Econ 108 FAll 2022
Problem Set 1

This problem set is due at 7pm, Friday, October 7th, 2022.

1. Based on your understanding of the lecture notes and the textbook, explain the difference between frequentist uncertainty and Bayesian uncertainty. Which school do you prefer?

2. Based on your understanding of the lecture notes and the textbook, explain why computer (Monte Carlo) simulation can be useful for the study of frequentist uncertainty.

3. Based on your understanding of the lecture notes and the textbook, explain the relation and difference between the roles of computer (Monte Carlo) simulation and the bootstrap method in the study of frequentist uncertainty.

4. Use R to reproduce the box-plot on page 25 of the lecture slides (01.pdf) and the scatter-plot on page 26-27.

   (a) The syntax of the commands on page 25 and on page 27 are almost identical. Why does one generate a box-plot yet the other one generates a scatter-plot?

   (b) Change the second command above to generate a box-plot of price for each year instead.

5. On page 8 of the lecture notes 01uncertainty.pdf (please use the most updated version on canvas), we simulate from a *log normal* distribution. First we draw from a normal random variable Z_i with mean (μ) 1 and standard deviation (σ) 0.5, and then we calculate $Y_i = \exp(Z_i)$. We plot the histogram of Y_i, $i = 1, \ldots, 50,000,000$.

   (a) Theoretical calculations have shown that $EY_i = Ee^{Z_i} = e^{\mu + \frac{\sigma^2}{2}}$. Calculate this number. Is this consistent with the simulation results on page 8 of 01uncertainty.pdf (You can also find the code in 01uncertainty.R posted on canvas)?

   (b) Change μ and σ to other values of your choice. For example you can try $\mu = 0, \sigma = 1$, etc. Recalculate $EY_i = Ee^{Z_i} = e^{\mu + \frac{\sigma^2}{2}}$, and rerun the simulation. Are the simulation results still consistent with the theoretical formula?

6. Consider the simulation of the sample mean on page 9 of the lecture notes 01uncertainty.pdf. Change the sample size n from the current 50 to 1, 2, 5, 10, 50, and rerun the simulation. How does the shape of the resulting histogram change when the sample size changes from small (e.g. 1 or 2) to become larger (e.g. 50)? What are the underlying statistical principles that are consistent with the observed change of the shape of the histogram of the sample mean when the sample size increases?

7. A variant of the bootstrap method is called m-out-of-n bootstrap, where the sample size in each bootstrap iteration $b = 1, \ldots, B$ is m, and m is typically less than the full sample size n.

   Let's take $m = n/4$.

   (a) Change the bootstrap program on page 22 of the lecture slides (01uncertainty.pdf), so that the sample size in each bootstrap iteration is m instead of n. (Hint: you only need to change one line of the code).

   (b) How does the new bootstrap standard deviation estimate differ from the original bootstrap standard deviation estimate (where the sample size is n in each bootstrap iteration)?

   (c) How does the histogram of the new bootstrap draws compare to the histogram of the original bootstrap draws on page 23?

   (d) Can you explain based on theoretical reasoning the differences you found in the previous two questions?