Statistics 3080
Homework 4
Due: Wednesday, October 25

Complete the following problems in a commented R file. Include any output requested as a comment following your code. Save any produced plots in a single PDF file.

Problem 1 (25 points): Consider a normally distributed population with mean 26.7 and standard deviation 4.4. Assume that the population mean is unknown to the practitioner, but that she knows the population standard deviation. The practitioner plans to take a random sample from the population to conduct a two-tailed hypothesis test using the one-sample z-test for the mean. For each part, repeat the following steps 10,000 times to conduct a Monte Carlo simulation to estimate the Type I error of the test the practicioner will conduct:

i. Draw a random sample of the given size from the population.

ii. Determine the test statistic and its associated p-value.
(You should *not* use a testing function.)

iii. Record whether the null hypothesis would be rejected or not.

(a) Using samples of 15 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

(b) Using samples of 30 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

(c) Using samples of 45 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

Problem 2 (15 points): Consider the distribution given in Problem 1. Assume that both the population mean and standard deviation are unknown to the practitioner. She plans to take a random sample from the population to conduct a two-tailed hypothesis test using the one-sample t-test for the mean. For each part, repeat the following steps 10,000 times to conduct a Monte Carlo simulation to estimate the Type I error of the test the practicioner will conduct:

i. Draw a random sample of the appropriate size from the population.

ii. Determine the test statistic and its associated p-value.
(You should *not* use a testing function.)

iii. Record whether the null hypothesis would be rejected or not.

(a) Using samples of 15 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

(b) Using samples of 30 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

(c) Using samples of 45 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

Problem 3 (15 points): Consider the distribution given in Problem 1. Assume that both the population mean and standard deviation are unknown to the practitioner. She plans to take a random sample from the population to conduct a two-tailed hypothesis test using the one-sample z-test for the mean. For each part, repeat the following steps 10,000 times to conduct a Monte Carlo simulation to estimate the Type I error of the test the practicioner will conduct:

i. Draw a random sample of the appropriate size from the population.

ii. Determine the test statistic and its associated p-value.
(You should *not* use a testing function.)

iii. Record whether the null hypothesis would be rejected or not.

(a) Using samples of 15 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

(b) Using samples of 30 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

(c) Using samples of 45 individuals, determine the proportion of samples that would lead the practitioner to reject the null hypothesis.

Problem 4 (20 points): Make a table that summarizes your results from Problems 1-3. What conclusions can you draw from these results?

Problem 5 (50 points): Consider a chi-squared distribution with 2 degrees of freedom, which corresponds to a mean of 2 and a standard deviation of 2.

(a) Plot the density curve of the population distribution and determine its shape.

(b) Repeat Problem 1 using this distribution.

(c) Repeat Problem 2 using this distribution.

(d) Repeat Problem 3 using this distribution.

(e) Create a table similar to the one created in Problem 4 to summarize the results from parts (b)-(d). What conclusions can you draw from these results.

(f) Compare your conclusions from Problem 4 and part (e). What overall conclusions can you draw?