

Central Limit Theorem

by Jordan DeHerrera

Synopsis

The central limit theorem states that the distribution of sample means of normalized variables approaches the standard normal distribution as sample size increases. We will demonstrate this by using the exponential distribution function in R. By increasing the number of iterations of the exponential distribution function, we will show that the sample means of the simulations will approach a normal distribution.

Assumptions

- Lambda set to 0.2
- 40 simulations
- Compare distributions of 1, 10, and 1000 iterations

Data

The following prepares the sample populations and then calculates the mean of each sample population:

```
# Create empty variables for means
means_10 = NULL
means_1000 = NULL

# Create empty variables for distributions
expdist_10 = NULL
expdist_1000 = NULL

# Create 40 simulations with Lambda of 0.2
expdist_1 = rexp(40,0.2)

# Create 10 iterations of 40 simulations
for (i in 1 : 10) expdist_10 = c(expdist_10 , rexp(40, 0.2))

# Create 1000 iterations of 40 simulations
for (i in 1 : 1000) expdist_1000 = c(expdist_1000, rexp(40, 0.2))

# Separate each of the iterations of 40 simulations
# into separate samples and then calculate the mean
# of those sample populations

# 1 iteration
means_1 <- mean(expdist_1)

# 10 iterations
```

```
means_10 <- unlist(lapply(split(expdist_10,
                              ceiling(seq_along(expdist_10)/40)),mean))

# 1000 iterations
means_1000 <- unlist(lapply(split(expdist_1000,
                                   ceiling(seq_along(expdist_1000)/40)),mean))
```

Results - Simulation Results vs Theory

Variance

The theoretical variance of an exponential distribution is the following:

```
0.2^-2
```

```
## [1] 25
```

Here are the results of the simulations:

```
var(expdist_1)
## [1] 20.84128

var(expdist_10)
## [1] 21.65653

var(expdist_1000)
## [1] 24.68754
```

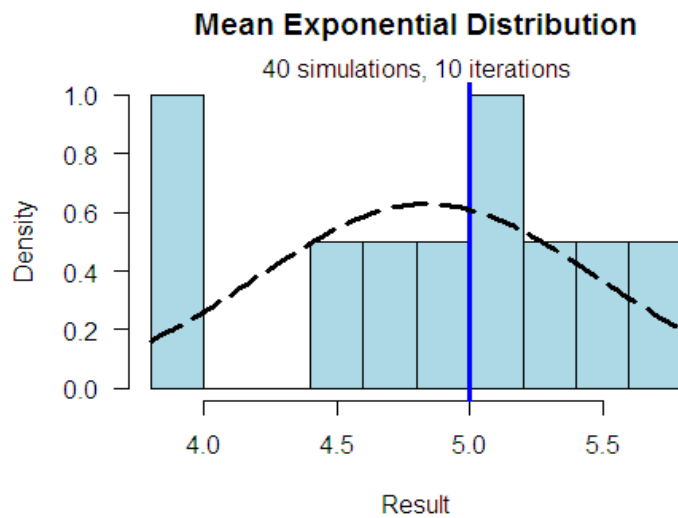
As shown in the above, as the number of iterations increases, the variance approaches the theoretical variance.

Distribution of the Mean

As the number of iterations increases, we can also see that the distribution begins to smooth out and appear to be closer to a normal distribution's "bell curve" type shape. The distribution is centered around the theoretical mean which is $1.0/0.2$ or 5.

```
# Calculate theoretical mean
theoretical_mean <- 1.0/0.2

# Exponential Distribution of 10 iterations
hist(means_10, main="Mean Exponential Distribution",
     xlab="Result",
     col="lightblue",
     las=1,
     breaks=10,
     prob = TRUE)
mtext("40 simulations, 10 iterations",3)
abline(v=theoretical_mean, col="blue", lwd=3)
curve(dnorm(x, mean=mean(means_10), sd=sd(means_10)), add=TRUE, lwd=3, lty=5)
```



```
# Exponential Distribution of 1000 iterations
hist(means_1000, main="Mean Exponential Distribution",
     xlab="Result",
     col="lightblue",
     las=1,
     breaks=200,
     prob = TRUE)
mtext("40 simulations, 1000 iterations",3)
abline(v=theoretical_mean, col="blue", lwd=3)
curve(dnorm(x, mean=mean(means_1000), sd=sd(means_1000)), add=TRUE, lwd=3, lty=5)
```

