Assignment 2

Danny Kearns

2023-09-21

Please turn this assignment in as a knitted R Markdown (either .html, .pdf or .docx file). Number your answers and provide some rationale to the steps of your code

You will be creating a couple of functions, demonstrate the usage of common functions in R (such as `rnorm`), figuring out how to use a new function and also use some control structures

1. Build a function that takes a vector of size 500 (call the function and argument whatever you would like). If the vector is normally distributed, find the 95% confidence interval and print that to the console (use the `cat()` function). If the vector is not normally distrubted, print a message back to the console that the vector is not normally distributed. Regardless of distribution, plot the density of the vector.

   To determine if the vector is normally distributed, use the `shapiro.test()` function. If the p-value is greater than or equal to 0.05, then the function is normally distributed.

   To calculate the 95% confidence interval, the first thing to do is to calculate the standard error at the 95% level

   $$\text{se} = \frac{z_{0.975} * s}{\sqrt{n}}$$

   Where $z_{0.975}$ is the corresponding value on the standard normal distribution $[\mathcal{N}(0, 1)]$ to the 97.5^{th} percentile (keep in mind that the 2.5 th percentile is just -$z_{0.975}$), s is the sample standard deviation and n is the size of the sample. The confidence interval can then be calculated by adding or subtracting the sample mean, $\bar{x}$. To $z_{0.975}$, use the `qnorm()` function:

   $$\text{CI} = (\bar{x} - \text{se},\ \bar{x} + \text{se})$$

   Run a test case on both functions where the data is normally distributed and one where the data is not. The best way would be to use some `r<distribution>()` functions (R has numerous distributions you can use to create vectors from with random values).

2. Create a function that accepts a dataset. For each column in the dataset, print a statement "`<column name>` is of type: `<variable type>`". Test out your function on the dataset attached to this assignment (show the steps of reading in the dataset, you may need to use the `str()` function and make a class conversion for one of the variables).

3. The code snippet below has three functions `f1`, `f2` and `f3`

```
f1 <- function(string, prefix) {
  substr(string, 1, nchar(prefix)) == prefix
}

f2 <- function(x) {
  if (length(x) <= 1) return(NULL)
  x[-length(x)]
```

```
}

f3 <- function(x, y) {
  rep(y, length.out = length(x))
}
```

What does each function do? What does each function return? Come up with better names for these functions