

(From Session 17 Group Quiz)

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, \dots, x_n .

o1-worksheet Maximum Likelihood Estimation

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

- a. Find the formula for the standard deviation of maximum likelihood estimate of $\hat{\theta}$ for observations x_1, \dots, x_n . Then compute the standard deviation estimate given by the Fisher information for $n = 50, 500, 5000$. What do these estimates mean? (Recall that $SD(\hat{\theta}) \approx 1/nI_1(\theta)$ where the Fisher information can be computed via the second derivative formula.)

These values of standard deviations are the lowest possible values that we can theoretically get with our estimators. Since this depends on our sample size n , we see that the higher our sample size, the lower the value of the lower bound becomes.

$$(a) I_1(\theta) = -E_\theta \left(\frac{\partial^2}{\partial \theta^2} \ln f_\theta(\theta | \mathbf{x}) \right)$$

$$= -E_\theta \left(-\frac{n}{\theta^2} \right) \quad \text{Since we want } I_{(1)}, n \text{ here} = 1$$

$$= -E_\theta \left(-\frac{1}{\theta^2} \right) = \boxed{\frac{1}{\theta^2}}$$

$$\text{Var}(\hat{\theta}) \geq \frac{1}{n I_1(\theta)} = \frac{1}{n} \cdot \theta^2 = \boxed{\frac{\theta^2}{n}}$$

$$SD(\hat{\theta}) \geq \sqrt{\theta^2/n}$$

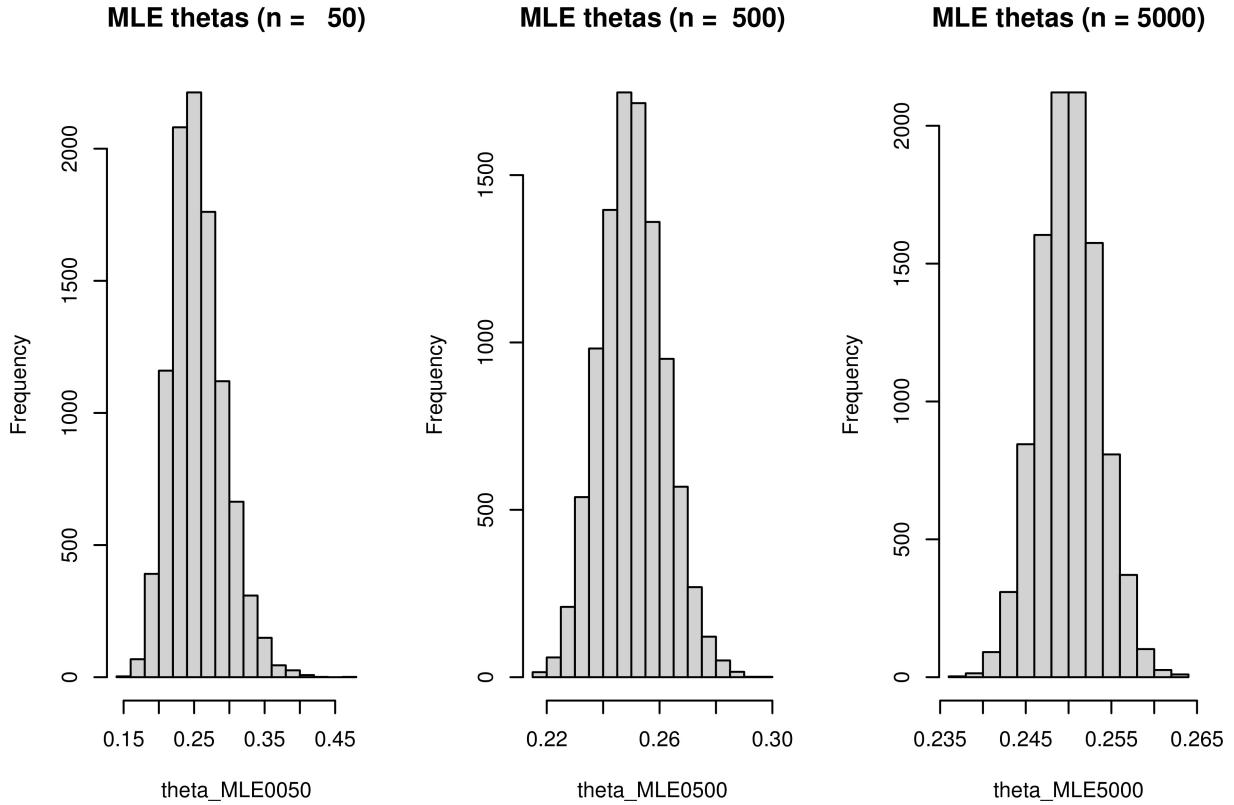
$$SD(\hat{\theta} | 50) \approx \sqrt{\frac{(1/4)^2}{50}} \approx 0.0354$$

$$SD(\hat{\theta} | 500) \approx \sqrt{\frac{(1/4)^2}{500}} \approx 0.0112$$

$$SD(\hat{\theta} | 5000) \approx \sqrt{\frac{(1/4)^2}{5000}} \approx 0.00354$$

Figure 1: Part(a)

★ 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 the RMD file.)

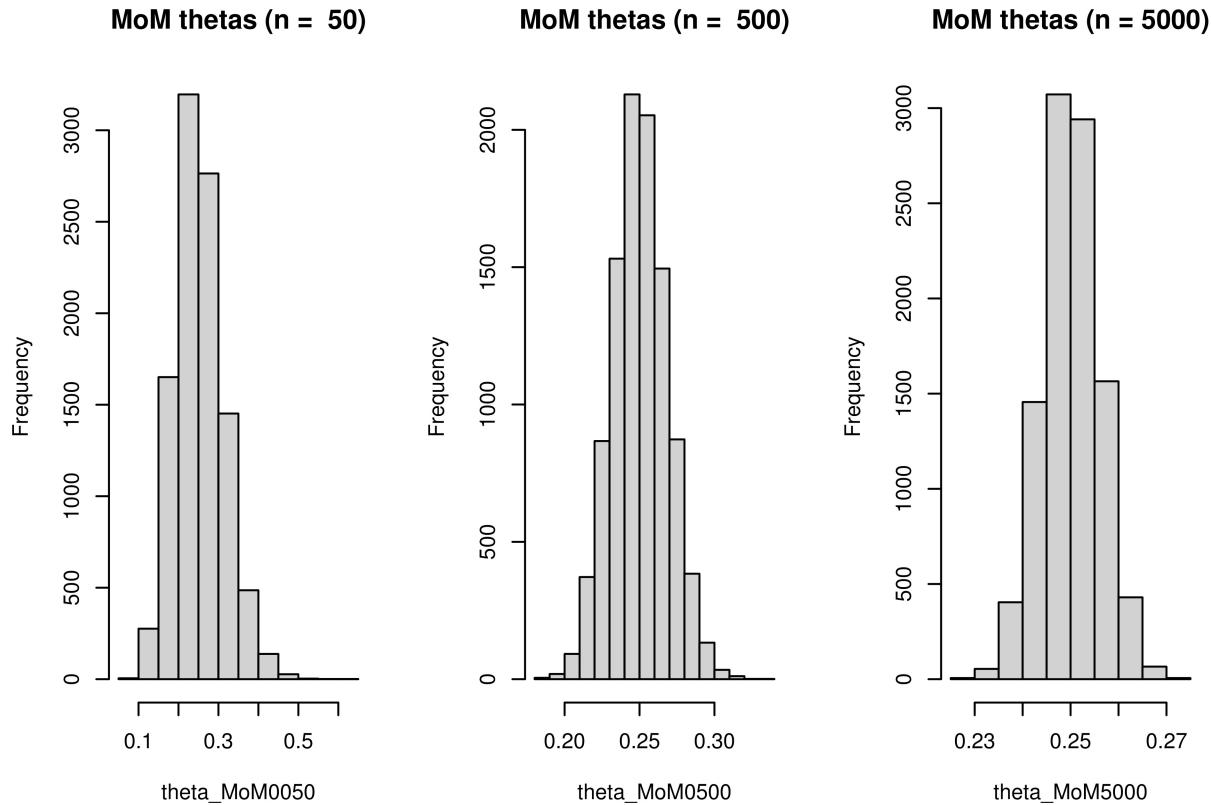


```
##           Mean      StandDev
## MLE n=50  0.2552448 0.037056025
## MLE n=500 0.2504668 0.011275141
## MLE n=5000 0.2500562 0.003574606
```

- b. Discuss what happens to the bias, standard deviation, and shape of the distributions of $\hat{\theta}$ s as n gets large in the context of the Asymptotic Properties of Maximum Likelihood Estimators.

We see that as the value of n increases, the shape of the graphs become more and more normal. Moreover, the skewness of the graph is also decreasing. Now since the skewness of the graph is decreasing, we know that the bias of the graph will also decrease. In terms of the Asymptotic properties: the graph is CONSISTENT since as the value of n grows larger, the value of the mean converges to value of the population mean, i.e., $1/4$. We have Asymptotic NORMALITY since the graph is growing less and less skewed, meaning the value of bias is also reducing. Finally, we also have asymptotic EFFICIENCY since the spread of the data, that is the standard deviation is growing smaller and smaller.

★ 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 the RMD file.)



```
##           Mean      StandDev
## MoM n=50   0.2529691  0.060905703
## MoM n=500  0.2502870  0.018516955
## MoM n=5000 0.2500688  0.005980134
```

- c. Compare the Maximum Likelihood Estimates simulations used in part (b) 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?

We can see that as the value of n increases, both the simulations show a decrease in the skewness of the distribution. Both of them are a bit skewed initially, but as the value of n increases, this skewness is seen to reduce significantly. Most importantly however, these simulations differ in terms of the values of standard deviations for different values of n . We can see that with the Maximum Likelihood Estimates, the values of the standard deviations are much lower compared to the values we got in the Method of moments simulation. Therefore, since the value of spread is less in the MLE simulation, this is a much more powerful estimator as compared to the MoM estimator.