M378K Introduction to Mathematical Statistics
Homework assignment #8

Please, provide your **final answer only** to the following problems.

Problem 8.1. *(5 points) Which of the following estimators is **not** unbiased for μ if $Y_1, \dots, Y_n$ is a random sample from the normal distribution $N(\mu, \sigma)$:*

(a) Y_n

(b) $\frac{1}{2}(Y_1 + Y_2)$

(c) $Y_1 - Y_2 + Y_3$

(d) $\bar{Y}$

(e) **All of the above are unbiased.**

Solution: The correct answer is **(e)** since (a)-(d) are all unbiased.

Problem 8.2. *(5 points) Let $Y_1, \dots, Y_n$ be a random sample of size $n \geq 2$, from $N(\mu, \sigma)$ and let the estimators $\hat{\mu}_1$, $\hat{\mu}_2$ and $\hat{\mu}_3$, for μ, be given by*

$$\hat{\mu}_1 = Y_1, \hat{\mu}_2 = \frac{1}{2}(Y_1 + Y_2) \text{ and } \hat{\mu}_3 = \bar{Y}.$$

Then, no matter what μ and σ are, we always have

(a) $MSE(\hat{\mu}_1) \leq MSE(\hat{\mu}_2) \leq MSE(\hat{\mu}_3)$

(b) $MSE(\hat{\mu}_3) \leq MSE(\hat{\mu}_2) \leq MSE(\hat{\mu}_1)$

(c) $MSE(\hat{\mu}_3) \leq MSE(\hat{\mu}_1) \leq MSE(\hat{\mu}_2)$

(d) $MSE(\hat{\mu}_1) \leq MSE(\hat{\mu}_3) \leq MSE(\hat{\mu}_2)$

(e) **None of the above.**

Solution: The correct answer is **(b)**.
All of these are unbiased, so their MSEs are just their variances. We have $\text{Var}[\hat{\mu}_1] = \sigma^2$, $\text{Var}[\hat{\mu}_2] = \frac{1}{4}(\text{Var}[\hat{\mu}_1] + \text{Var}[\hat{\mu}_2]) = \sigma^2/2$ and $\text{Var}[\hat{\mu}_3] = \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}[Y_i] = \sigma^2/n$.

Please, provide your **complete solutions** to the following problems. Final answers only, even if correct will earn zero points for those problems.

Problem 8.3. *(30 points) Let (Y_1, Y_2) be a random sample (of size $n = 2$) from the uniform distribution $U(0, \theta)$, with $\theta > 0$ unknown.*

1. *($2 + 3 + 10 = 15$ points) Find constants c_1, c_2 and c_3 such that the following estimators*

$$\hat{\theta}_1 = c_1 Y_1, \ \hat{\theta} = c_2 Y_2 \text{ and } \hat{\theta}_3 = c_3 \max(Y_1, Y_2),$$

   are unbiased. (Hint: For $\hat{\theta}_3$, integrate the function $\max(y_1, y_2)$ multiplied by the joint density of Y_1, Y_2. Split the integral over $[0, \theta] \times [0, \theta]$ into two parts - one where $y_1 \geq y_2$ and the other where $y_1 < y_2$ and note that $\max(y_1, y_2) = y_1 1_{\{y_1 \geq y_2\}} + y_2 1_{\{y_1 < y_2\}}$.)

2. *($2 + 3 + 10 = 15$ points) With values c_1, c_2 and c_3 as above, compute mean-squared errors MSE($\hat{\theta}_1$), MSE($\hat{\theta}_2$) and MSE($\hat{\theta}_3$) of $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{\theta}_3$.*

3. *(10 points) Sketch the graphs of MSE($\hat{\theta}_1$), MSE($\hat{\theta}_2$) and MSE($\hat{\theta}_3$) as functions of θ. Is one of the three clearly better (in the mean-square sense) than the others?*

Solution:

1. Using the uniform pdf, we compute

$$\mathbb{E}[c_1 \hat{\theta}_1] = \mathbb{E}[c_1 Y_1] = c_1 \int_0^\theta y \frac{1}{\theta} \, dy = c_1 \, \tfrac{\theta}{2},$$

   and it follows that $\hat{\theta}_1$ is unbiased when $c_1 = 2$. The same computation yields $c_2 = 2$. To compute c_3, we remember that Y_1 and Y_2 are independent and, so, their joint density is given by

$$f_{Y_1, Y_2}(y_1, y_2) = \frac{1}{\theta^2} 1_{\{0 \leq y_1, y_2 \leq \theta\}}.$$

   Therefore,

$$\begin{aligned} \mathbb{E}[c_3 \max(Y_1, Y_2)] &= c_3 \int_0^\theta \int_0^\theta \frac{1}{\theta^2} \max(y_1, y_2) \, dy_2 \, dy_1 \\ &= \tfrac{c_3}{\theta^2} \int_0^\theta \int_0^\theta y_1 1_{\{y_1 \geq y_2\}} \, dy_2 \, dy_1 + \tfrac{c_3}{\theta^2} \int_0^\theta \int_0^\theta y_2 1_{\{y_2 > y_1\}} \, dy_2 \, dy_1 \end{aligned}$$

   We compute the two integrals separately:

$$\int_0^\theta \int_0^\theta y_1 1_{\{y_1 \geq y_2\}} \, dy_2 \, dy_1 = \int_0^\theta \int_0^{y_1} y_1 dy_2 \, dy_1 = \int_0^\theta y_1^2 \, dy_1 = \tfrac{\theta^3}{3},$$

   and

$$\int_0^\theta \int_0^\theta y_2 1_{\{y_2 > y_1\}} \, dy_2 \, dy_1 = \int_0^\theta \int_{y_1}^\theta y_2 dy_2 \, dy_1 = \int_0^\theta \tfrac{1}{2}(\theta^2 - y_1^2) \, dy_1 = \tfrac{1}{2}(\theta^3 - \theta^3/3) = \tfrac{\theta^3}{3}.$$

We put this all together to get that $\mathbb{E}[c_3 \max(Y_1, Y_2)] = \frac{2c_3\theta}{3}$ and we conclude that $\hat{\theta}_3$ is unbiased for $c_3 = 3/2$.

2. The values c_1, c_2 and c_3 are chosen to make $\hat{\theta}_1, \hat{\theta}_2$ and $\hat{\theta}_3$ unbiased, so we have the following formula
$$\mathrm{MSE}(\hat{\theta}_1) = \mathrm{Var}[\hat{\theta}_1] = \mathbb{E}[\hat{\theta}_1^2] - \mathbb{E}[\hat{\theta}_1]^2 = \mathbb{E}[\hat{\theta}_1^2] - \theta^2.$$
Since
$$\mathbb{E}[\hat{\theta}_1^2] = \frac{1}{\theta}\int_0^\theta (2y)^2\, dy = \frac{4\theta^2}{3},$$
we have $\mathrm{MSE}(\hat{\theta}_1) = \theta^2(\frac{4}{3} - 1) = \frac{1}{3}\theta^2$. Similarly, $\mathrm{MSE}(\hat{\theta}_2) = \frac{1}{3}\theta^2$. A similar computation, but now involving a double integral (which we split just like in 1. above) yields:
$$\begin{aligned}\mathbb{E}[\hat{\theta}_3^2] &= \frac{c_3^2}{\theta^2}\int_0^\theta \int_0^\theta \max(y_1, y_2)^2\, dy_2\, dy_1 \\ &= \frac{c_3^2}{\theta^2}\Big(\int_0^\theta \int_0^\theta y_1^2 1_{\{y_1 \geq y_2\}}\, dy_2\, dy_1 + \int_0^\theta \int_0^\theta y_2^2 1_{\{y_1 < y_2\}}\, dy_2\, dy_1\Big).\end{aligned}$$
We skip the steps in the evaluation of the two integrals, and report that $\mathbb{E}[\hat{\theta}_3^2] = \frac{c_3^2}{2\theta^2} = \frac{9}{8}\theta^2$. Therefore,
$$\mathrm{MSE}(\hat{\theta}_3) = \frac{9}{8}\theta^2 - \theta^2 = \frac{1}{8}\theta^2.$$

3. Here is the graph of $\mathrm{MSE}(\hat{\theta}_1) = \mathrm{MSE}(\hat{\theta}_2)$ and $\mathrm{MSE}(\hat{\theta}_3)$ for $\theta \in [0, 10]$. We see that $\hat{\theta}_3$ has a strictly smaller mean-square error then either $\hat{\theta}_1$ or $\hat{\theta}_2$. Therefore, $\hat{\theta}_3$ is better (at least in the mean-square sense).