Exponential Distribution

Abhijith Asok, Chris Hilger, Liam Loscalzo, Katherine Wang

Point Estimation: Methods

We compared the following point estimators:

- Maximum Likelihood Estimator
- Unbiased correction for the MLE
- Second Method of Moment Estimator

add criteria for comparing estimators!

Point Estimation Selection

We compared the point estimators listed below, by looking at plots of Mean Squared Error and attributes like Bias, Asymptotic Unbiasedness and Consistency.

The first three estimators were determined before looking at the MSE plot. Obviously, the MLE and Unbaised correction of the MLE would be decent estimates, we also considered the Method of Moments estimator. However, the first moment, $E[X] = \frac{n}{\sum x_i}$ is the same as the MLE. Therefore, we looked at the second moment instead. $E[X^2] = 2 * \bar{x}^2$

The last two estimators were developed in an attempt to achive a lower MSE for a given range of λ, and n, $4 < \lambda < 6$ and $n = 10, 25$ respectively.

The chosen estimators

Estimator list		
Symbol	Estimator	Value
$\hat{\lambda}_0$	Maximum Likelihood Estimator	$\frac{n}{\sum x_i}$
$\hat{\lambda}_1$	Unbiased correction for the MLE	$\frac{n-1}{\sum x_i}$
$\hat{\lambda}_2$	Second Method of Moment Estimator	$2\bar{x}^2$
$\hat{\lambda}_3$	Hard estimate, centered around 5	5
$\hat{\lambda}_4$	Similar to the MLE, without n in numerator	$\frac{1}{\sum x_i}$

Maximum Likelihood Estimator : $\hat{\lambda}_0 \frac{n}{\sum x_i}$

Given how we parameterized the exponential distribution, we cannot apply the E to a fraction since $\sum x_i$ is in the denominator. However, we recognize that $\sum x_i$ for an exponential is a gamma function. To make things easier,

Calculating the Expectation

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:

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

Maximum Likelihood Estimator : $\hat{\lambda}_0 \frac{n}{\sum x_i}$

Calculating the Variance

We need to calculate $Var[\hat{\lambda}_0]$
Rather than use the gamma function again, we will use the expected information $I = -l''(\lambda_0)$

$$I = -E[l''(\lambda_0)] = -E\left[-\frac{\lambda^2}{n}\right]$$

and we know: $Var = \frac{1}{I}$
Therefore:

$$Var[\hat{\lambda}_0] = \frac{\lambda^2}{n}$$

Maximum Likelihood Estimator : $\hat{\lambda}_0 \ \frac{n}{\sum x_i}$

Calculating the bias:

$$Bias = E[\hat{\lambda}_0 - \lambda]$$

Which is

$$\lambda\left(\frac{n}{n-1}\right) - \lambda$$

Calculating the Mean Squared Error:

$$MSE = Bias^2 + Var = (E[\lambda_0 - \lambda])^2 + Var([\hat{\lambda}_0])$$

$$MSE = \left(\lambda\frac{n}{n-1} - \lambda\right)^2 + \frac{\lambda^2}{n}$$

Classifying the Estimator

Unbiased: No, since the $E[\hat{\lambda}_0] \neq \lambda$

Asymptotically Unbiased: Yes, since the $\lim_{n\to\infty} \lambda \left(\frac{n}{n-1}\right)$ equals λ

Consistent: Yes, since it is Asymptotically Unbiased and $\lim_{n\to\infty} \frac{\lambda^2}{n}$ equals 0

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

Values for Each Estimator					
Estimator	Value	$E[\lambda]$	$Var[\lambda]$	$Bias[\lambda]$	$MSE[\lambda]$
$\hat{\lambda}_0$	$\frac{n}{\sum x_i}$	$\lambda\frac{n}{n-1}$	$\frac{\lambda^2}{n}$	$\lambda\frac{n}{n-1} - \lambda$	$(\lambda\frac{n}{n-1} - \lambda)^2 + \frac{\lambda^2}{n}$
$\hat{\lambda}_1$	$\frac{n-1}{\sum x_i}$	λ	$\frac{\lambda^2}{n}\frac{n-1}{n}^2$	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	$5 - \lambda$	$(5 - \lambda)^2$
$\hat{\lambda}_4$	$\frac{1}{\sum x_i}$	$\frac{\lambda}{n-1}$	$\frac{\lambda^2}{n^3}$	$\frac{\lambda}{n-1} - \lambda$	$(\frac{\lambda}{n-1} - \lambda)^2 + \frac{\lambda^2}{n^3}$

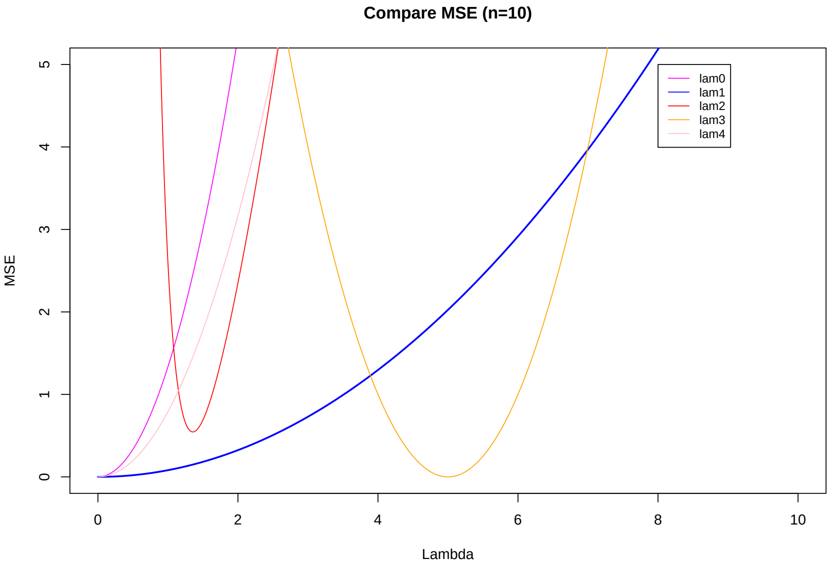
Table of Properties for All Estimators

Properties for Each Estimator				
Estimator	Value	Unbiased	Asymp. Unbaised	Consistent
$\hat{\lambda}_0$	$\frac{n}{\sum x_i}$	No	Yes	Yes
$\hat{\lambda}_1$	$\frac{n-1}{\sum x_i}$	Yes	Yes	Yes
$\hat{\lambda}_2$	$2\bar{x}^2$	No	No	No
$\hat{\lambda}_3$	5	No, unless $\hat{\lambda} = 5$	No	No
$\hat{\lambda}_4$	$\frac{1}{\sum x_i}$	No	Yes	Yes

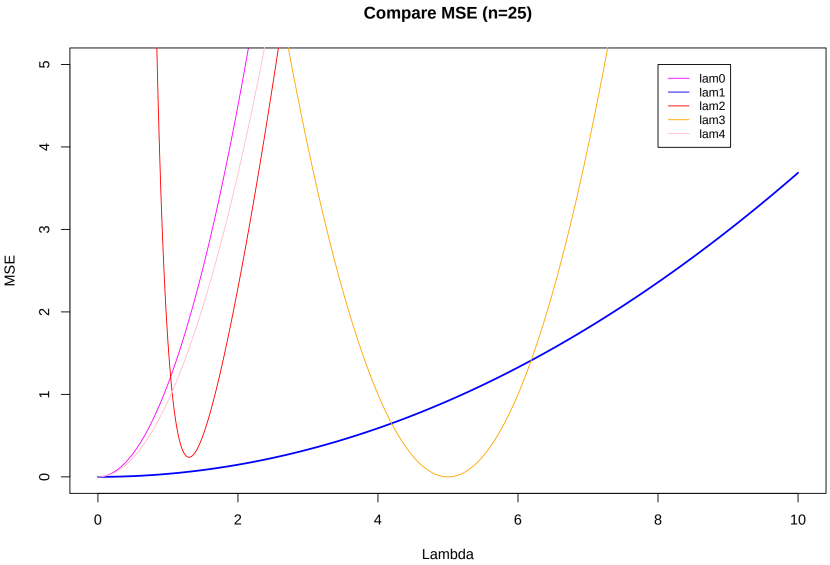
Plotting the MSE

- Now that we have the MSE calculations for all estimators, we can plot them for various n over a range of λ 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 achive a lower MSE for a given range of λ, specifically $4 < \lambda < 6$ at $n = 10 and 25$.
- 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 unseful 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.

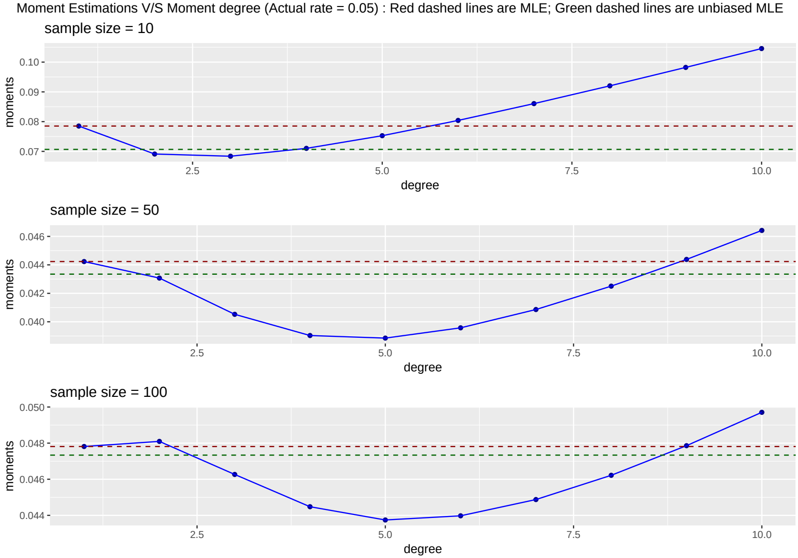
Plotting the MSE



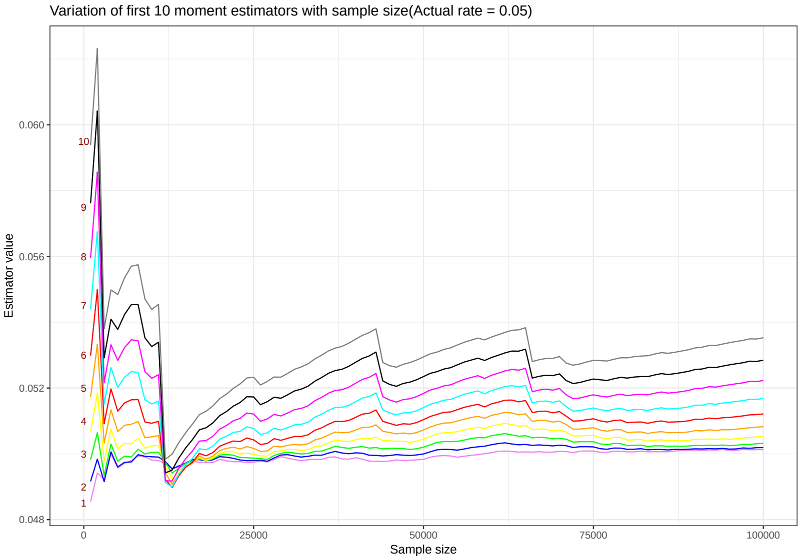
Plotting the MSE



For low sample sizes, MoM estimates decrease and then increase with increasing degree



For a range of sample sizes, the first 10 MoM estimators, although slightly irregular, 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

add criteria for comparing CIs!!

Wald Confidence Interval

```
wald_ci <- function(N, rate, alpha = 0.05){
  x <- rexp(N, rate = rate)
  x_bar <- mean(x)
  se <- x_bar / sqrt(N)
  ci <- x_bar + c(-1, 1) * qnorm(1 - (alpha / 2))*se
  return(ci)
}
```

Gamma Confidence Interval

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 yth percentile of this distribution. Then we can say:

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

And therefore, a $(1 - \alpha)\%$ confidence interval for λ is

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

Score Confidence Interval

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

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

So the $1-\alpha$ score confidence interval for the mean can be obtained by solving

$$\frac{\frac{n}{\lambda_o} - \sum x_i}{\sqrt{n/\lambda_o^2}} = \sqrt{n}(1-\lambda_o\overline{X}) \leq z_{1-\alpha/2}$$

for $\overline{X}$, giving us a confidence interval of

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

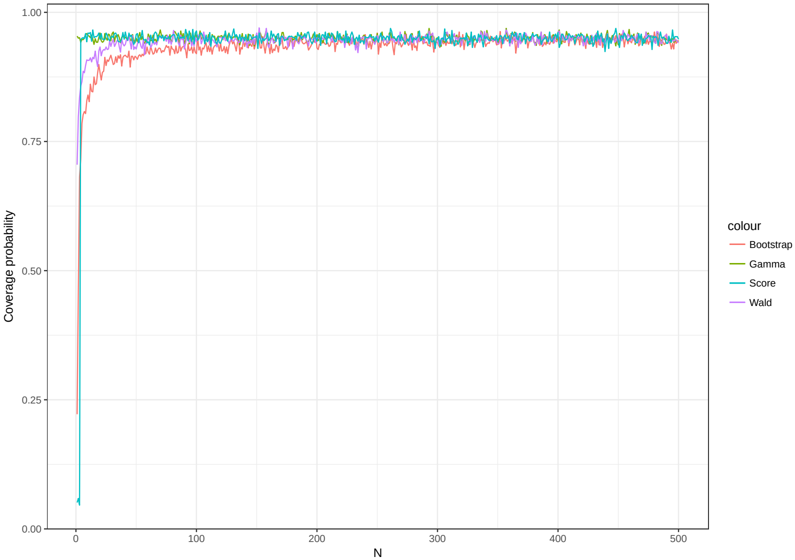
Bootstrap Confidence Interval

- 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 $1/\lambda$
- Not a problem when we use simulated data

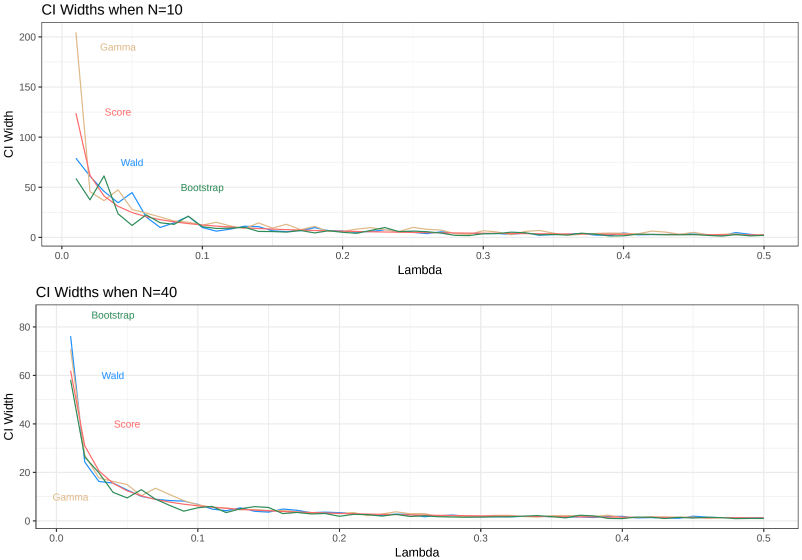
Bootstrap Confidence Interval

```
bootstrap_ci <- function(N, rate, alpha = 0.05){
  # Function to calculate bootstrap CI
  x <- rexp(N, rate = rate)
  x_bar <- mean(x)
  # Number of bootstrap samples
  nb <- 1000
  # Take boostrap samples
  bootstrap_samples <- sample(x, N * nb, replace = TRUE) %>%
    matrix(nrow = N, ncol = nb)
  # Get means of columns
  means <- colMeans(bootstrap_samples)
  # Get deltas (x* - x)
  deltas <- means - x_bar
  deltas <- sort(deltas)
  # Calculate CIs
  ci <- x_bar - quantile(deltas, probs = c(alpha / 2, 1 - (al
  return(c(mean(ci), max(ci)))
```

Coverage probabilities



Plot CI Widths



Summary of Findings and Recommendations

```r
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00  
```