

Exponential Distribution

Main goal

Illustrate the properties of the distribution of the mean of 40 exponentials.

1. Show the sample mean and compare it to the theoretical mean of the distribution

Initialize an array to store the means

```
means <- 1:1000
```

Run the simulations

Set the given parameters

```
lambda <- 0.2
observations <- 40
```

For each simulation, generate the exponential distribution, calculate its mean and store it in the `means` array

```
set.seed(20150124)
for (simulation in 1:1000) { means[simulation] <- mean(rexp(n = observations, rate = lambda)) }
```

Compare the mean of the simulations means and the theoretical mean

The theoretical sample mean for an exponential distribution is

$$\bar{X} = \mu = \frac{1}{\lambda}$$

Calculate the difference between sample and theoretical means

```
simulations_mean_of_means <- mean(means)
theoretical_mean <- 1/lambda
mean_difference <- abs(theoretical_mean - simulations_mean_of_means)
```

simulations_mean_of_means: 4.9846

theoretical_mean: 5

difference: 0.0154

This minimal difference shows that the means are very similar, therefore the predicted mean is a good approximation.

2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution

The theoretical variance of the sample mean distribution is the population variance divided by the sample size

$$\frac{\sigma^2}{\text{observations}}$$

The value of the standard deviation σ for an exponential distribution is, we are told,

$$\frac{1}{\lambda}$$

Calculate the difference between sample and theoretical variances

```
simulations_variance <- var(means)
theoretical_variance <- (1/lambda)^2 / observations
variance_difference <- abs(theoretical_variance - simulations_variance)
```

simulations_variance: 0.648

theoretical_variance: 0.625

variance_difference: 0.023

This minimal difference shows that the variances are very similar.

3. Show that the distribution is approximately normal

That can be shown using the *Central Limit Theorem*, because these exponential distributions follow the prerequisites: they are random samples generated by simulation.

The *CLT* states 3 consequences:

The mean of the distribution of sample means equals the population mean.

We verified that in step 1 above.

The standard deviation of the distribution of sample means equals the standard error.

We can verify that as follows:

The theoretical standard deviation for an exponential distribution is

$$\sigma = \frac{1}{\lambda}$$

The standard deviation for a distribution of sample means is equal to the standard error

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{\text{observations}}}$$

```
simulations_sd <- sd(means)
theoretical_sd <- (1/lambda) / sqrt(observations)
sd_difference <- abs(theoretical_sd - simulations_sd)
```

simulations_sd: 0.805

theoretical_sd: 0.7906

sd_difference: 0.0144

The theoretical standard deviation and the observed standard deviation are almost equal.

The distribution of sample means will be approximately normally distributed.

Graphically, the theoretical mean (in blue) and the sample mean (in red) almost overlap, which means that the sample means really are centered around the theoretical mean. The resulting distribution is characteristically bell-shaped and symmetrical.

```
hist(means, breaks=40, ylab="density")  
abline(v = theoretical_mean, col = "blue")  
abline(v = simulations_mean_of_means, col = "red")
```

