# QUIZ 03

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

## Your Information

**Your Name:** ______________________________
(*First and Last*)

**Your NetID:** ______________________________
(*NOT Perm Number*)

**Your Section:** **2pm (Hyuk-Jean)** **3pm (Hyuk-Jean)** **4pm (Minwoo)** **5pm (Minwoo)**
(*Circle One*)

## Instructions

- **PLEASE DO NOT REMOVE ANY PAGES FROM THIS QUIZ.**
- You have **20 minutes** to complete this quiz.
- The use of **calculators** is permitted, but the use of any other aids (including, but not limited to, notes, phones, laptops, etc.) is prohibited.
- Ensure you show **all your work**; correct answers with no supporting justification will not receive full credit.
- Kindly note that any instances of academic misconduct (including, but not limited to: unauthorized collaboration, or the use of unauthorized resources) will be dealt with in the strictest manner possible.
- Please refrain from discussing the contents of this quiz with **ANYONE**, until given the all-clear by the Instructor.
- **Good Luck!**

**This quiz contains 15 points in total**

**The Quiz Begins on the Next Page**

**Problem 1.** Let $Y_1, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} f(y; \theta)$ where

$$f(y;\theta) = \frac{3}{2\theta^3} \cdot y^2 \cdot \mathbb{1}_{\{-\theta \leq y \leq \theta\}}$$

and $\theta > 1$ is an unknown parameter.

(a) (4 points) Derive an expression for $\widehat{\theta}_{\text{MM}}$, the method of moments estimator for $\theta$.

**Solution:** First note that

$$\mathbb{E}[Y_i] = \int_{-\theta}^{\theta} \frac{3}{2\theta^3} y^3 \, \mathrm{d}y = 0$$

Hence, we need the second moment:

$$\mathbb{E}[Y_i^2] = \int_{-\theta}^{\theta} \frac{3}{2\theta^3} y^4 \, \mathrm{d}y = \frac{3}{2\theta^3} \cdot \frac{1}{5}(\theta^5 + \theta^5) = \frac{3}{5}\theta^2$$

Therefore, $\widehat{\theta}_{\text{MM}}$ must satisfy

$$\frac{3}{5}\widehat{\theta}_{\text{MM}}^2 = M_{2,n} \implies \widehat{\theta}_{\text{MM}} = \sqrt{\frac{5}{3} M_{2,n}} = \sqrt{\frac{5}{3n} \sum_{i=1}^{n} Y_i^2}$$

(b) (4 points) Suppose that, for a given sample $\vec{y}$ from this distribution, we find that the maximum likelihood estimate for $\theta$ is given by $1.98$. What is the maximum likelihood estimate of $\mathbb{P}(Y_1 \geq 1)$? You do **not** need to simplify your answer, but you should justify your logic clearly.

**Solution:** First note that

$$\mathbb{P}(Y_1 > 1) = \int_{1}^{\theta} \frac{3}{2\theta^3} y^2 \, \mathrm{d}y = \frac{3}{2\theta^3} \cdot \frac{1}{3}(\theta^3 - 1) = \frac{1}{2} - \frac{1}{2\theta^3}$$

By the **equivariance property** of the maximum likelihood estimator,

$$\widehat{\mathbb{P}(Y_1 \geq 1)}_{\text{MLE}} = \widehat{\left( \frac{1}{2} - \frac{1}{2\theta^3} \right)}_{\text{MLE}} = \frac{1}{2} - \frac{1}{2\widehat{\theta}_{\text{MLE}}^3}$$

meaning our maximum likelihood estimate for $\mathbb{P}(Y_1 > 1)$ is given by

$$\frac{1}{2} - \frac{1}{2(1.98)^3}$$

**QUIZ CONTINUES ON NEXT PAGE**

**Problem 2.** (7 points) The wait time (in minutes) of a randomly-selected customer at the Goleta DMV (Department of Motor Vehicles) follows a Gamma$(3, \theta)$ distribution for some unknown $\theta > 0$. Let $\vec{\boldsymbol{Y}} := \{Y_i\}_{i=1}^n$ denote the wait times (in minutes) of an i.i.d. sample of $n$ individuals (where $n > 1$), selected from the Goleta DMV.

Derive an expression for $\widehat{\theta}_{\mathsf{MLE}}$, the maximum likelihood estimator for $\theta$. Show all your work!

**Solution:** We start by finding the likelihood:

$$\begin{aligned}\mathcal{L}_{\vec{\mathbf{Y}}}(\theta) &= \prod_{i=1}^{n}(f_Y;\theta) = \prod_{i=1}^{n}\left[\frac{1}{2\theta^3}Y_i^2 e^{-Y_i/\theta}\cdot \mathbb{1}_{\{Y_i\geq 0\}}\right] \\ &= \left(\frac{1}{2}\right)^n \cdot (\theta)^{-3n}\cdot \prod_{i=1}^{n}(Y_i^2)\cdot \exp\left\{-\frac{1}{\theta}\sum_{i=1}^{n}Y_i\right\}\cdot \prod_{i=1}^{n}\mathbb{1}_{\{Y_i\geq 0\}}\end{aligned}$$

The log-likelihood is therefore given by

$$\ell_{\vec{\mathbf{Y}}}(\theta) = \ln \mathcal{L}_{\vec{\mathbf{Y}}}(\theta) = -3n\ln(\theta) + \sum_{i=1}^{n}\ln(Y_i^2) - \frac{1}{\theta}\sum_{i=1}^{n}Y_i + \sum_{i=1}^{n}\ln \mathbb{1}_{\{Y_i\geq 0\}}$$

Taking the first derivative yields

$$\frac{\partial}{\partial\theta^2}\ell_{\vec{\mathbf{Y}}}(\theta) = -\frac{3n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n}Y_i$$

which means

$$-\frac{3n}{\widehat{\theta}_{\mathsf{MLE}}} + \frac{1}{\widehat{\theta}_{\mathsf{MLE}}^2}\sum_{i=1}^{n}Y_i = 0$$

Solving for $\widehat{\theta}_{\mathsf{MLE}}$ yields

$$\boxed{\widehat{\theta}_{\mathsf{MLE}} = \frac{1}{3n}\sum_{i=1}^{n}Y_i = \frac{\overline{Y}_n}{3}}$$

**END OF QUIZ**

**Selected Discrete Distributions**

| Distribution | PMF | Expec./Var. | MGF |
|---|---|---|---|
| $X \sim \text{Bin}(n, p)$ | $p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k \in \{0, \cdots, n\}$ | $\mathbb{E}[X] = np$ <br> $\text{Var}(X) = np(1-p)$ | $M_X(t) = (1 - p + pe^t)^n$ |
| $X \sim \text{Geom}(p)$ | $p_X(k) = p(1-p)^{k-1}$, $k \in \{1, 2, \cdots\}$ | $\mathbb{E}[X] = \frac{1}{p}$ <br> $\text{Var}(X) = \frac{1-p}{p^2}$ | $M_X(t) = \frac{pe^t}{1-(1-p)e^t}$, $t < -\ln(1-p)$ |
| $X \sim \text{Geom}(p)$ | $p_X(k) = p(1-p)^k$, $k \in \{0, 1, \cdots\}$ | $\mathbb{E}[X] = \frac{1-p}{p}$ <br> $\text{Var}(X) = \frac{1-p}{p^2}$ | $M_X(t) = \frac{p}{1-(1-p)e^t}$, $t < -\ln(1-p)$ |
| $X \sim \text{NegBin}(r, p)$ | $p_X(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$, $k \in \{r, r+1, \cdots\}$ | $\mathbb{E}[X] = \frac{r}{p}$ <br> $\text{Var}(X) = \frac{r(1-p)}{p^2}$ | $M_X(t) = \left[\frac{pe^t}{1-(1-p)e^t}\right]^r$, $t < -\ln(1-p)$ |
| $X \sim \text{Pois}(\lambda)$ | $p_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$, $k \in \{0, 1, \cdots\}$ | $\mathbb{E}[X] = \lambda$ <br> $\text{Var}(X) = \lambda$ | $M_X(t) = \exp\left\{\lambda(e^t - 1)\right\}$ |

**Selected Continuous Distributions**

| Distribution | PDF | Expec./Var. | MGF |
|---|---|---|---|
| $X \sim \text{Unif}[a, b]$ | $f_X(x) = \dfrac{1}{b-a} \cdot \mathbb{1}_{\{x \in [a,b]\}}$ | $\mathbb{E}[X] = \frac{a+b}{2}$ <br> $\text{Var}(X) = \frac{(b-a)^2}{12}$ | $M_X(t) = \begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)} & \text{if } t \neq 0 \\ 1 & \text{if } t = 0 \end{cases}$ |
| $X \sim \text{Exp}(\beta)$ | $f_X(x) = \dfrac{1}{\beta} e^{-x/\beta} \cdot \mathbb{1}_{\{x \geq 0\}}$ | $\mathbb{E}[X] = \beta$ <br> $\text{Var}(X) = \beta^2$ | $M_X(t) = (1 - \beta t)^{-1},\ t < 1/\beta$ |
| $X \sim \text{Gamma}(\alpha, \beta)$ | $f_X(x) = \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \cdot \mathbb{1}_{\{x \geq 0\}}$ | $\mathbb{E}[X] = \alpha\beta$ <br> $\text{Var}(X) = \alpha\beta^2$ | $M_X(t) = (1 - \beta t)^{-\alpha},\ t < 1/\beta$ |
| $X \sim \mathcal{N}(\mu, \sigma^2)$ | $f_X(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}$ | $\mathbb{E}[X] = \mu$ <br> $\text{Var}(X) = \sigma^2$ | $M_X(t) = \exp\left\{\mu t + \frac{\sigma^2}{2} t\right\}$ |

**Calculus Results**

- **Product Rule:** $\dfrac{\mathrm{d}}{\mathrm{d}x}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x)$
- **Power Rule:** $\dfrac{\mathrm{d}}{\mathrm{d}x}(x^n) = nx^{n-1}$
- **Gamma Function:** $\Gamma(r) := \displaystyle\int_0^\infty t^{r-1} e^{-t}\, \mathrm{d}t$ if $r > 0$
- **Quotient Rule:** $\dfrac{\mathrm{d}}{\mathrm{d}x}\left[\dfrac{f(x)}{g(x)}\right] = \dfrac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$
- **Chain Rule:** $\dfrac{\mathrm{d}}{\mathrm{d}x} f(g(x)) = f'(g(x)) \cdot g'(x)$
- **Recursiveness of Gamma Fnt.:** $\Gamma(r) = (r-1) \cdot \Gamma(r-1)$

**Selected Statistics Results**

- **Standarization of Normals:** If $Y \sim \mathcal{N}(0,1)$, then $Y^2 \sim \chi^2_1$.
- **Closure of $\chi^2$ under Sums:** Given independent random variables $\{Y_i\}_{i=1}^n$ with $Y_i \sim \chi^2_{\nu_i}$, then

$$(Y_1 + \cdots + Y_n) \sim \chi^2_{\nu_1 + \cdots + \nu_n}$$

- **Closure of Gamma under Scaling:** if $Y \sim \text{Gamma}(\alpha, \beta)$, then $(cY) \sim \text{Gamma}(\alpha, c\beta)$ for $c > 0$
- **Sample Mean:** $\overline{Y}_n := \frac{1}{n}\sum_{i=1}^{n} Y_i$ **Sample Variance:** $S_n^2 := \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y}_n)^2$

  **Sample Standard Deviation:** $S_n := \sqrt{S_n^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y}_n)^2}$
- **Modified Standardization Result:** Given $Y_1, Y_2, \cdots \overset{\text{i.i.d.}}{\sim} \mathcal{N}(\mu, \sigma^2)$,

$$\sqrt{n}\left(\frac{\overline{Y}_n - \mu}{S_n}\right) \sim t_{n-1}$$

- **Sampling Distribution of First Order Statistic:** $f_{Y_{(1)}} = n[1 - F_Y(y)]^{n-1} f_Y(y)$

  **Sampling Distribution of $n^{\text{th}}$ Order Statistic:** $f_{Y_{(n)}}(y) = n[F_Y(y)]^{n-1} f_Y(y)$
- **Factorization Theorem:** $\mathcal{L}_{\vec{\boldsymbol{Y}}}(\theta) = g(U, \theta) \times h(\vec{\boldsymbol{Y}}) \iff U$ is sufficient for $\theta$.
- **Equivariance Property of MLE:** $\widehat{\tau(\theta)}_{\text{MLE}} = \tau\left(\widehat{\theta}_{\text{MLE}}\right)$

**You may use the remainder of this page for scratch work;** please note that nothing written on this page will be graded.