

# 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!

# Maximum Likelihood Estimator

```
##code here
```

# Unbiased correction for the MLE

```
##code here
```

## Second Method of Moment Estimator

```
##code here
```

# 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  $X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Exponential}(\lambda)$ , 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:

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

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

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

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

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

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

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

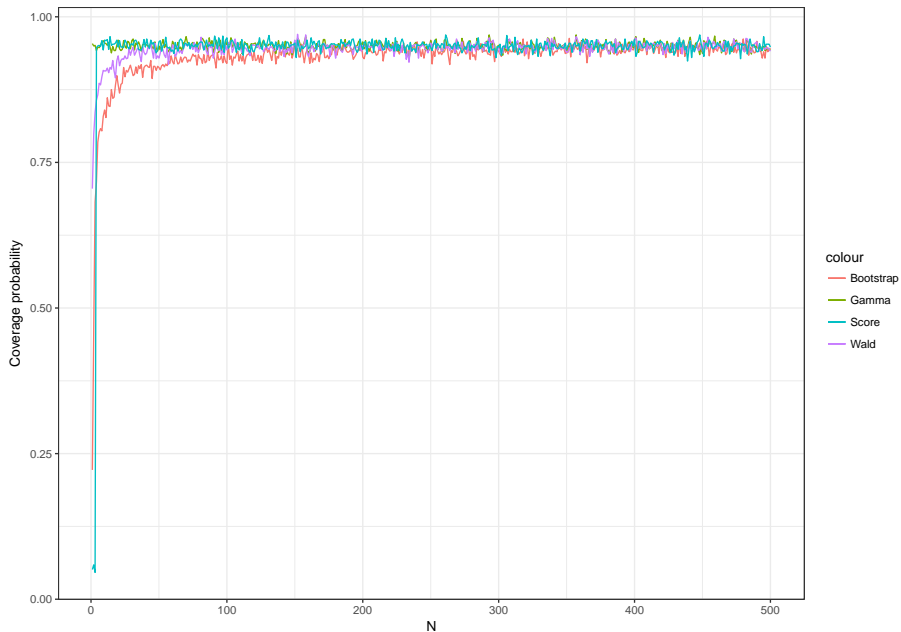
# Bootstrap Confidence Interval

- 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 bootstrap 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 - (alpha / 2)))  
  return(c(mean(ci), max(ci)))  
}
```

# Coverage probabilities



## Plot CI Widths

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

#ggplot() +
  #stat_function(aes(x = 0:1), fun = wald_ci, args = list(N=10, rate=1, alpha=0.05))

# width
set.seed(123)
wald_ci_width <- 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
```

# Summary of Findings and Recommendations

```
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   |