Simulation of emergency room wait times: a comparison of Weibull shape and scale estimators

Kenny Chen, Nia Martinez Jaime, Adeli Hutton

November 29, 2022

1 Motivation

In January of this year, hospitals in Massachusetts saw a surge in emergency room (ER) visits due to Covid-19 symptoms. The exhaustion of ER doctors and nurses, staffing levels at 70%, and the limited number of beds further contributed to overflowing ER waiting rooms.[1] Responsive urgent care is crucial for the health and safety of patients, and knowing the expected wait time to get a room provides valuable information to both health-care workers and to the patients themselves. Knowing approximate ER wait times will better inform hospital administration and patient home-care to prevent the need to go to the ER. Thus, the goals of this project are to simulate the wait time for an ER room and to evaluate and compare properties of different estimators to see which one might help us best model wait times.

2 Methods

Because we are interested in modeling how much time a patient must wait in the ER until the patient gets assigned a room, we chose to look at the Weibull distribution. It can accurately model failure times while allowing more flexibility by incorporating two parameters, a shape parameter γ and a scale parameter λ. The pdf of the Weibull is as follows:

$$f(x) = \begin{cases} \frac{\gamma}{\lambda}(\frac{x}{\lambda})^{\gamma-1}e^{-(\frac{x}{\lambda})^{\gamma}}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1}$$

where $\gamma > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter.

The Weibull distribution is often used in these scenarios and in survival analyses more generally since it considers the amount of time waited before an event occurs. When $\gamma > 1$, the longer the wait time, the higher the chance of getting a room, which emulates waiting in the ER. When $\gamma = 1$, the Weibull distribution represents the exponential distribution, which is memoryless, and does not consider the time one has already waited for the event to occur. When $\gamma < 1$, the longer a patient waits, the less likely the patient gets a room, which does not make sense in our scenario. Therefore, we will only consider the Weibull distribution when $\gamma \geq 1$.

To estimate the shape and scale parameters, we consider the following three estimators (equations for the estimators can be found in the Appendix):

1. Maximum Likelihood Estimation (MLE)
2. Method of Moments Estimator (MME)
3. Method of Moments Estimator Based on Unbiased Estimator of Variance (MMUE)

We considered these three functions since the MLE and the MME are estimators we covered extensively in class, and we wanted to extend them to a new distribution we have not practiced this semester. Selecting these estimators also allows us to utilize the function `eweibull` from the `EnvStats` package in R, which calculates the MLE, MME, and MMUE for both the shape and scale parameter of our data.[2]

3 Metrics

We compare the three estimators based on their bias and Mean Squared Error (MSE) which are defined respectively, where $\hat{\theta}$ is an estimator for θ:

$$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta \quad \text{and} \quad MSE(\hat{\theta}) = Bias(\hat{\theta})^2 + Var(\hat{\theta})$$

We measure mean bias and MSE as our metrics since we want to measure the tradeoff between those two metrics. Having a smaller bias / MSE is indicative of better estimation, so we want to see which one performs better under different simulation conditions to choose the best one.

4 Simulation Design

For our simulation setup, we use the function `rweibull` from the `stats` package in R to simulate the data under the various simulation conditions with a seed set for reproducibility. We generate samples of size $n = 10, 25, 50, 75, 100$ and consider two shape values, $\gamma = 1, 3$, and two scale values, $\lambda = 2, 4$, leading to $5 \times 2 \times 2 = 20$ simulation conditions. We repeat each simulation scenario 1000 times to calculate mean estimates, biases, and MSEs. We plot the mean biases and MSEs in the figures in Section 6 below and include a table of the mean estimates stratified by sample size, shape, scale, and estimator method.

5 Discussion

Figures 1 and 2 display the mean bias of the shape and scale parameters respectively under the various simulation conditions. We observe that the bias of both parameters differs depending on the magnitude of the shape. Our results suggest that higher shape parameters are associated with higher variability in the mean bias of γ with changes in sample sizes. Also, the MMUE achieves lower mean bias with higher shape parameters. However, when analyzing the bias of λ, higher shape parameters were associated with negative mean bias and lower shape parameters with a positive one.

Figures 3 and 4 display the mean MSE of the shape and scale parameters respectively under the various simulation conditions. Similar to the analysis above, the mean MSE of both parameters also differ depending on the magnitude of the shape. For the MSE of γ, lower shape parameters are associated with low MSE regardless of sample size, while higher shape parameters are associated with high MSE for small sample sizes. However, we observed the opposite for the MSE of λ. Here, lower shape sizes are associated with higher mean MSE for smaller sample sizes when compared to higher shape sizes, which have small mean MSE regardless of sample size.

From the figures, we observed a common trend of decreasing mean bias and MSE as the sample size increased, which points to the importance of having an adequate sample size. Table 1, which shows the mean shape and scale estimates also demonstrates this pattern, as the sample size increases, the mean estimates grow closer to the true parameters. Overall, we saw more variability in the mean bias and MSE of shape estimators when sample size was small (n=10 or 25) and shape = 3. However, the changes from 25 to 50 or 50 to 75 were not as dramatic, which suggests that a sample size at around 50 might provide decent estimates for all three methods considered and at $n = 50$, all three methods had similar metrics for both shape and scale estimators. We also observed that the MMUE, on average, achieves the smallest mean bias and MSE than our other two estimators, which suggests that the MMUE is the most appropriate estimator to use, especially when sample size is small. Due to the constraints of this project, we only considered a few values for the shape and scale parameter, so further work would involve looking at higher values for these parameters to see how these estimators act.

6 Results

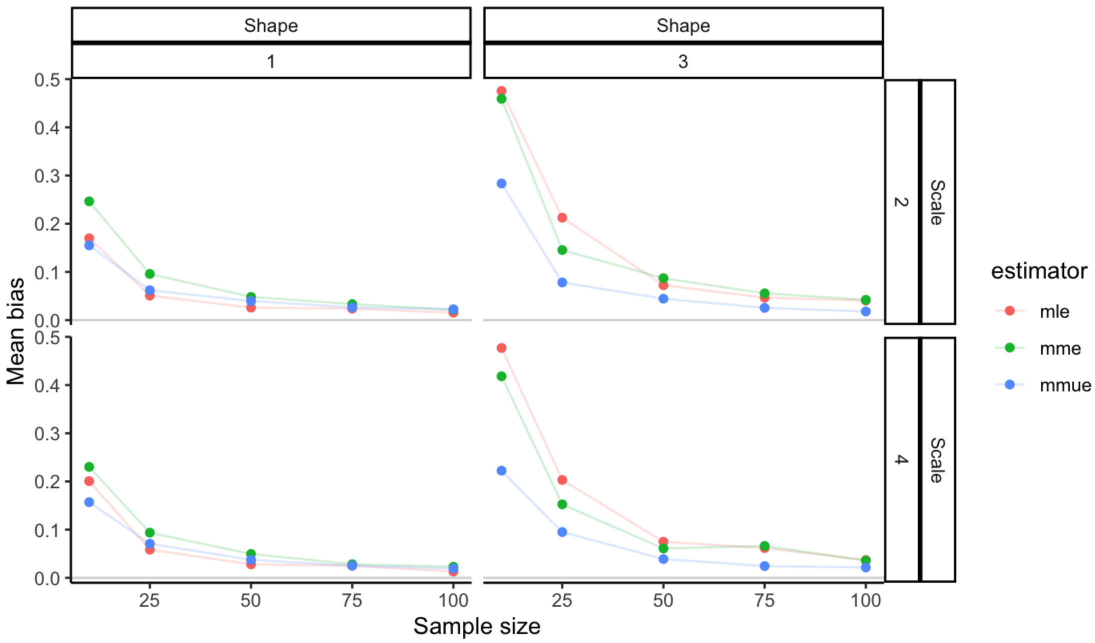


Figure 1. Mean Bias of Shape Estimators Faceted by Shape and Scale Values

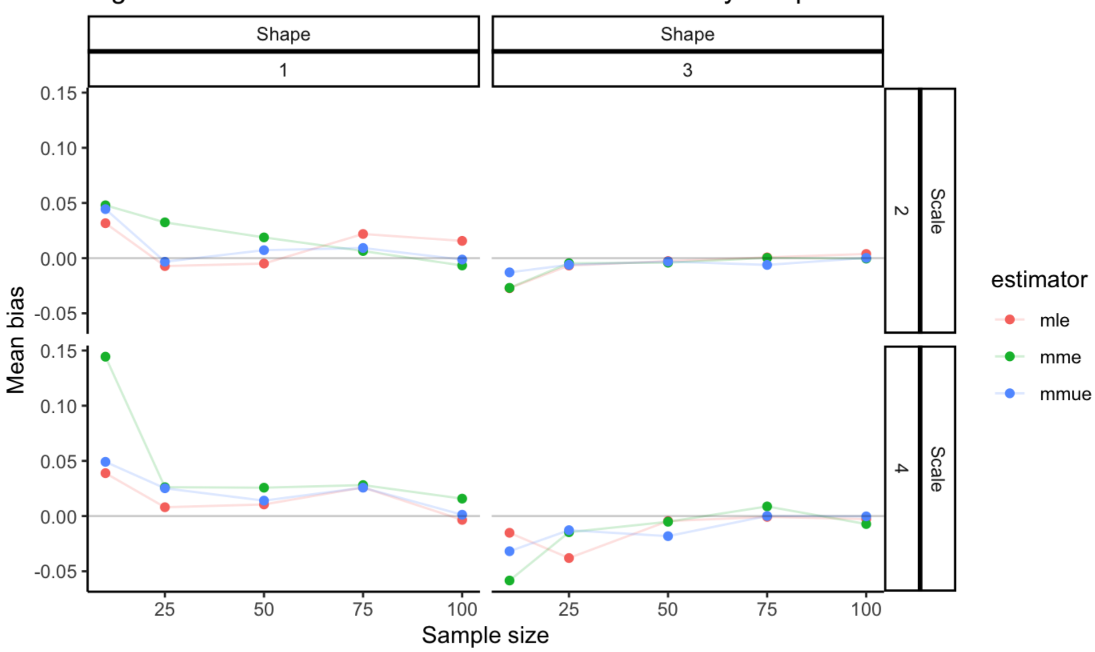


Figure 2. Mean Bias of Scale Estimators Faceted by Shape and Scale Values

Figure 3. Mean MSE of Shape Estimators Faceted by Shape and Scale Values

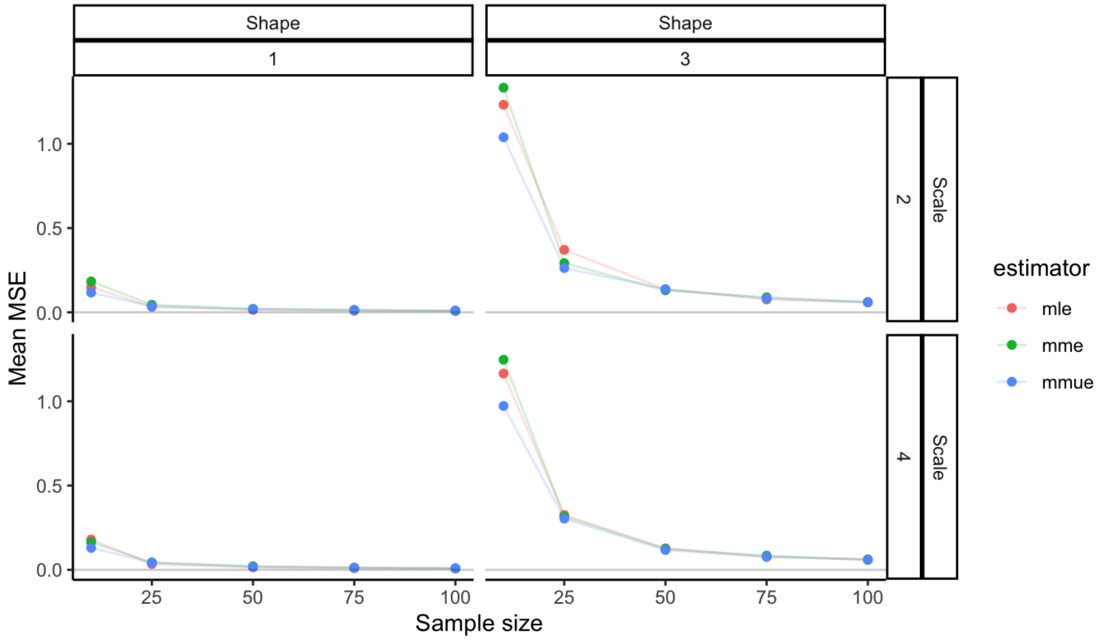
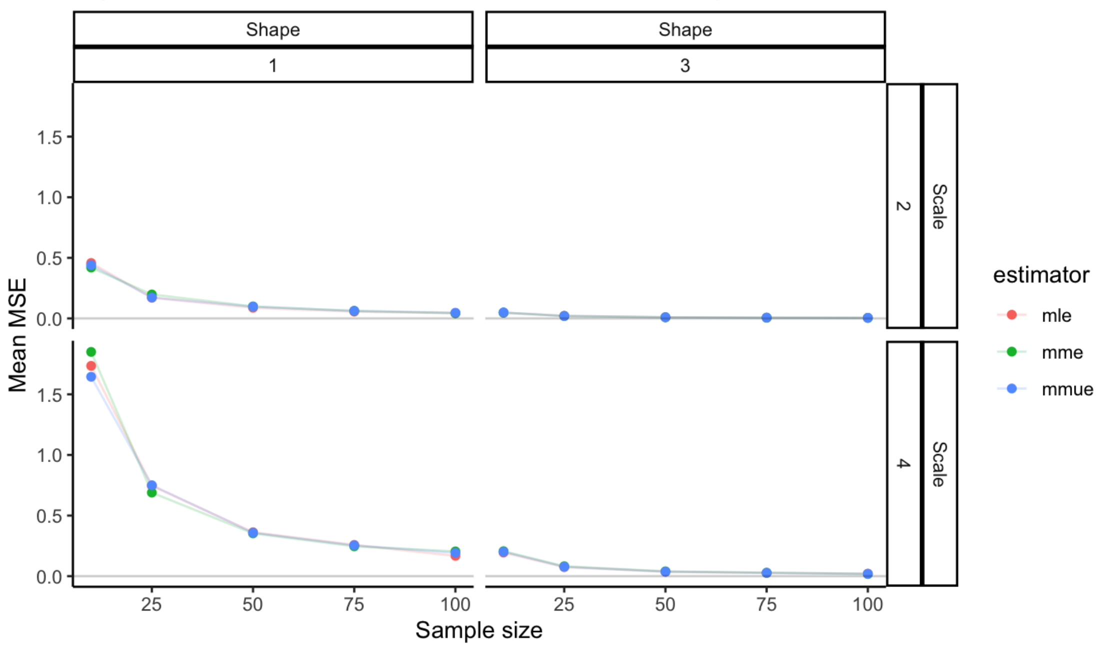


Figure 4. Mean MSE of Scale Estimators Faceted by Shape and Scale Values



	Shape	Scale	Estimator	n=10	n=25	n=50	n=75	n=100
1	1.00	2.00	mle	[1.17, 2.03]	[1.05, 1.99]	[1.03, 2]	[1.02, 2.02]	[1.02, 2.02]
2	1.00	2.00	mme	[1.25, 2.05]	[1.1, 2.03]	[1.05, 2.02]	[1.03, 2.01]	[1.02, 1.99]
3	1.00	2.00	mmue	[1.16, 2.04]	[1.06, 2]	[1.04, 2.01]	[1.03, 2.01]	[1.02, 2]
4	1.00	4.00	mle	[1.2, 4.04]	[1.06, 4.01]	[1.03, 4.01]	[1.03, 4.03]	[1.01, 4]
5	1.00	4.00	mme	[1.23, 4.14]	[1.09, 4.03]	[1.05, 4.03]	[1.03, 4.03]	[1.02, 4.02]
6	1.00	4.00	mmue	[1.16, 4.05]	[1.07, 4.03]	[1.04, 4.01]	[1.03, 4.03]	[1.02, 4]
7	3.00	2.00	mle	[3.48, 1.97]	[3.21, 1.99]	[3.07, 2]	[3.05, 2]	[3.04, 2]
8	3.00	2.00	mme	[3.46, 1.97]	[3.15, 2]	[3.09, 2]	[3.06, 2]	[3.04, 2]
9	3.00	2.00	mmue	[3.28, 1.99]	[3.08, 1.99]	[3.04, 2]	[3.03, 1.99]	[3.02, 2]
10	3.00	4.00	mle	[3.48, 3.98]	[3.2, 3.96]	[3.07, 4]	[3.06, 4]	[3.04, 4]
11	3.00	4.00	mme	[3.42, 3.94]	[3.15, 3.99]	[3.06, 3.99]	[3.07, 4.01]	[3.04, 3.99]
12	3.00	4.00	mmue	[3.22, 3.97]	[3.09, 3.99]	[3.04, 3.98]	[3.02, 4]	[3.02, 4]

Table 1: Mean estimators after 1000 simulations under different simulation scenarios. First value is the mean shape estimator followed by the mean scale estimator.

7 Appendix

Equations of the three estimators of γ and λ.[3,4]

1. MLE Estimators for γ and λ:

$$\hat{\gamma}_{MLE} = \frac{n}{(\frac{1}{\hat{\lambda}_{MLE}})^{\hat{\gamma}_{MLE}} \sum_{i=1}^{n} (x_i^{\hat{\gamma}_{MLE}} \log(x)) - \sum_{i=1}^{n} log(x_i)} \text{ and } \hat{\lambda}_{MLE} = (\frac{1}{n} \sum_{i=1}^{n} x_i^{\hat{\gamma}_{MLE}})^{1/\hat{\gamma}_{MLE}}$$

2. MME Estimators for γ and λ:

$$\hat{\gamma}_{MME} = \frac{s}{\bar{x}} = \{\frac{\Gamma[(\hat{\gamma}_{MME}+2)/\hat{\gamma}_{MME}]}{\{\Gamma[(\hat{\gamma}_{MME}+1)/\hat{\gamma}_{MME}]\}^2} - 1\}^{1/2} \text{ and } \hat{\lambda}_{MME} = \frac{\bar{x}}{\Gamma[(\hat{\gamma}_{MME}+1)/\hat{\gamma}_{MME}]}$$

   where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x$ is the sample mean and $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ and Γ is the gamma function.

3. MMUE Estimators for γ and λ: Same as above but with an unbiased estimator for the variance: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

8 Citations

1. "Massachusetts Emergency Room Doctors, Nurses Say They're Overwhelmed". CBS Boston, Jan 4, 2022. https://www.cbsnews.com/boston/news/massachusetts-emergency-room-doctors-nurses-overwhelmed-covid-19-surge-omicron-variant/

2. Millard SP (2013). EnvStats: An R Package for Environmental Statistics. Springer, New York. ISBN 978-1-4614-8455-4, https://www.springer.com.

3. Forbes, C., M. Evans, N. Hastings, and B. Peacock. (2011). Statistical Distributions. Fourth Edition. John Wiley and Sons, Hoboken, NJ.

4. Johnson, N. L., S. Kotz, and N. Balakrishnan. (1994). Continuous Univariate Distributions, Volume 1. Second Edition. John Wiley and Sons, New York.