I_{(\theta \le t \le 2\theta)} \\
F_{X_i}(t) &= \frac{t-\theta}{\theta} I_{(\theta \le t \le 2\theta)} + I_{(2\theta \le t)} \\
f_{X_{(n)}}(t) &= n(F_{X_i}(t))^{n-1} f_{X_i}(t) \\
&= n\frac{(t-\theta)^{n-1}}{\theta^n} I_{(\theta \le t \le 2\theta)} \\
E(X_{(n)}) &= \int_\theta^{2\theta} \frac{nt(t-\theta)^{n-1}}{\theta^n} dt \\
&= \int_0^\theta \frac{n(y+\theta)y^{n-1}}{\theta^n} dy \qquad y = t-\theta \\
&= \frac{2n+1}{n+1}\theta \\
E(\hat{\theta}) &= \frac{2n+1}{2n+2}\theta \\
E(X_{(n)}^2) &= \int_\theta^{2\theta} \frac{nt^2(t-\theta)^{n-1}}{\theta^n} dt \\
&= \int_0^\theta \frac{n(y+\theta)^2 y^{n-1}}{\theta^n} dy \qquad y = t-\theta \\
&= \frac{4n^2+8n+2}{(n+1)(n+2)}\theta^2 \\
Var(X_{(n)}) &= E(X_{(n)}^2) - E(X_{(n)})^2 = \frac{n}{(n+1)^2(n+2)}\theta^2 \\
Var(\hat{\theta}) &= \frac{1}{4} Var(X_{(n)}) = \frac{n}{4(n+1)^2(n+2)}\theta^2
\end{aligned}
$$

(d) By part(b), we have $\tilde{\theta} = \frac{2}{3}\bar{X}$ is an unbiased estimator of θ.
By part(c), take S = $\frac{n+1}{2n+1}X_{(n)}$ and E(S) = θ, so S is an unbiased estimator of θ.

(e)

$$
\begin{aligned}
Var(\tilde{\theta}) &= \frac{\theta^2}{27n} \\
Var(S) &= \frac{(n+1)^2}{(2n+1)^2} Var(X_{(n)}) = \frac{n}{(2n+1)^2(n+2)}\theta^2
\end{aligned}
$$

Therefore, when n is large, Var(S) is much smaller than Var($\tilde{\theta}$), and since they are both unbiased, S is better than $\tilde{\theta}$.

4. If $X_1, X_2, \ldots, X_n$ are independently and normally distributed with the same mean μ but different **known** variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$.

(a) **(3 marks)** Find the estimator from the method of moment. Prove that it is an unbiased estimator for μ. Find its variance.

(b) **(3 marks)** Find the estimator from the method of maximum likelihood. Prove that it is an unbiased estimator for μ. Find its variance.

(c) **(2 mark)** Which of the estimators from part (a) or part (b) is more efficient? Explain in details

(d) **(4 marks)** Let $W = \sum_{i=1}^{n}(X_i/\sigma_i^2)/\sum_{j=1}^{n}(1/\sigma_j^2)$. Find its distribution. Hence or otherwise, construct the $(1-\alpha)100\%$ confidence interval for μ.

(e) **(2 marks)** Find the distribution of $X_i - W$.

(f) **(3 marks)** Are W and $X_i - W$ independent? Explain in details.

(g) **(1 mark)** Find the distribution of

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma_i}\right)^2 .$$

(h) **(1 mark)** Find the distribution of

$$\sum_{i=1}^{n}\frac{1}{\sigma_i^2}(W-\mu)^2 .$$

(i) **(4 marks)** Hence or otherwise, find the distribution of

$$\sum_{i=1}^{n}(X_i-W)^2/\sigma_i^2 .$$

Solutions:

(a) For MME:

$$\begin{aligned}
M_1' &= \widetilde{E(X)} \\
\frac{1}{n}\sum_{i=1}^{n}x_i &= \widetilde{\mu} \\
\tilde{\mu} &= \bar{X} \\
E(\bar{X}) &= \frac{1}{n}\sum_{i=1}^{n}E(X_i) = \mu \\
Var(\bar{X}) &= \frac{1}{n^2}\sum_{i=1}^{n}Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2
\end{aligned}$$

(b)

$$
\begin{aligned}
f_{\mathbf{X}}(\mathbf{x}) &= \Pi_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma_i^2}} \\
l(\mu) &= log f_{\mathbf{X}}(\mathbf{x}) = -\frac{1}{2}\sum_{i=1}^{n} log(2\pi\sigma_i^2) - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma_i^2} \\
l'(\mu) &= \sum_{i=1}^{n} \frac{(x_i-\mu)}{\sigma_i^2}
\end{aligned}
$$

Taking $l'(\mu) = 0$, we have $\mu = \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}$ and $l''(\mu) < 0$ at $\mu = \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}\sum_{i=1}^{n}\frac{x_i}{\sigma_i^2}$.Therefore, we get the MLE for μ, which is $\hat{\mu} = \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2}$.

$$
\begin{aligned}
E(\hat{\mu}) &= \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}\sum_{i=1}^{n} E(\frac{X_i}{\sigma_i^2}) = \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}\sum_{i=1}^{n}\frac{\mu}{\sigma_i^2} = \mu \\
Var(\hat{\mu}) &= \frac{1}{(\sum_{i=1}^{n}\frac{1}{\sigma_i^2})^2}\sum_{i=1}^{n}\frac{Var(X_i)}{(\sigma_i^2)^2} = \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}
\end{aligned}
$$

(c)

$$
\begin{aligned}
\frac{Var(\bar{X})}{Var(\hat{\mu})} &= \frac{\frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2}{\frac{1}{\sum_{j=1}^{n}\frac{1}{\sigma_j^2}}} = \frac{1}{n^2}(\sum_{i=1}^{n}\sigma_i^2)(\sum_{j=1}^{n}\frac{1}{\sigma_j^2}) \\
&= \frac{1}{n^2}(\sum_{i=j}\frac{\sigma_i^2}{\sigma_j^2} + \sum_{i\neq j}\frac{\sigma_i^2}{\sigma_j^2}) = \frac{1}{n^2}(n + \sum_{i\neq j}\frac{\sigma_i^2}{\sigma_j^2}) \\
&= \frac{1}{n^2}(n + \sum_{i>j}(\frac{\sigma_i^2}{\sigma_j^2} + \frac{\sigma_j^2}{\sigma_i^2})) \\
&\geq \frac{1}{n^2}(n + \sum_{i>j} 2) = 1
\end{aligned}
$$

Therefore $Var(\bar{X}) \geq Var(\hat{\mu})$, which implies $\hat{\mu}$ is more efficient.

(d)

$$
\begin{aligned}
X_i &\sim N(\mu, \sigma_i^2) \\
\frac{X_i}{\sigma_i^2} &\sim N(\frac{\mu}{\sigma_i^2}, \frac{1}{\sigma_i^2}) \\
\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2} &\sim N(\mu\sum_{i=1}^{n}\frac{1}{\sigma_i^2}, \sum_{i=1}^{n}\frac{1}{\sigma_i^2}) \\
W = \frac{\sum_{i=1}^{n}\frac{X_i}{\sigma_i^2}}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}} &\sim N(\mu, \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}})
\end{aligned}
$$

Therefore, a $(1-\alpha)100\%$ C.I. for μ is $W \pm z_{\alpha/2}\sqrt{\frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}}$

(e) Take $m = Var(W) = \frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}}$, which is a constant.

$$\begin{aligned}
E(X_i - W) &= E(X_i) - EW = 0 \\
Cov(X_i, W) &= Cov(X_i, \frac{m}{\sigma_i^2}X_i) = \frac{m}{\sigma_i^2}\sigma_i^2 = m \\
Var(X_i - W) &= Var(X_i) + Var(W) - 2Cov(X_i, W) \\
&= \sigma_i^2 + m - 2Cov(X_i, \frac{m}{\sigma_i^2}X_i) \\
&= \sigma_i^2 + m - 2\frac{m}{\sigma_i^2}\sigma_i^2 \\
&= \sigma_i^2 - m \\
&= \sigma_i^2 - \frac{1}{\sum_{j=1}^{n}\frac{1}{\sigma_j^2}}
\end{aligned}$$

Therefore, $X_i - W \sim$ N(0, $\sigma_i^2 - \frac{1}{\sum_{j=1}^{n}\frac{1}{\sigma_j^2}}$)

(f)

$$Cov(W, X_i - W) = Cov(W, X_i) - Var(W) = m - m = 0$$

,where m is the same as the one defined in part(e).Then if W and $X_i - W$ follow bivariate normal distribution, then W and $X_i - W$ are independent.
The covariance matrix between W and $X_i - W$ is $\Sigma = \begin{pmatrix} m & 0 \\ 0 & \sigma_i^2 - m \end{pmatrix}$. Let L = $\begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$. Then LL^T is $\begin{pmatrix} a^2 & ab \\ ab & b^2 + c^2 \end{pmatrix}$. We may take a = $\sqrt{m}$, b=0 and c=$\sqrt{\sigma_i^2 - m}$, and then LL^T = Σ. Therefore, there exist 2 independent standard normal variables, i.e., $Z_1, Z_2 \sim$ N(0,1), a 2-vector of μ and a 2×2 matrix L = $\begin{pmatrix} \sqrt{m} & 0 \\ 0 & \sqrt{\sigma_i^2 - m} \end{pmatrix}$ = Chol(Σ) such that $\mathbf{X} = \mu +$ Chol(Σ)$\mathbf{Z}$, so the random vector of $(W, X_i - W)^T$ has a bivariate normal distribution.
Or since $\begin{pmatrix} W \\ X_i - W \end{pmatrix}$ is a linear transformation of $\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$, so the random vector $\begin{pmatrix} W \\ X_i - W \end{pmatrix}$ follows a bivariate normal distribution.
Therefore, together with $Cov(W, X_i - W) = 0$, W and $X_i - W$ are independent.

(g) since $X_i \overset{\text{i.i.d.}}{\sim}$ N(μ, σ_i^2), we have

$$\begin{aligned}
\frac{X_i - \mu}{\sigma_i} &\sim N(0,1) && (1) \\
(\frac{X_i - \mu}{\sigma_i})^2 &\sim \chi_1^2 && (2) \\
\sum_{i=1}^{n}(\frac{X_i - \mu}{\sigma_i})^2 &\sim \chi_n^2 && (3)
\end{aligned}$$

The (3) comes from (2) and the independence of X_i, i=1,2,...n.

(h) From part(d), we have $W \sim N(\mu, \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}})$, so

$$\begin{aligned}
\frac{W-\mu}{\sqrt{\frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}}} &\sim N(0,1) \\
(\frac{W-\mu}{\sqrt{\frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}}})^2 &\sim \chi_1^2 \\
\sum_{i=1}^n \frac{1}{\sigma_i^2}(W-\mu)^2 &\sim \chi_1^2
\end{aligned}$$

(i)

$$\begin{aligned}
\sum_{i=1}^n (\frac{X_i-\mu}{\sigma_i})^2 &= \sum_{i=1}^n (\frac{(X_i-W)+(W-\mu)}{\sigma_i})^2 \\
&= \sum_{i=1}^n (\frac{X_i-W}{\sigma_i})^2 + \sum_{i=1}^n (\frac{W-\mu}{\sigma_i})^2
\end{aligned}$$

because the cross-product term is equal to

$$2\sum_{i=1}^n \frac{(X_i-W)(W-\mu)}{\sigma_i^2} = 2(W-\mu)\sum_{i=1}^n \frac{X_i-W}{\sigma^2} = 0$$

Then,

$$\begin{array}{ccc}
\sum_{i=1}^n (\frac{X_i-\mu}{\sigma_i})^2 & = \sum_{i=1}^n (\frac{X_i-W}{\sigma_i})^2 + & \sum_{i=1}^n (\frac{W-\mu}{\sigma_i})^2 \\
\sim \chi_n^2 & & \sim \chi_1^2
\end{array}$$

Since W and $X_i - W$ are independent, thus $\sum_{i=1}^n (\frac{X_i-W}{\sigma_i})^2 \sim \chi_{n-1}^2$. The detailed proof can be carried out using investigating on the MGFs.

5. **(Bonus)** Consider a random sample $\{X_1, X_2\}$ from density

$$f_X(x|\theta) = \frac{3x^2}{\theta^3} I_{(0<x<\theta)},$$

where $\theta > 0$.

(a) **(2 marks)** Are $\hat{\theta}_1 = \frac{2}{3}(X_1+X_2)$ and $\hat{\theta}_2 = \frac{7}{6}\max(X_1, X_2)$ unbiased for θ?

(b) **(4 marks)** Find the mean squared errors (MSEs) of $\hat{\theta}_1$ and $\hat{\theta}_2$, and compare those estimators.

(c) **(4 marks)** Prove that in the sense of MSE, $T_{8/7}$ is the best estimator of θ among the estimators in form of $T_c = c\max(X_1, X_2)$.

Solutions:

(a) For $\hat{\theta}_1$,

$$\begin{aligned} EX &= \int_0^\theta x\frac{3x^2}{\theta^3}dx = \frac{3}{4}\theta \\ EX^2 &= \int_0^\theta x^2\frac{3x^2}{\theta^3}dx = \frac{3}{5}\theta^2 \\ Var(X) &= EX^2 - (EX)^2 = \frac{3}{80}\theta^2 \\ E\hat{\theta}_1 &= \frac{2}{3}(EX_1 + EX_2) = \theta \end{aligned}$$

Therefore, $\hat{\theta}_1$ is unbiased.
For $\hat{\theta}_2$, take T = max(X_1,X_2).

$$\begin{aligned} f_T(t) &= 2F_X(t)f_X(t)\frac{6t^5}{\theta^6}I_{(0\leq t\leq\theta)} \\ ET &= \int_0^\theta t\frac{6t^5}{\theta^6}dt = \frac{6}{7}\theta \\ E\hat{\theta}_2 &= \theta \end{aligned}$$

Therefore, $\hat{\theta}_2$ is unbiased.

(b) Since both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased, MSEs of $\hat{\theta}_1$ and $\hat{\theta}_2$ are equal to their variance respectively.

$$\begin{aligned} Var(\hat{\theta}_1) &= \frac{4}{9}(Var(X_1) + Var(X_2)) = \frac{1}{30}\theta^2 \\ E(T^2) &= \int_0^\theta t^2\frac{6t^5}{\theta^6}dt = \frac{3}{4}\theta^2 \\ Var(T) &= ET^2 - (ET)^2 = \frac{3}{196}\theta^2 \\ Var(\hat{\theta}_2) &= \frac{1}{48}\theta^2 \end{aligned}$$

(c)

$$\begin{aligned} T_c &= cmax(X_1, X_2) = cT \\ E(T_c) &= \frac{6c}{7}\theta \\ Var(T_c) &= \frac{3c^2}{196}\theta^2 \\ MSE\ of\ T_c &= Var(T_c) + (ET_c - \theta)^2 \\ &= \frac{3c^2}{196}\theta^2 + (\frac{6c}{7}\theta - \theta)^2 \\ &= (\frac{3}{4}c^2 - \frac{12}{7}c + 1)\theta^2 \end{aligned}$$

Therefore, MSE of T_c achieves its minimum at $c = -\frac{-\frac{12}{7}}{2\frac{3}{4}} = \frac{8}{7}$, so $T_{\frac{8}{7}}$ is the best estimator of θ among T_c in the sense of MSE.