part1

FarzanehBina

2024-03-01

Part1: Simulation Exercise

Overview: This part is going to execute simulations and data analysises to illustrate application of the central limit theorem. R programming will be the major tool to realize the mentioned goal.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(1337)
lambda = 0.2
exponentials = 40
```

Run Simulations with variables

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

Sample Mean versus Theoretical Mean

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

```
mean(simMeans)
```

```
## [1] 5.055995
```

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

```
lambda^-1
```

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

Comparison There is only a slight difference between the simulations sample mean and the exponential distribution theoretical mean.

```r
abs(mean(simMeans)-lambda^-1)
```

```
## [1] 0.05599526
```

Sample Variance versus Theoretical Variance

Sample Variance Calculating the variance from the simulation means with give the sample variance.

```r
var(simMeans)
```

```
## [1] 0.6543703
```

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

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

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

Comparison There is only a slight difference between the simulations sample variance and the exponential distribution theoretical variance.

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

```
## [1] 0.0293703
```

Distribution

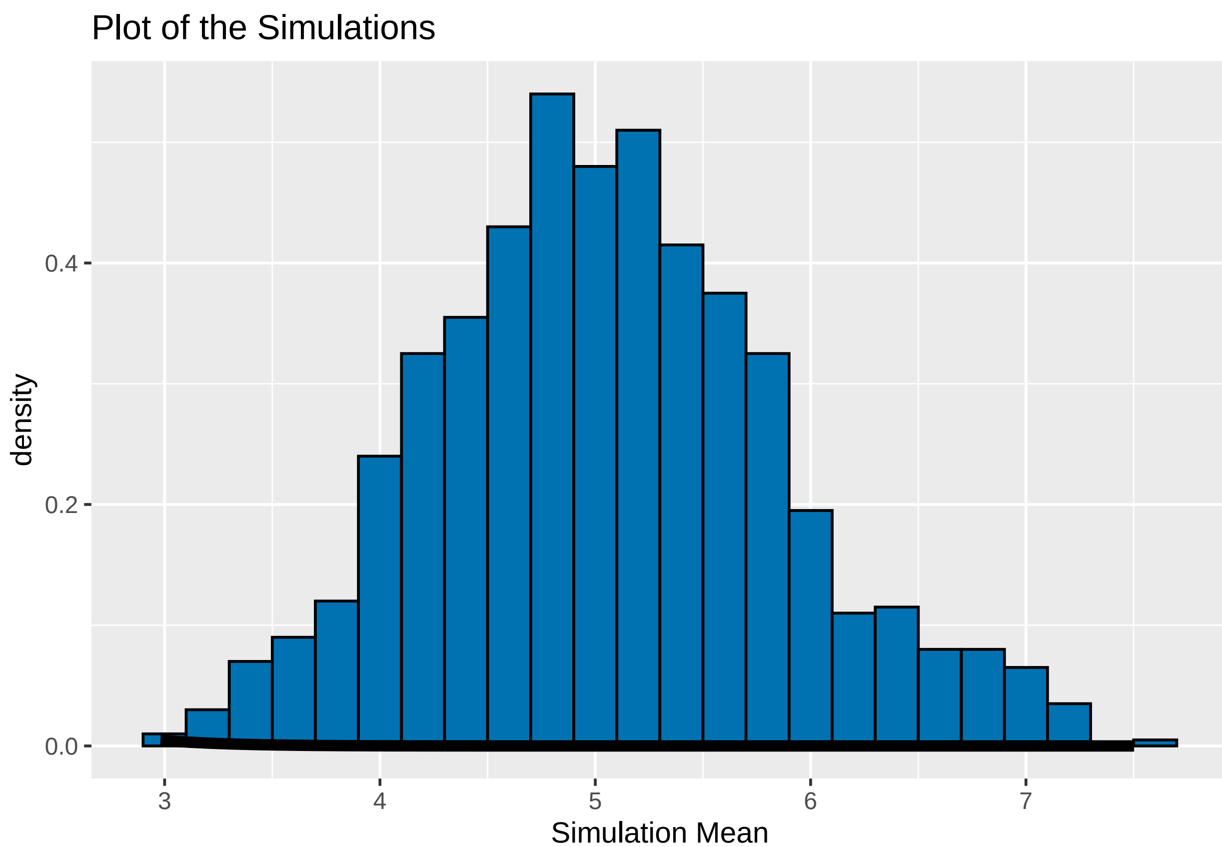
This is a density histogram of the 1000 simulations. There is an overlay with a normal distribution that has a mean of lambda^-1 and standard deviation of (lambda*sqrt(n))^-1, the theoretical normal distribution for the simulations.

```r
library(ggplot2)
ggplot(data.frame(y=simMeans), aes(x=y)) +
  geom_histogram(aes(y=..density..), binwidth=0.2, fill="#0072B2",
                 color="black") +
  stat_function(fun=dnorm, arg=list(mean=lambda^-1,
                                    sd=(lambda*sqrt(exponentials))^-1),
                size=2) +
  labs(title="Plot of the Simulations", x="Simulation Mean")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
## Warning in stat_function(fun = dnorm, arg = list(mean = lambda^-1, sd = (lambda
## * : Ignoring unknown parameters: `arg`
```

```
## Warning: The dot-dot notation (`..density..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(density)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```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
```