

≡ Item Navigation

✓ Peer-graded Assignment: Statistical Inference Course Project

[Instructions](#) [My submission](#)[Discussions](#)

The project consists of two parts:

1. A simulation exercise.
2. Basic inferential data analysis.

You will create a report to answer the questions. Given the nature of the series, ideally you'll use knitr to create the reports and convert to a pdf. (I will post a very simple introduction to knitr). **However, feel free to use whatever software that you would like to create your pdf.**

Each pdf report should be no more than 3 pages with 3 pages of supporting appendix material if needed (code, figures, etcetera).

Review criteria

less ^

- Did you show where the distribution is centered at and compare it to the theoretical center of the distribution?
- Did you show how variable it is and compare it to the theoretical variance of the distribution?
- Did you perform an exploratory data analysis of at least a single plot or table highlighting basic features of the data?
- Did the student perform some relevant confidence intervals and/or tests?
- Were the results of the tests and/or intervals interpreted in the context of the problem correctly?
- Did the student describe the assumptions needed for their conclusions?

Part 1: Simulation Exercise Instructions

less ^

In this project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem. The exponential distribution can be simulated in R with `rexp(n, lambda)` where `lambda` is the rate parameter. The mean of exponential distribution is $1/\lambda$ and the standard deviation is also $1/\lambda$. Set `lambda = 0.2` for all of the simulations. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials. You should

1. Show the sample mean and compare it to the theoretical mean of the distribution.
2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution.
3. Show that the distribution is approximately normal.

In point 3, focus on the difference between the distribution of a large collection of random exponentials and the distribution of a large collection of averages of 40 exponentials.

As a motivating example, compare the distribution of 1000 random uniforms

```
1 hist(runif(1000))
```

and the distribution of 1000 averages of 40 random uniforms

```
1 mns = NULL
2 for (i in 1 : 1000) mns = c(mns, mean(runif(40)))
3 hist(mns)
```

