M378K Introduction to Mathematical Statistics
Homework assignment #7

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

Problem 7.1. *(5 points) Let $Y_1, \dots, Y_n$ be a random sample from $U(0, \theta)$, with an unknown $\theta > 0$. For what value of the constant c is the estimator $\hat{\theta} = c \sum_{i=1}^{n} Y_i$ unbiased for θ?*

(a) 1

(b) $1/n$

(c) $2/n$

(d) n

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

Solution: The correct answer is **(c)**.
We have

$$\mathbb{E}[c \sum_{i=1}^{n} Y_i] = c \sum_{i=1}^{n} \tfrac{\theta}{2} = \tfrac{nc}{2}\theta,$$

so c must be equal to $2/n$.

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

Problem 7.2. *(20 points)* Source: "Probability" by Pitman
Four people agree to meet at a cafe at noon. Suppose that each person arrives at a time normally distributed with mean 12noon and standard deviation of 5 minutes, independently of all the others.

1. *(5 points) What is the chance that the first person to get to the cafe arrives before 11:50am?*
2. *(5 points) What is the chance that some of the four have still not arrived at 12:15pm?*
3. *(10 points) Approximately, what is the chance that the second person to arrive gets there within 10 seconds on 12noon?*

Solution: Let the random variables X_1, X_2, X_3 and X_4 denote the arrival times of the four people centered around 12noon and measured in minutes. Then, $X_i, i = 1, \dots, 4$ are independent and all have distribution $N(0, \sigma = 5)$. When we create the order statistics of this random sample, we will have ordered the people in order of arrival from the earliest to the latest.

1. $$\mathbb{P}[X_{(1)} < -10] = 1 - (\mathbb{P}[X_1 \geq -10])^4 = 1 - (\mathbb{P}[X_1/5 \geq -2])^4 = 1 - (\Phi(2))^4 = 1 - (0.9772)^4 = 0.0881.$$

2.

$$\mathbb{P}[X_{(4)} > 15] = 1 - \mathbb{P}[X_{(4)} \leq 15] = 1 - (\mathbb{P}[X_1 \leq 15])^4 = 1 - (\mathbb{P}[X_1/5 \leq 3])^4 = 1 - (0.9986)^4 = 0.0056.$$

3.

$$\mathbb{P}[-1/6 < X_{(2)} < 1/6] = \int_{-1/6}^{1/6} g_{(2)}(x)\, dx$$

with $g_{(2)}$ denoting the density function of the second order statistic, i.e.,

$$g_{(2)}(x) = 4\binom{4-1}{2-1} f_X(x) F_X(x)(1 - F_X(x))^{4-2}.$$

We can approximate the required probability as:

$$\begin{aligned}
\mathbb{P}[-1/6 < X_{(2)} < 1/6] &\approx \frac{1}{3} g_{(2)}(0) = \frac{1}{3} \times 12 f_X(0) F_X(0)(1 - F_X(0))^2 \\
&= 4 \times \frac{1}{5\sqrt{2\pi}} (1/2)(1/2)^2 \\
&= \frac{1}{10\sqrt{2\pi}} \approx 0.0399.
\end{aligned}$$

Problem 7.3. *(5 points) Consider an estimator $\hat{\theta}$ for a parameter θ. Let's say that*

$$\mathbb{E}[\hat{\theta}] = \kappa_1 \theta + \kappa_2$$

for some constants $\kappa_i \neq 0, i = 1, 2$. Is the estimator $\hat{\theta}$ unbiased? If so, justify your answer; if not, how would you transform the estimator $\hat{\theta}$ to obtain an unbiased estimator?

Solution: The condition for unbiasedness is

$$\mathbb{E}[\hat{\theta}] = \theta.$$

So, the only scenario in which $\hat{\theta}$ is unbiased is if $\kappa_2 = 0$ and $\kappa_1 = 1$. In general, we could define a new estimator $\hat{\theta}'$ as follows:

$$\hat{\theta}' = \frac{\hat{\theta} - \kappa_2}{\kappa_1}.$$

Indeed, we would have

$$\mathbb{E}[\hat{\theta}'] = \mathbb{E}\left[\frac{\hat{\theta} - \kappa_2}{\kappa_1}\right] = \frac{1}{\kappa_1}(\mathbb{E}[\hat{\theta}] - \kappa_2) = \frac{1}{\kappa_1}(\kappa_1\theta + \kappa_2 - \kappa_2) = \theta.$$

Problem 7.4. *(20 points) Let $Y_1, \ldots, Y_n$ be a random sample from the uniform distribution on $[0, \theta]$, where $\theta > 0$ is an unknown parameter. We consider the estimator*

$$\hat{\theta} = c\,\frac{1}{n}\sum_{i=1}^{n} Y_i^2.$$

where c is a constant (not dependent on θ or on $Y_1, \ldots, Y_n$).

1. *(10 points) For what value of the constant c will $\hat{\theta}$ be an unbiased estimator for θ^2? Is there such a value if $\hat{\theta}$ is used as an estimator for θ instead of θ^2?*
2. *(10 points) Using the value of c obtained above, compute the mean squared error of $\hat{\theta}$ (when interpreted as an estimator of θ^2).*

Solution:

1. For any i, we have $\mathbb{E}[Y_i^2] = \frac{1}{\theta}\int_0^\theta y^2\,dy = \frac{1}{3}\theta^2$. Therefore

   $$\mathbb{E}[\hat{\theta}] = \tfrac{c}{n}\mathbb{E}[\textstyle\sum_{i=1}^n Y_i^2] = \tfrac{c}{n}\textstyle\sum_{i=1}^n \mathbb{E}[Y_i^2] = \tfrac{c}{3}\theta^2.$$

   It follows that $\hat{\theta}$ is unbiased for θ^2 when $c = 3$, but that there is no c that will make it unbiased for θ.

2. We take $c = 3$ as above. Since $\hat{\theta}$ is unbiased, its mean-squared error is the same as its variance, so that

   $$MSE(\hat{\theta}) = \mathrm{Var}[\hat{\theta}] = c^2\frac{1}{n^2}\textstyle\sum_{i=1}^n \mathrm{Var}[Y_i^2].$$

   Since $\mathbb{E}[Y_i^2] = \frac{1}{3}\theta^2$, we have $\mathrm{Var}[Y_i^2] = \mathbb{E}[Y_i^4] - \mathbb{E}[Y_i^2]^2 = \frac{1}{\theta}\int_0^\theta y^4\,dy - \frac{1}{9}\theta^4 = \frac{4}{45}\theta^4$. Hence

   $$MSE(\hat{\theta}) = c^2\frac{1}{n^2} \times n\tfrac{4}{45}\theta^4 = \tfrac{4}{5n}\theta^4.$$