Statistical_Inference_Part1

Sharath G

```
    library(knitr)
    library(ggplot2)
    opts_chunk$set(echo=TRUE, warning=FALSE, message=FALSE)
```

In this analysis, we aim to validate that the mean and variance of the sampling distribution of means of a distribution is related to the population distribution. We also prove that the sampling distribution of means follows a normal distribution, validating central tendency theorem.

Create a sampling distribution of means from the exponential distribution

```
set.seed(3)
exp_dist <- rep(NA, 1000)

for (i in 1:1000) {
  exp_dist [i] <- mean(rexp(40,0.2))
}
```

Now, let's calculate the mean and the variance of the sampling distribution

```
sample_mean <- mean(exp_dist)
sample_variance <- var(exp_dist)
```

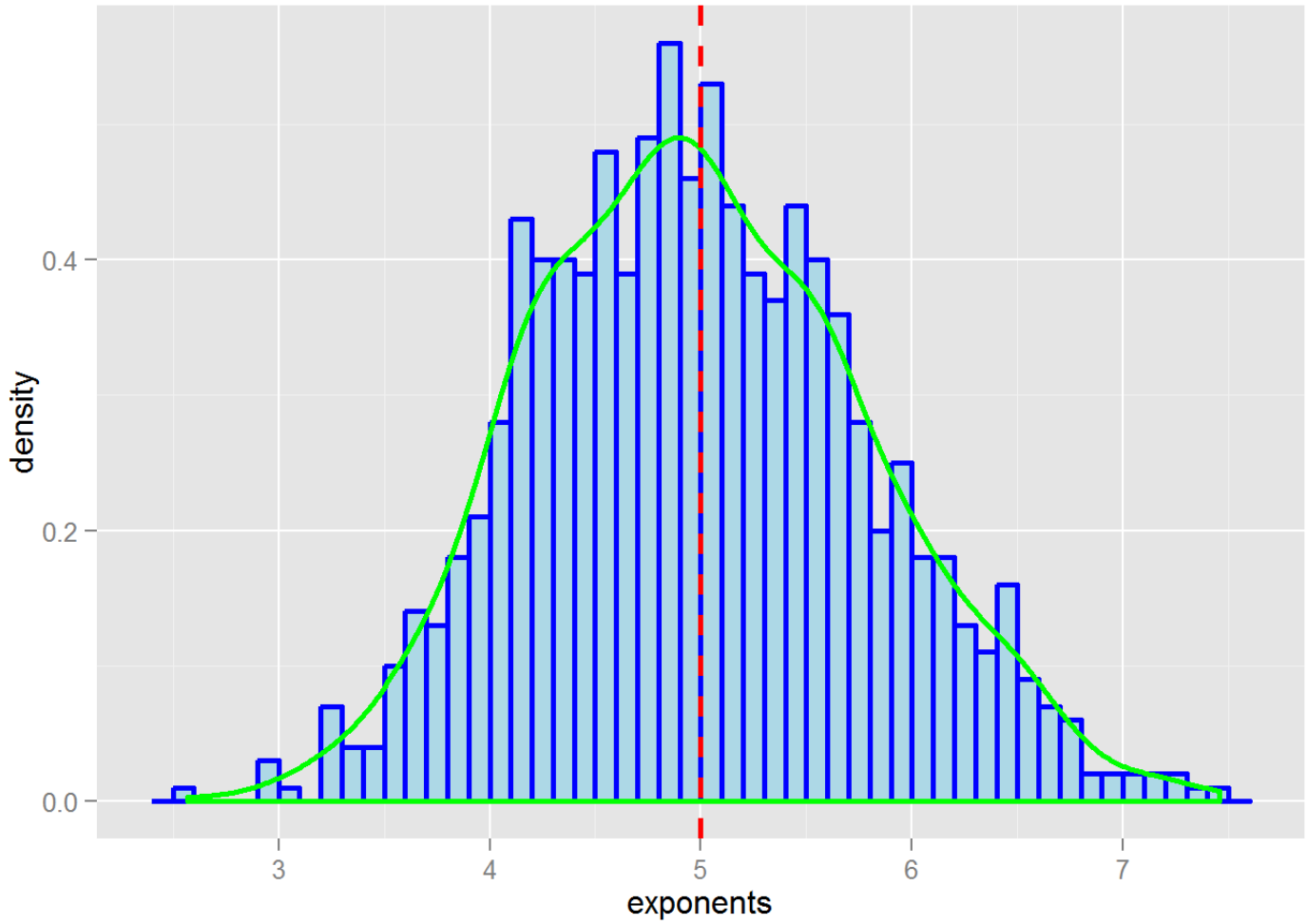
Now, let's calculate the theoretical mean and variance

```
theoretical_mean <- 1/0.2
theoretical_var <- (1/0.2)^2/40
```

We have validated that the theoretical parameters are almost equal to the sample parameters. The sample mean is 4.99 almost equivalent to the theoretical mean 5. Similarly, the sample variance is 0.63 almost equivalent to the theoretical variance 0.62.

Now, let's construct a plot to determine if the sampling distibution of means of the exponential distribution is normal.

```
library(ggplot2)
exp_dist <- data.frame(exp_dist)
names(exp_dist) [1] <- "exponents"
g <- ggplot(exp_dist, aes(exponents))
g <- g + geom_histogram(aes(y = ..density..), fill = "lightblue", binwidth=1/10, color = "blue", lwd
= 1)
g <- g + geom_density(lwd = 1, lty = 1, color = "green")
g + geom_vline(xintercept = 5, lwd = 1, lty =2, color = "red")
```



We observe that the distribution of exponents is almost normal, thus validating the central tendency theorem.