Statistical Inference P1 August

Zamora, Juan Carlo

August 19, 2018

Overview

The purpose of this data analysis is to investigate the exponential distribution and compare it to the Central Limit Theorem. For this analysis, the lambda will be set to 0.2 for all of the simulations. This investigation will compare the distribution of averages of 40 exponentials over 1000 simulations.

Simulations

Set the simulation variables lambda, exponentials, and seed.

```
ECHO=TRUE
set.seed(6969)
lambda = 0.2
exponentials = 40
```

Run Simulations with variables

```
Means = NULL
for (i in 1 : 1000) Means = c(Means, mean(rexp(exponentials, lambda)))
```

Sample Mean versus Theoretical Mean

Sample Mean

Calculating the mean from the simulations with give the sample mean.

```
mean(Means)
```

```
## [1] 4.990647
```

Theoretical Mean

The theoretical mean of an exponential distribution is lambda^-1.

```
lambda^-1
```

```
## [1] 5
```

Difference of Means

Minimal difference from the simulation and the theoretical values.

```
abs(mean(Means)-lambda^-1)
```

```
## [1] 0.00935296
```

Sample Variance versus Theoretical Variance

Sample Variance

Calculate the variance

```
var(Means)
```

```
## [1] 0.5778058
```

Theoretical Variance

The theoretical variance of an exponential distribution is (lambda * sqrt(n))^-2.

```
(lambda * sqrt(exponentials))^-2
```

```
## [1] 0.625
```

Difference of variance

Minimal difference from the simulation and the theoretical values.

```
abs(var(Means)-(lambda * sqrt(exponentials))^-2)
```

```
## [1] 0.04719419
```

Distribution

This is a density histogram of the 1000 simulations. There is an overlay with a normal distribution that has a mean of lambda^-1 and standard deviation of (lambda*sqrt(n))^-1, the theoretical normal distribution for the simulations.

```
library(ggplot2)
ggplot(data.frame(y=Means), aes(x=y)) +
  geom_histogram(aes(y=..density..), binwidth=0.0625, fill="#E70F00",
                 color="pink") +
  stat_function(fun=dnorm, args = list(mean=lambda^-1,
                                       sd=(lambda*sqrt(exponentials))^-1),
                size=1.5) +
  labs(title="Simulation", x="Mean")
```

Simulation