Simulation Basics

Your Name

Due: dd/mm/yyyy

Project tasks

1. Simulation Intro

1.1 [4 pts] Generate Random Normal Data

Task 1.1.1: Simulate n = 20 random samples from a Normal distribution with mean 0 and standard deviation 3 (hint: use the `rnorm` function), and show a density plot of your sample.

Task 1.1.2: Simulate n = 1000 random samples from a Normal distribution with mean 0 and standard deviation 3, and show a density plot of your sample.

1.2 [5 pts] Describe Random Normal Samples

Corresponding to each plot, [A] report the summary statistics requested, AND [B] explain (one sentence/phrase) how your observations for each sample relate to your understanding of the Normal Distribution from STAT 200 (or similar introductory statistics course). You may use the bullet list below to organize your observations

- **Task 1.2.1: Sample 1 (n = 20)**
  - mean:
  - standard deviation:
  - first quartile:
  - median:
  - third quartile:
- **Task 1.2.2: Sample 3 (n = 1000)**
  - mean:
  - standard deviation:
  - first quartile:
  - median:
  - third quartile:

1.3 [9 pts] Generate and Describe Random Gamma Samples

Repeat the above exercises, for a Gamma distribution (hint: use function `rgamma`). A Gamma distribution has two parameters: shape and scale.

Task 1.3.1: Simulate n = 20 random samples from a Gamma(shape = 4, scale = 3). Show a density plot of your sample, as well as provide and describe the sample mean, median, standard deviation and quantiles again (bullet list is fine).

Task 1.3.2: Simulate n = 1000 random samples from a Gamma(shape = 4, scale = 3). Show a density plot of your sample, as well as provide and describe the sample mean, median, standard deviation and quantiles again (bullet list is fine).

2. [9 pts] Quantile-quantile plots

For each of the following pairs of samples, [A] make a q-q plot (quantile-quantile plot), [B] provide a brief summary to describe the appearance, and [C] to interpret what the q-q plots indicate about the relationship between Sample1 and Sample2.

Task 2.1: Plot, describe, and interpret a q-q plot for:
- Sample1: Normal(mean = 3, sd = 8); n = 100 samples
- Sample2: Normal(mean = -2, sd = 4); n = 100 samples

Task 2.2: Plot, describe, and interpret a q-q plot for:
- Sample1: Normal(mean = 0, sd = 1); n = 100 samples
- Sample2: t-distribution with mean 0; n = 100 samples

Task 2.3: Plot, describe, and interpret a q-q plot for:
- Sample1: Normal(mean = 0, sd = 1); n = 100 samples
- Sample2: Gamma(shape=3, scale=8); n = 100 samples

3. Back on the bicycle... (Random Subsets)

Return to the bike sharing data set from last week (you may need to consult your previous assignment). You will draw and explore several random samples–without replacement–from the data.

Task 3.1: [9 pts] Make 3 scatterplots for 3 different random data subsets of size 500 each. Show a scatter plot of *propCasual* and *temp* for each random subset of the data. Make sure to correct for overplotting, and color the points according to the day of the week or weather, whichever is more informative. What do you see? Are these plots as informative as the previous assignment? Are they roughly consistent across random samples?

Task 3.2: [9 pts] Repeat the exercise for 3 different samples of size 2,000 each.