# The Exponential Distribution

## Estimators and Confidence Intervals

Abhijith Asok, Chris Hilger, Liam Loscalzo, Katherine Wang

# Point estimator selection and comparison

We compared the point estimators listed on the following slide by looking at plots of mean squared error (MSE) and estimator properties like consistency, bias, and asymptotic bias.

The first three estimators were chosen before looking at the MSE plot. We know that the MLE and unbiased correction to the MLE have some desirable characteristics. We also chose to investigate a method of moments estimator. Since the first moment $E[X] = \frac{n}{\sum x_i}$ is the same as the MLE, we looked at the second moment instead: $E[X^2] = 2\bar{x}^2$

The last two estimators were developed in an attempt to achieve a lower MSE for different values of $\lambda$ and $n$ ($4 < \lambda < 6$ and $n = 10, 25$, respectively)

## Exponential parameterization

$$f(x, \lambda) = \lambda e^{-\lambda x}$$

# Classifying the estimators

We calculate values for each of the estimators to construct the table below:

| Values for each estimator | | | | | |
|---|---|---|---|---|---|
| Estimator | Value | $E[\lambda]$ | $Var[\lambda]$ | $Bias[\lambda]$ | $MSE[\lambda]$ |
| $\hat{\lambda}_0$ | $\frac{n}{\sum x_i}$ | $\lambda\frac{n}{n-1}$ | $\frac{\lambda^2}{n}$ | $\lambda\frac{n}{n-1}-\lambda$ | $(\lambda\frac{n}{n-1}-\lambda)^2+\frac{\lambda^2}{n}$ |
| $\hat{\lambda}_1$ | $\frac{n-1}{\sum x_i}$ | $\lambda$ | $\frac{\lambda^2}{n}\frac{n-1}{n}^2$ | $0$ | $\frac{\lambda^2}{n}\frac{n-1}{n}^2$ |
| $\hat{\lambda}_2$ | $2\bar{x}^2$ | $\frac{2}{\lambda^2}$ | $\frac{16}{n\lambda^4}$ | $\frac{2}{\lambda^2}-\lambda$ | $\left(\frac{2}{\lambda^2}-\lambda\right)^2-\frac{16}{n\lambda^4}$ |
| $\hat{\lambda}_3$ | $5$ | $5$ | $0$ | $5-\lambda$ | $(5-\lambda)^2$ |
| $\hat{\lambda}_4$ | $\frac{1}{\sum x_i}$ | $\frac{\lambda}{n-1}$ | $\frac{\lambda^2}{n^3}$ | $\frac{\lambda}{n-1}-\lambda$ | $(\frac{\lambda}{n-1}-\lambda)^2+\frac{\lambda^2}{n^3}$ |

# Maximum likelihood estimator: $\hat{\lambda}_0$

Given how we parameterized the exponential distribution, we cannot easily take the expected value of $\hat{\lambda}_0$ since $\sum x_i$ is in the denominator. However, we recognize that $\sum x_i$ for an exponential variable is gamma-distributed. Below we show how we use the gamma to calculate the expectation and use similar methods to compute the values shown in the slide above.

## Expectation calculation

Let $y = \sum x_i$ and $E[\hat{\lambda}_0] = E\left[\frac{n}{\sum x_i}\right] = E[\frac{n}{y}]$

So:

$$E\left[\frac{n}{y}\right] = \int_0^\infty \frac{1}{(n-1)!\lambda^n} x^{(n-1)} e^{\left(\frac{-x}{\lambda}\right)} dy$$

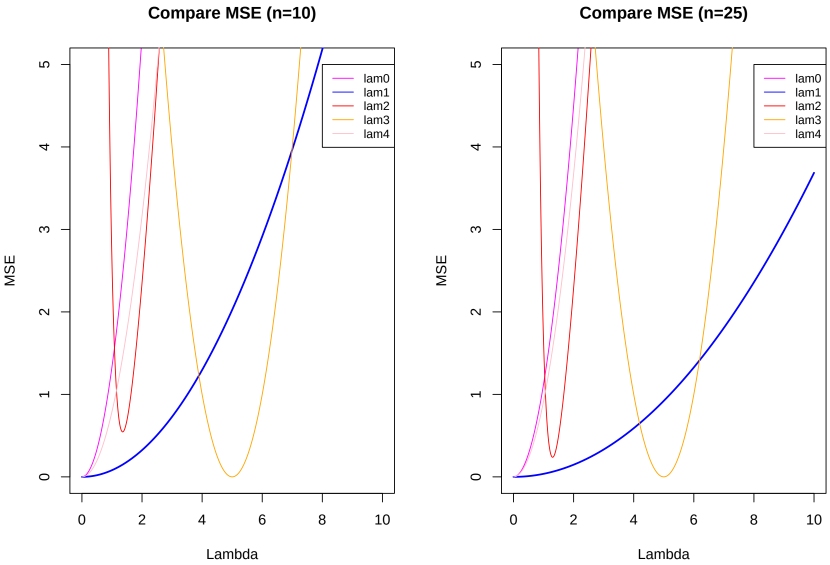
Computing this integral yields:

$$E[\hat{\lambda}_0] = \lambda \left(\frac{n}{n-1}\right)$$

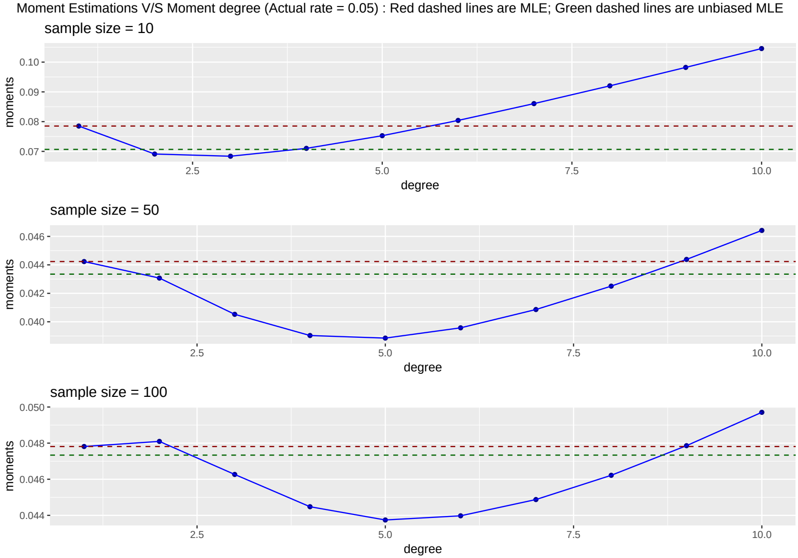
# Plotting the MSE

- Now that we have the MSE calculations for all estimators, we can plot them for various $n$ over a range of $\lambda$ values. Specifically, we look at $n = 10$ and $n = 25$.
- Looking at the MSE for $\hat{\lambda}_1$, which is the unbiased version of the MLE, we find that the MSE is lowest compared to all other estimates, except for that of $\hat{\lambda}_3$ which equals 5.
- As discussed earlier, the last two estimators $\hat{\lambda}_3$ and $\hat{\lambda}_4$ were developed in an attempt to achieve a lower MSE for a given range of $\lambda$, specifically $4 < \lambda < 6$ at $n = 10, 25$.
- We find that $\hat{\lambda}_3$ achieves this goal nicely, but $\hat{\lambda}_4$ does not.
- Plotting the estimators for n $=$ 10 and n $=$ 25, we see that the hard $\lambda = 5$ estimate would be less useful because $\hat{\lambda}_1$, which is the unbiased version of the MLE, has an MSE which is much closer to it than at the $n = 10$ level.

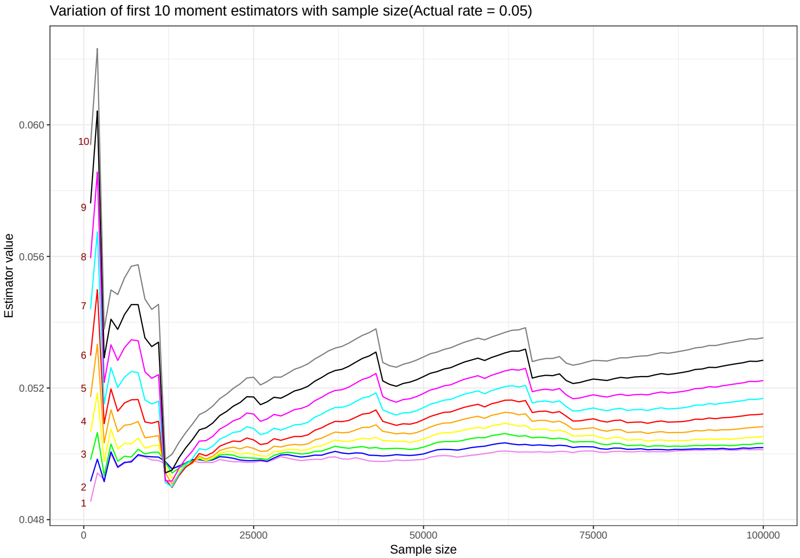
# Plotting the MSE



# For small samples, MoM estimates decrease, then increase with increasing degree



# For a range of sample sizes, the first 10 MoM estimators follow similar patterns



# Confidence intervals: Methods

We compared the following confidence intervals:

- Wald-based confidence interval
- Gamma-based confidence interval
- Score-based confidence interval
- Bootstrap confidence interval

## Comparing confidence intervals

To compare confidence intervals, we are interested in understanding both

- The coverage probability of each confidence interval at a given $\alpha$ and $\lambda$ value, evaluated at different sample sizes
- How large the confidence interval is for a given $\alpha$ level, across different sample sizes and values of $\lambda$

A good confidence interval should be as small as possible while providing a coverage probability corresponding to the chosen $\alpha$ level. We perform our assessments using simulated data.

# Wald-based confidence interval

The "standard" confidence interval based on the CLT, which says that as $n$ gets large, the Wald test statistic

$$T = \frac{\lambda - \lambda_0}{\text{s.e.}(\lambda)} \overset{\text{app}}{\sim} N(0, 1)$$

We reject the null hypothesis when $|T| \geq z_{1-\alpha/2}$. This can be inverted to obtain the $(1-\alpha) \cdot 100\%$ confidence interval

$$\hat{\lambda} \pm z_{1-\alpha/2} \cdot \text{s.e.}(\lambda)$$

# Gamma-based confidence interval

## Relation of exponential distribution to gamma distribution

If $X_1, X_2, ..., X_n \overset{\text{iid}}{\sim} \text{Exponential}(\lambda)$, then

$$\sum_{i}^{n} x_i \sim \text{Gamma}(n, \lambda) \implies \lambda \bar{x} \sim \text{Gamma}(n, n)$$

Let $g_y$ be the $y$th percentile of this distribution. Then we can say:

$$1 - \alpha = P(g_{\alpha/2} \leq \lambda \bar{x} \leq g_{1-\alpha/2}) = P(g_{\alpha/2}/\bar{x} \leq \lambda \leq g_{1-\alpha/2}/\bar{x})$$

Therefore, a $(1 - \alpha) \cdot 100\%$ confidence interval for $\lambda$ is

$$(g_{\alpha/2}/\bar{x},\, g_{1-\alpha/2}/\bar{x})$$

Which can be trivially converted to a confidence interval for the mean $\mu = 1/\lambda$

# Score-based confidence interval

For moderate to large sample sizes, the following score test statistic is approximately normally distributed

$$\frac{U(\lambda_o)}{\sqrt{\mathrm{Var}[U(\lambda_o)]}} \overset{\text{app}}{\sim} N(0,1)$$

$(1-\alpha)\cdot 100\%$ confidence interval for the mean can be obtained by solving

$$\left|\frac{\frac{n}{\lambda_o}-\sum x_i}{\sqrt{n/\lambda_o^2}}\right| = \left|\sqrt{n}(1-\lambda_o\bar{x})\right| \geq z_{1-\alpha/2}$$

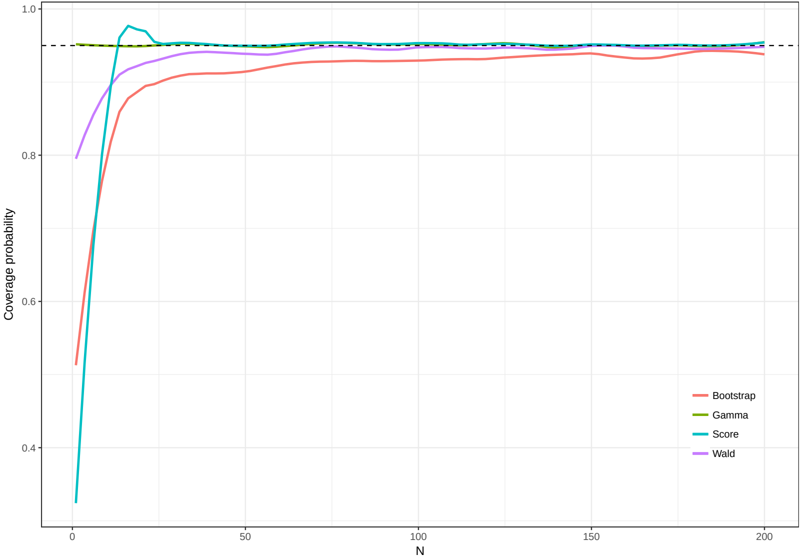
for $\lambda_o$, giving us a confidence interval of

$$\left(\frac{1}{\bar{x}}(1-\frac{1}{\sqrt{n}}z_{1-\alpha/2}), \frac{1}{\bar{x}}(1+\frac{1}{\sqrt{n}}z_{1-\alpha/2})\right)$$

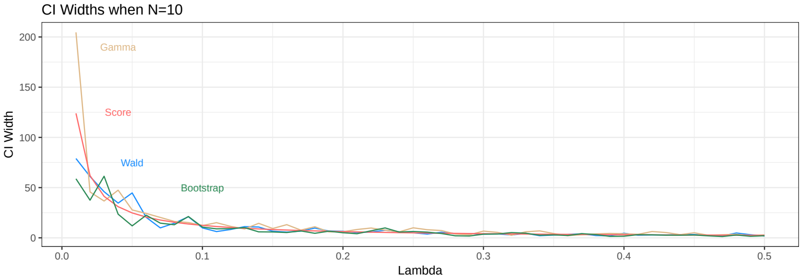
# Bootstrap confidence interval

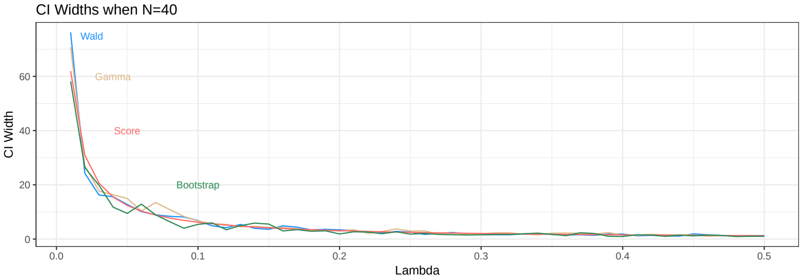
- Empirical method which does not require knowledge of underlying distribution for $X$
- Based on resampling data (**with replacement**) many times to create an empirical distribution $U^*$ which approximates the true (unknown) distribution $U$
- We estimate the variation of $\bar{x}$ around the true mean $1/\lambda$ using the variation of $\bar{x}^*$ in bootstrapped samples
- As an empirical method, depends on the original data
- We expect that $\bar{x}^*$ will approximate $\bar{x}$ well, but no guarantee it will be a good estimate of $\mu = 1/\lambda$
- Not a problem when we use simulated data

# Coverage probabilities: $\alpha = 0.05, \lambda = 1/12$



# Confidence Interval Widths: $\alpha = 0.05$



# Summary of Findings and Recommendations

## Estimators

- For small $\lambda$, all estimators seem to do well. For anything larger than $\lambda > 2\ \hat{\lambda}_1$, the unbiased MLE is the best.

## Confidence intervals

- Gamma-based confidence interval is exact and coverage probability aligns with chosen $\alpha$ for all $n$, but tends to be large for small $n$ and $\lambda$
- Only useful when the underlying distribution is known (not realistic)
- Wald- and score-based intervals perform similarly, with score providing better coverage at small $n$
- Bootstrap provides great flexibility, as does not require knowledge distribution underlying data
- "Slowest" to provide coverage corresponding to $\alpha$, but could possibly be improved using more bootstrap samples (restricted here for computational time)