Worksheet 09 - Method of Moments, Maximum Likelihood Estimators and Bias

Shawn Kim

***Directions**: Please upload a PDF to Gradescope that includes both your written responses and corresponding R code inputs/outputs (if requested) for each problem.*

Special Directions This problem starts with a past Group Activity. If you use an equation derived in the Group Activity, simply cite it as "From GA:". Be sure to demonstrate the correct use of expectation notation and mathematical notation/reasoning in your work. Clearly show your work/reasoning by entering all necessary algebra/calculations as text or inserting a clear well-cropped image of your work using an R chunk.

Method of Moments 2

Problem 1.

Recall from Group Activity 16

Daily rainfall data, in millimeters, is modeled as having a $\Gamma(1/2, \beta)$ distribution. The density is

$$f_X(x|1/2, \beta) = \begin{cases} 0 & \text{for } x \leq 0, \\ \frac{\beta^{1/2}}{\sqrt{\pi}} x^{-1/2} e^{-\beta x} & \text{for } x > 0. \end{cases}$$

(Do Not Write Up Solutions Again) You already found the method of moments estimator for β based on rainfall amounts $x_1, x_2, \ldots, x_n$. Hint: For a $\Gamma(\alpha, \beta)$ the distributional mean is $\mu = \dfrac{\alpha}{\beta}$ and the distributional variance is $\sigma^2 = \dfrac{\alpha}{\beta^2}$.

(Do Not Write Up Solutions Again) You already gave the estimate $\hat{\beta}$ for the monsoon rainfall amounts in millimeters during July and August, 2017 for Tucson, Arizona.

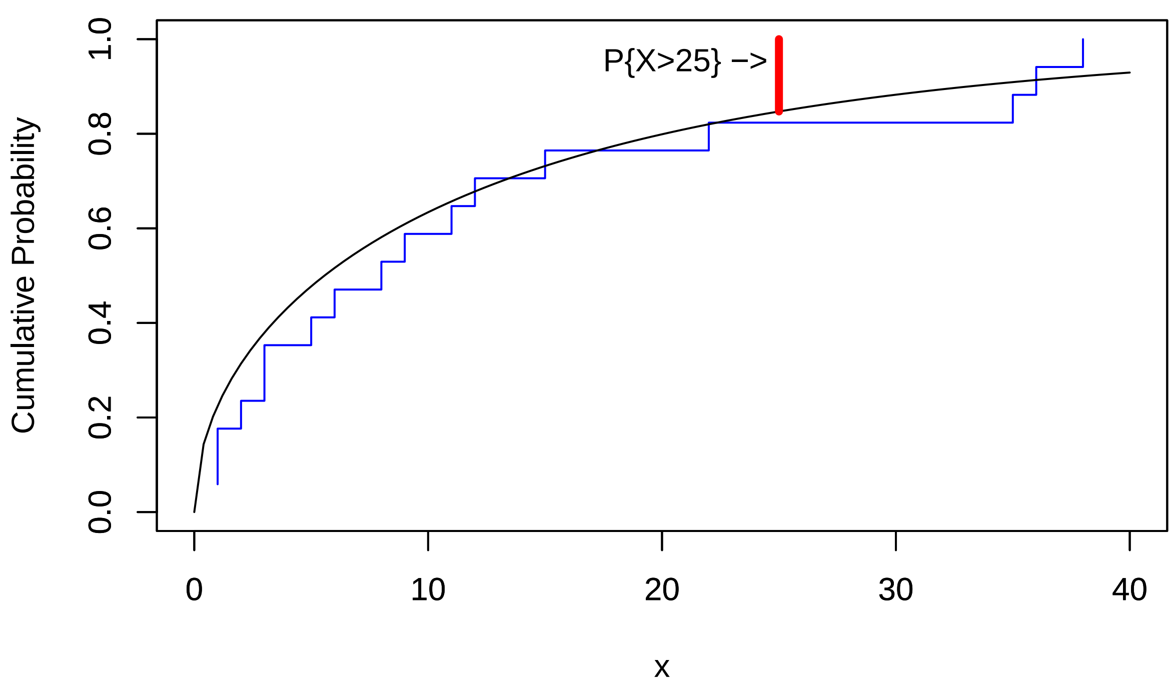
3	15	1	37	5	1	8	11	6	9	12	35	22	3	38	1	2

★ Consider the Method of Moments estimator for β based on the distribution $X \sim \text{Gamma}\left(\frac{1}{2}, \hat{\beta}\right)$.

Problem 1 Part a) On a single figure, plot both the empirical cumulative distribution function for the monsoon rainfall data above and the appropriate gamma distribution function with $\alpha = \frac{1}{2}$ and $\beta = \hat{\beta}$.

```r
# NOTE: The Data and estimator from GA16

x <- c(3, 15, 1, 36, 5, 1, 8, 11, 6, 9, 12, 35, 22, 3, 38, 1, 2)

xbar <- mean(x)
```

```r
beta_hat <- 1/(2 * xbar)

# NOTE: The first plot is of the empirical cumulative distribution function

plot(sort(x), (1:length(x))/length(x), type = "s", xlim = c(0, 40), ylim = c(0, 1),
    col = "blue", xlab = "x", ylab = "Cumulative Probability")

# NOTE: FILL IN the necessary command for overlaying plots

par(new = TRUE)

# NOTE: The second plot is of the true cumulative distribution function for
# Gamma(0.5, beta_hat)

curve(pgamma(x, 0.5, beta_hat), xlab = "", ylab = "", from = 0, to = 40, ylim = c(0,
    1))

# NOTE: Space for adding to the plot for b)

P_g25 = pgamma(25, 0.5, beta_hat)

segments(25, P_g25, 25, 1, col = "red", lwd = 4)

# NOTE: This is what I used to label the indication

text(21, 0.95, "P{X>25} ->")
```



Problem 1 Part b) Use $\hat{\beta}$ and the appropriate gamma distribution commands in `R` to estimate the probability that a monsoon rain exceeds 25 mm. Indicate this value on the plot from part (a).

$1 - pgamma(25, 0.5, beta_{hat}) = 0.1528808$

Hint: Use the gamma distribution family command with the correct prefix in `R` to estimate the probability that a monsoon rain exceeds 25 mm. If computed correctly, $P(X > 25) \approx 0.153$.

Hint: To indicate the value on the plot, you can 1) create new copy of the plot below, 2) add to the plot you created above, or 3) annotating on the plot and uploading an image as shown in part d.

NOTE: To do this task, enter your code in the R chunk below by filling in the blanks denoted with `FILL IN` *and uncommenting all non-NOTE lines.*

```r
# NOTE: calculate P(X > 25) where X ~ Gamma(0.5, beta_hat)
1 - pgamma(25, 0.5, beta_hat)
```

```
## [1] 0.1528808
```

Problem 1 Part c) If $\mu > 0$, confirm that the estimator $\hat{\beta}$ is biased upward. Why is it reasonable to assume that $\mu > 0$?

Hint: Recall that bias can be analyzed via concavity of the estimator. You will need to show that $\beta(\mu)$ is concave up. Any estimator that is a concave up function, i.e. convex, is known to be biased upward .

1c) $\hat{\beta} = \frac{1}{2\bar{x}} = \frac{1}{2}\bar{x}^{-1}$

$\hat{\beta}' = -\frac{1}{2}(\bar{x})^{-2}$

$\hat{\beta}'' = (\bar{x})^{-3} = \frac{1}{\bar{x}}$

second derivative of estimator is positive which means estimator is concave up function which means estimator is biased upwards

it's reasonable to assume that $\mu > 0$ because you can't have negative rainfall (although evaporation would be something like that), also you can assume it would rain at least once during the monsoon season

Problem 1 Part d) Use the delta method to estimate the variance in the estimator using the value obtained for $\hat{\beta}$ for the $n = 17$ monsoon rainfall amounts.

1d)

$$\sigma^2_{\hat{\beta}(\bar{x})} = \left(\frac{d}{dx}\left(\frac{1}{2x}\right)\right)^2 \frac{\sigma_x^2}{n_x}$$

$$= \left(-\frac{1}{2}(\bar{x})^{-2}\right)^2 \frac{\frac{1}{2(0.04086538)^2}}{17}$$

$$= \left(-\frac{1}{2}(12.23529)^{-2}\right)^2 \frac{\frac{1}{2(0.04086538)^2}}{17}$$

$$\boxed{\sigma^2_{\hat{\beta}(\bar{x})} = 1.9647 \times 10^{-4}}$$

$$\sigma_x^2 = \frac{d}{\beta^2}$$

$$\sigma = \sqrt{\frac{d}{\beta^2}}$$

$$\sigma_x^2 = \frac{1}{2\beta^2}$$

$$n_x = 17$$

Problem 2.

Recall from Group Activity 17

For a parameter $\theta > 0$, we model the accuracy of a dart player by the $Beta(\theta, 1)$ density

$$f_X(x|\theta) = \begin{cases} 0 & \text{if } x < 0, \\ \theta x^{\theta-1} & \text{if } 0 \leq x < 1, \\ 0 & \text{if } 1 \leq x, \end{cases}$$

for a continuous random variable X, the distance the dart is from the center of the board.

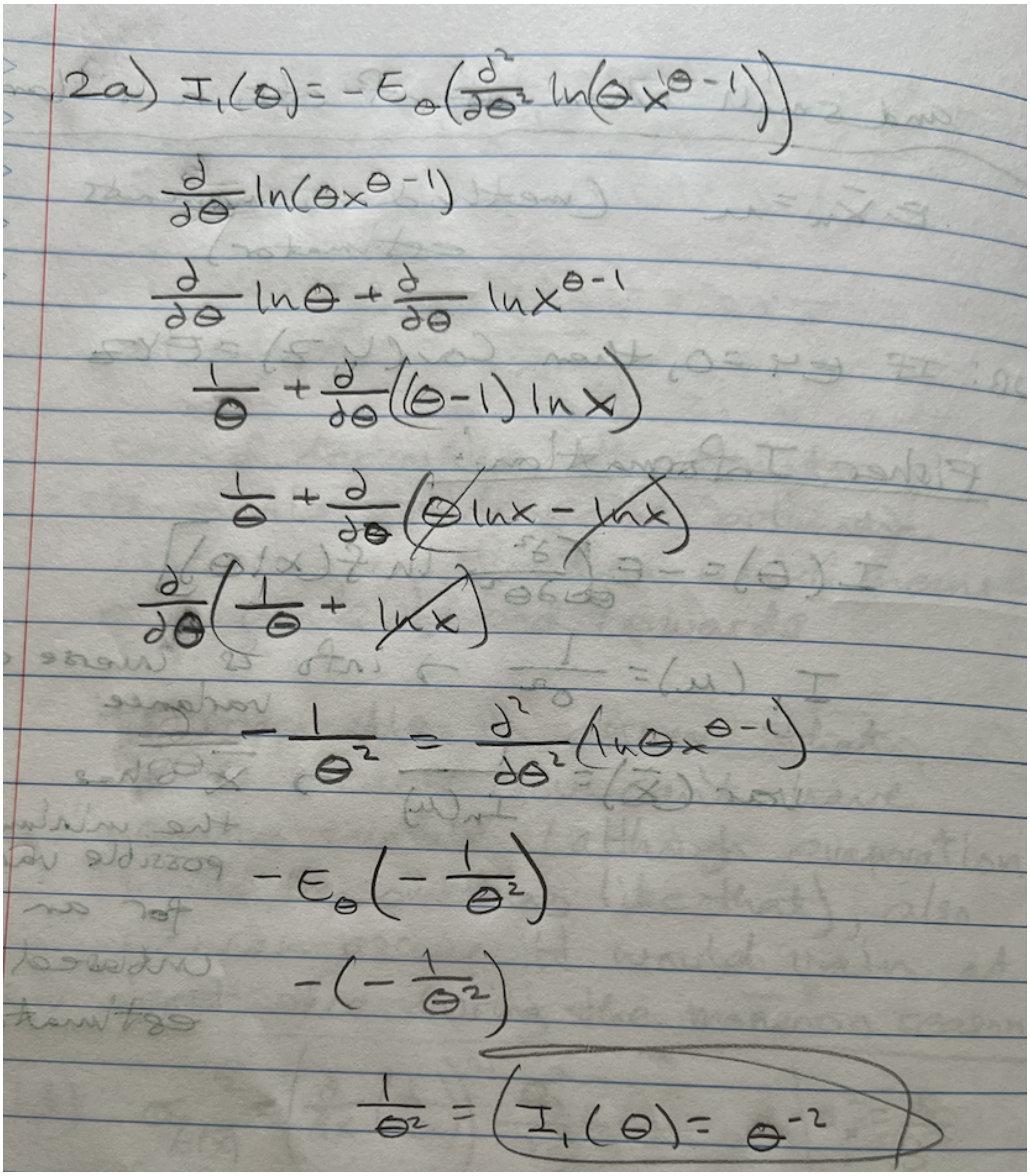
(Do Not Repeat Solutions) For n observations, you already found the likelihood function.

(Do Not Repeat Solutions) You already found the maximum likelihood estimate for observations $x_1, \ldots, x_n$.

★ Consider the maximum likelihood estimates of $\hat{\theta}$ based on the values $\theta = 1/4$.

Problem 2 Part a) In this context, verify that the Fisher Information $I_1(\theta) = \theta^{-2}$.

Hint: Recall that the Fisher Information $I_1(\theta) = -E_\theta \left[\frac{\partial^2}{\partial \theta^2} \ln(f_X(x|\theta)) \right]$. Note that the subscript on the Fisher Information $I_1(\theta)$ implies that $n = 1$.



Problem 2 Part b) Compute the standard deviation estimate $\text{SD}(\hat{\theta})$ given by the Fisher Information for $\theta = 1/4$ and $n = 50, 500,$ and 5000.

Hint: Recall $\text{Var}(\hat{\theta}) \geq \frac{1}{n \cdot I_1(\theta)}$ and $I_1(\theta) = \theta^{-2}$. You will need to compute the lower bound for $\text{SD}(\hat{\theta}|n = 50)$, $\text{SD}(\hat{\theta}|n = 500)$, and $\text{SD}(\hat{\theta}|n = 5000)$ where $\text{SD}(\hat{\theta}) \geq \frac{1}{\sqrt{n \cdot I_1(\theta)}}$.

2b) $SD(\hat{\theta} \mid n=50) \geq \dfrac{1}{\sqrt{50 \cdot 1/(1/4)^2}}$

$\geq \boxed{0.035355}$

$SD(\hat{\theta} \mid n=500) \geq \boxed{0.011180}$

$SD(\hat{\theta} \mid n=5000) \geq \boxed{0.0035355}$

Problem 2 Part c) Below are 10000 simulations of Maximum Likelihood Estimates of $\hat{\theta}$ based on the values $\theta = 1/4$ and $n = 50, 500, 5000$.

NOTE: The R-code for these simulations is in a separate RMD file titled 'WS17_Simulation_Code' *since the simulations cause knitting to PDF to take a while.*

Based upon the distributions of $\hat{\theta}$, discuss whether the Asymptotic Properties of Maximum Likelihood Estimators are met, i.e.,

1. **Consistency**: As the number of observations n increase, the distribution of the MLE $\hat{\theta}$ becomes more and more concentrated about $\theta = 1/4$, the true state of nature.
2. **Asymptotic normality**: As the number of observations n increase, the distribution of the MLE $\hat{\theta}$ becomes asymptotically normal.
3. **Asymptotic efficiency**: As the number of observations n increase, the variance of the distribution of the MLE $\hat{\theta}$ approaches $\frac{1}{n \cdot I_1(\theta)}$.

NOTE: Enter your response as text below.

1. Consistency is met because we can see as we increase in observations n, the graph distribution becomes more peaked about the true value of 1/4, meaning smaller standard deviation and more values more concentrated around the true value. 2. Asymptotic normality is met becuase again we can see from the graphs that as n increases, the graph appearance approaches to a normal distribution (n=50 was more right skewed and progressively evens out with larger n) 3. Asymptotic efficiency is met because based on the standard deviations calculated in part 2b and the stds given in 2c, we can see that the stds are closer in value to each other as the number of observations n increases which implies that their variances become closer in value to each other as well, although the differences are more prounounced due to squaring for the variance

Problem 2 Part d) Below are 10000 simulations of the Method of Moments estimates of $\hat{\theta}$ based on the values $\theta = 1/4$ and $n = 50, 500, 5000$.

NOTE: The R-code for these simulations is in a separate RMD file titled 'WS17_Simulation_Code' *since the simulations cause knitting to PDF to take a while.*

Compare the Maximum Likelihood Estimates simulations used in part (c) with the equivalent Method of Moments simulations above. How do these results explain why Maximum Likelihood Estimation is preferred to Method of Moments for parameter estimation?

Hint: Think about the standard deviations, but don't forget to address the means as well.

NOTE: Enter your response as text below.

We can compare the standard deviations and means to see that the maximum likelihood estimator is much more accurate and precise than the method of moments estimation due to closer averages to the true value and lower standard deviations for each n observations