HOMEWORK 04

PSTAT 120B: Mathematical Statistics, I
Summer Session A, 2024 with Instructor: Ethan P. Marzban

---

1. *Gaucho Gourmande*, a popular restaurant chain, has just opened several locations in the Central Coast area. Corporate is interested in estimating the shortest amount of time a randomly-selected customer needs to wait in line at a *Gaucho Gourmande* location before being served. To that end, they hire an analyst who takes an i.i.d. sample $\{Y_i\}_{i=1}^n$ of customer wait times; additionally, the analyst decides to model wait times using a Pareto(θ, α) distribution. That is,

$$f_{Y_i}(y) = \frac{\alpha \theta^\alpha}{y^{\alpha+1}} \cdot \mathbb{1}_{\{y \geq \theta\}}$$

for some fixed $\alpha > 3$. Notice that θ is therefore the parameter of interest - i.e. in the context of the problem, it represents the shortest amount of time a randomly-selected *Gaucho Gourmande* customer needs to wait in line before being served.

   (a) Show that if $Y_i \sim$ Pareto(α, θ) then

$$\mathbb{E}[Y_i] = \frac{\alpha\theta}{\alpha - 1} \text{ if } \alpha > 1 \qquad \text{and} \qquad \text{Var}(Y_i) = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)} \text{ if } \alpha > 2$$

   (b) Given that θ is a population minimum, it seems plausible to use the sample minimum as an estimator for θ. Is $Y_{(1)}$ an unbiased estimator for θ? Is it an asymptotically unbiased estimator for θ?

   Hint: On a previous Discussion Worksheet, you actually showed that $Y_{(n)}$ follows a Pareto Distribution! Hence, identify the appropriate parameters and utilize your result from part (a) to avoid having to perform any integrals in this part.

   (c) Compute MSE$(Y_{(1)}, \theta)$.

2. In some cases it becomes necessary to appeal to the *definition* of consistency in order to determine whether a given estimator is consistent or not. Let $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim}$ Unif$[0, \theta]$ for $n \geq 2$, and consider using $Y_{(n)}$, the sample maximum, as an estimator for θ.

   (a) Derive the sampling distribution of $Y_{(n)}$.

   (b) Use your answer to part (a) to show that $Y_{(n)}$ is a biased estimator for θ.

   (c) For a fixed $\varepsilon > 0$, show that

$$\mathbb{P}(|Y_{(n)} - \theta| < \varepsilon) = \begin{cases} 1 - \left(1 - \frac{\varepsilon}{\theta}\right)^n & \text{if } \varepsilon < \theta \\ 1 & \text{if } \varepsilon > \theta \end{cases}$$

   Be careful with you reasoning here!

   (d) Take the limit of your answer to part (c) as $n \to \infty$, to show that

$$\lim_{n \to \infty} \mathbb{P}(|Y_{(n)} - \theta| < \varepsilon) = 1$$

   thereby establishing that $Y_{(n)}$ is a consistent estimator for θ.

3. The IRS (Internal Revenue Service) offers a phone line which can be used to answer certain tax-related questions - however, after dialing the line, it is common to have to wait in a queue before your call is answered. Suppose the average amount of time spent waiting in the queue is 34 minutes, and that the standard deviation of all wait times is 20 minutes. Further suppose we collect a random sample of 81 independent wait times.

   (a) What is the approximate probability that the average wait time among the 81 sampled wait times lies within 2 minutes of the true average of 34 minutes?

   (b) What is the approximate probability that the *aggregate* (i.e. total) wait time of the 81 sampled wait times is greater than 2750 minutes?

4. (a) Let $Y_1, \cdots, Y_n$ denote an i.i.d. sample from the distribution with density $f_Y(y) = e^{-(y-\theta)} \cdot \mathbb{1}_{\{y \geq \theta\}}$. Find $\widehat{\theta}_{\mathrm{MM}}$, the Method of Moments estimators for θ.

   (b) Let $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim}$ Gamma(α, β) for some $\alpha > 0$ and $\beta > 0$. Find $\widehat{\alpha}_{\mathrm{MM}}$ and $\widehat{\beta}_{\mathrm{MM}}$, the Method of Moments estimators for α and β, respectively.