Simulation Exercise

GS

4/8/2021

Background

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.

```
library(tidyverse)
library(tinytex)
theme_set(theme_bw())
```

Simulations

Setting lambda y exponentials

```
set.seed(2021) #for reproducibility

lambda = .2
expo = 40

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

Theoretical mean vs Sample mean

Sample mean (from the simulation)

```
print(paste('The sample mean is ', round(mean(sim_means), 4)))
```

```
## [1] "The sample mean is  5.0086"
```

```
print(paste('The theoretical mean is ', 1/lambda))
```

```
## [1] "The theoretical mean is  5"
```

```r
mean_dif <- abs(mean(sim_means) - (1/lambda))
print(paste('The mean difference is ', round(mean_dif, 4)))
```

```
## [1] "The mean difference is  0.0086"
```

Theoretical variance vs Sample variance

```r
print(paste('The sample variance is ', round(var(sim_means), 4)))
```

```
## [1] "The sample variance is  0.6313"
```

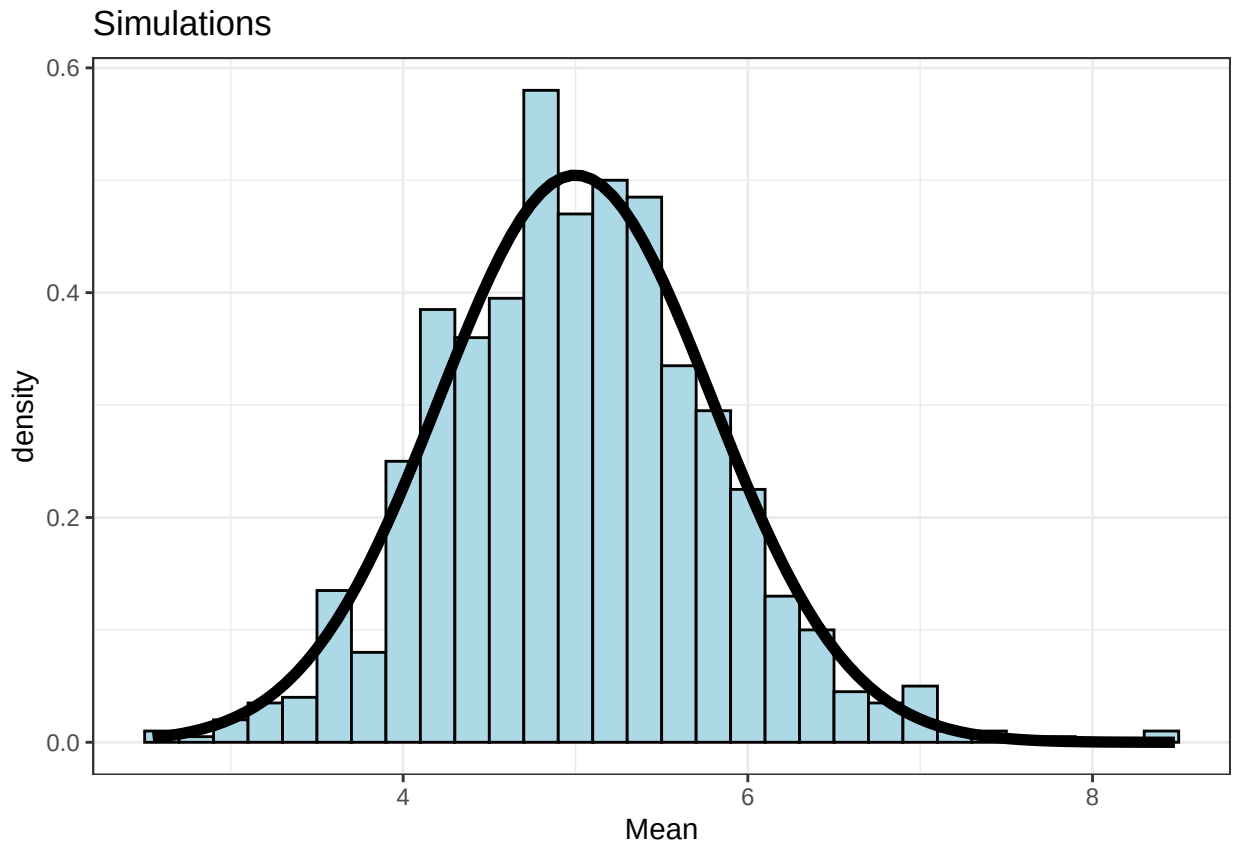
```r
print(paste('The theoretical variance is ', (lambda * sqrt(expo))^-2))
```

```
## [1] "The theoretical variance is  0.625"
```

```r
var_dif <- var(sim_means) - (lambda * sqrt(expo))^-2
print(paste('The variance difference is ', round(var_dif, 4)))
```

```
## [1] "The variance difference is  0.0063"
```

```r
ggplot(data.frame(y=sim_means), aes(x=y)) +
  geom_histogram(aes(y=..density..), binwidth=0.2, fill="lightblue",
                 color="black") +
  stat_function(fun=dnorm, args=list(mean=1/lambda,
                                     sd=(lambda*sqrt(expo))^-1), size=2) +
  xlab('Mean') +
  ggtitle('Simulations')
```




```
sessionInfo()
```

```
## R version 4.1.0 (2021-05-18)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19043)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=Spanish_Mexico.1252  LC_CTYPE=Spanish_Mexico.1252
## [3] LC_MONETARY=Spanish_Mexico.1252 LC_NUMERIC=C
## [5] LC_TIME=Spanish_Mexico.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] tinytex_0.32    forcats_0.5.1   stringr_1.4.0   dplyr_1.0.7
##  [5] purrr_0.3.4     readr_1.4.0     tidyr_1.1.3     tibble_3.1.2
##  [9] ggplot2_3.3.5   tidyverse_1.3.1
##
## loaded via a namespace (and not attached):
##  [1] tidyselect_1.1.1  xfun_0.24         haven_2.4.1       colorspace_2.0-2
##  [5] vctrs_0.3.8       generics_0.1.0    htmltools_0.5.1.1 yaml_2.2.1
##  [9] utf8_1.2.1        rlang_0.4.11      pillar_1.6.1      glue_1.4.2
## [13] withr_2.4.2       DBI_1.1.1         dbplyr_2.1.1      modelr_0.1.8
```

```
## [17] readxl_1.3.1      lifecycle_1.0.0   munsell_0.5.0     gtable_0.3.0
## [21] cellranger_1.1.0  rvest_1.0.0       evaluate_0.14     labeling_0.4.2
## [25] knitr_1.33        fansi_0.5.0       highr_0.9         broom_0.7.8
## [29] Rcpp_1.0.7        scales_1.1.1      backports_1.2.1   jsonlite_1.7.2
## [33] farver_2.1.0      fs_1.5.0          hms_1.1.0         digest_0.6.27
## [37] stringi_1.6.2     grid_4.1.0        cli_3.0.0         tools_4.1.0
## [41] magrittr_2.0.1    crayon_1.4.1      pkgconfig_2.0.3   ellipsis_0.3.2
## [45] xml2_1.3.2        reprex_2.0.0      lubridate_1.7.10  assertthat_0.2.1
## [49] rmarkdown_2.9     httr_1.4.2        rstudioapi_0.13   R6_2.5.0
## [53] compiler_4.1.0
```