Appendix Part I: Code for `plot_sample_means()`

Tamer Köksal

February 13, 2015

```
plot_sample_means <- function(f_sample, n, lambda, nosim100, nosim1000, ylim, a, b,
                      title, ...) {

  # set.seed for reproducibility
    set.seed(42)

  # define vectors to hold our sample means
  means100 <- double(0)
  means1000 <- double(0)

  # generate 100 samples of size n and store their means
  for(i in 1:nosim100) means100[i] = mean(f_sample(n, rate = lambda))

  # generate 1000 samples of size n and store their means
  for(i in 1:nosim1000) means1000[i] = mean(f_sample(n, rate = lambda))

  # set up a two panel plot
  par(mfrow=c(1,2))
  par(mar=c(5,2,5,1)+0.1)

  # plot histogram and density of the 100 sample means
  hist(means100, probability = TRUE, col="light grey", border="grey",
       ylim = c(0, ylim), main="\n\n\nBased on 100 samples", xlab = "Sample means")
  lines(density(means100))
  abline(v = mean(means100), col = "red"
         )
  text(x=5, y = .1, labels="M", col = "red")

  # overlay the normal curve (with mu=1/lambda and variance 1/(n*lambda^2))
  # in blue for comparison  (for the 100-simulation)
  curve(dnorm(x, 1/lambda, 1/(lambda*sqrt(n))), from = a, to = b, col='blue', add=T)

# plot histogram and density of 1000 sample means
  hist(means1000, probability = TRUE, col="light grey", border="grey",
       ylim = c(0, ylim), main="\n\n\nBased on 1000 samples", xlab = "Sample means")
  lines(density(means1000))
  abline(v = mean(means1000), col = "red")
  text(x=5, y = .1, labels="M", col = "red")

  # overlay the normal curve (with mu=1/lambda and variance 1/(n*lambda^2))
  # in blue for comparison (for the 1000-simulation)
  curve(dnorm(x, 1/lambda, 1/(lambda*sqrt(n))), from = a, to = b, col='blue', add=T)

  # return margins to normal and go back to one panel
  par(mar=c(5,4,4,2)+0.1)
  par(mfrow=c(1,1))
```

```
  # add a title
  par(omi=c(0,0,0.75,0))
  title(paste(title, ", n=", n, sep=""), outer=T)
  par(omi=c(0,0,0,0))

  # return the means and variances of sample means  (without printing)
  return(invisible(list(df = data.frame(means100, means1000), means=c(mean(means100),
                                                                      mean(means1000)),
                        variances=c(var(means100), var(means1000)))))
}
```

This code for the function `plot_sample_means()` was originally adopted from an R-bloggers article titled Using R for Introductory Statistics 6, Simulations and modified accordingly to the needs of this project assignment.

This function plots 2 distributions side by side for 100 and 1000 simulations, respectively. User provided input `f_sample` to the function accepts built in R distributions such as `runif`, `rbinom`, `rpois`, etc. (in our case `rexp`). The input `n` is for the sample size; `lambda` for the rate parameter of the exponential distribution; `nosim100` and `nosim1000` for the number of simulations (number of samples to be drawn): 100 and 1000, respectively; `ylim` for the upper limit of the y-axis; `a` and `b` for the lower and upper limits (on the x-axis) of the theoretical normal distribution; and `title` for the title of the plot.