BAX-441 Homework 1

Mehul Rangwala

Fall 2022

Assignment Weight: 5% of your grade

Due Date: Sunday, October 9, 2022 11:59 PM (Both Sections)

Instructions

1. There are 4 questions. Some have multiple parts.

2. All questions should be completed using `R`.

3. Submit an RMD file knitted as HTML. The knitted file should show the code and the result (output) below it.

4. Email me if you have any questions. mrangwala@ucdavis.edu

Question 1 *(10 points)*

The following figure shows the normal quantile plot of daily stock returns in 2010 on the value-weighted total U.S. market index.
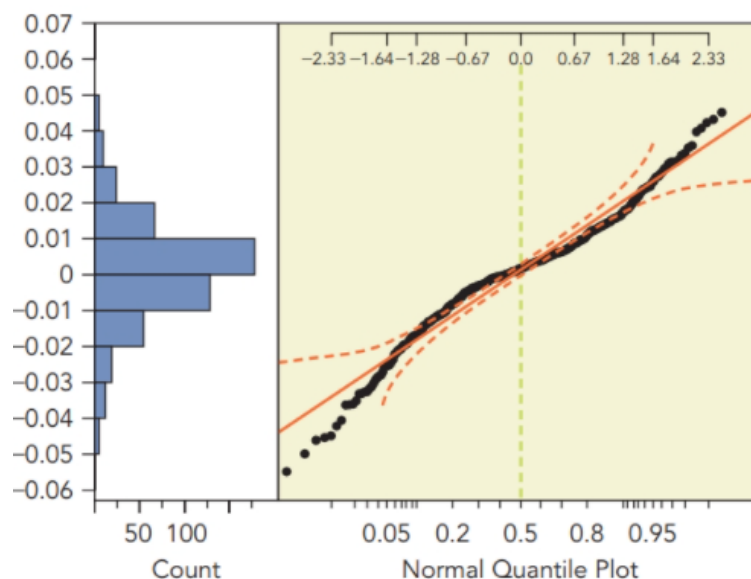
The following table counts the number of returns falling into eight intervals. The table includes the count expected under the assumption that these data are normally distributed, using the sample mean $\bar{x} = 0.0009874$ with $SD = 0.0151$.

Range	Count	Expected Count
$x \le -0.03$	18	10.02
$-0.03 < x \le -0.02$	19	31.24
$-0.02 < x \le -0.01$	56	76.16
$-0.01 < x \le 0$	128	121.42
$0 < x \le 0.01$	178	126.61
$0.01 < x \le 0.02$	66	86.37
$0.02 < x \le 0.03$	23	38.53
$0.03 < x$	16	13.66
Total	504	

a) What does the normal quantile plot indicate about the normality of returns?

b) The table groups all returns that are less than -0.03 and more than 0.03. Why not use more categories to separate very high or low returns?

c) Compute the chi-squared test of goodness of fit and its p-value.

d) Does the chi-squared test agree with the normal quantile plot?

2

e) What's the advantage of using a normal quantile plot to check for normality? The advantage of using the chi-squared test?

Question 2 *(10 points)*

The data set provided on the CSV file *Stock Market* summarizes the number of days the stock market went up or down during each trading day of 2018.

a) Use an appropriate statistical test to determine if these data indicate that trading on some days is better or worse (more or less likely to earn positive returns) than any other.

b) How does the test used in (a) differ from comparing the proportion positive for each day with 0.5?

c) These data only indicate the direction of the market. How does that limit the conclusions we might draw?

Question 3 *(5 points)*

The following figures are from the United States Census in 2014:

Education	
Less than high school	12.3%
High school	29.6%
Some college including junior college	19.4%
College graduate including bachelor's and graduate degrees	38.6%

Can we infer that the General Social Survey in 2014 overrepresented at least one education category? Data set available on the CSV file *GSS2014*. The document *General Social Survey List of Variables* contains the description of all the variables in the GSS data set.

Question 4 *(10 points)*

Over the past few years, immigration was a hot-button issue in American politics. One question that arises is: Are immigrants' educational attainments (DEGREE) different from those born in the United States? Use the General Social Survey data provided on the CSV file *GSS2014* to conduct a statistical test to answer the question. The document *General Social Survey List of Variables* contains the description of all the variables in the GSS data set. Choose whichever variables are relevant to answering the question.