Financial Time Series I and Methods of Statistical Prediction

Project 1: Review on Basics

Due Date:

1. (Set up hypothesis) (a) In the legal world, a null hypothesis might be "This person is innocent." A type I error would be judging the person guilty when he is innocent. A Type II error would involve declaring the person innocent when he is guilty. If one accepts the thought that it is better to release a guilty person than to convict an innocent one, then it would be important to minimize the chances of Type I error or the chances of Type II error.
   (b) In the world of medicine, a null hypothesis might be "This drug will cure an illness." A Type I error would be concluding that the drug does not work when it actually does. A Type II error would conclude that the drug does work when it actually doesn't. If one agrees that spending time and money on a useless drug would replace what might be some other effective treatment, what kind of error should be minimized?

2. (hypothesis testing) Assume that two samples of people have the indicated ethnic distributions. The sample sizes are 25 and 20, respectively.

	African-American	Native American	Caucasian	Oriental
Sample #1	3	1	15	6
Sample #2	10	4	3	3

   Here is your challenge: Choose one sample at random, then randomly pick one individual from the chosen sample. Based on your observation of the individual, you must make a conjecture as to which sample the chosen individual belongs. Here is a null hypothesis, H_o.

   $$H_o: \text{ The individual came from Sample \#1.}$$

   In this case, a rejection of H_o is, of course, equivalent to asserting that the individual came from Sample #2. Here are four possible strategies among many relating to H_o.

   - Strategy #1: Accept H_o if the randomly chosen individual is Caucasian.
   - Strategy #2: Accept H_o if the randomly chosen individual is Caucasian or Oriental.
   - Strategy #3: Accept H_o if the randomly chosen individual is not Native-American.
   - Strategy #4: Accept H_o if the randomly chosen individual is not African-American.

   Find the probabilities of committing Type I and Type II errors for the above four strategies.

3. (graphical display) A plot of the number of defective items produced during 20 consecutive days at a factory is shown in page 3.

(a) Draw a histogram that shows the frequencies of the number of defective items.
(b) Give one fact that is obvious from the histogram but is not obvious from the scatterplot.
(c) Give one fact that is obvious from the scatterplot but is not obvious from the histogram.

4. (one sample test for proportion) A large university provides housing for 10% of its graduate students to live on campus. The university's housing office thinks that the percentage of graduate students looking for housing on campus may be more than 10%. The housing office decides to survey a random sample of graduate students, and 62 of the 481 respondents say that they are looking for housing on campus.
(a) On the basis of the survey data, would you recommend that the housing office consider increasing the amount of housing on campus available to graduate students? Give appropriate evidence to support your recommendation.
(b) In addition to the 481 graduate students who responded to the survey, there were 19 who did not respond. If these 19 had responded, is it possible that your recommendation would have changed? Explain.

5. (goodness of fit test) The Colorado Rocky Rescue Service wishes to study the behavior of lost hikers. If more were known about the direction in which lost hikers tend to walk, then more effective search strategies could be devised. Two hundred hikers selected at random from those applying for hiking permits are asked Whether they would head uphill, downhill, or remain in the same place if they became lost while hiking. Each hiker in the sample was also classified according to whether he or she was an experienced or novice hiker. The resulting data are summarized in the following table.

	Direction		
	Uphill	Downhill	Remain in Same Place
Novice	20	50	50
Experienced	10	30	40

Do these data provide convincing evidence of an association between the level of hiking Expertise and the direction the hiker would head if lost? Give appropriate statistical evidence to support your conclusion.

6. (probabilitry and normal distribution) The manager of a cultured pearl farm has received a special order for two pearls between 7 millimeters and 9 millimeters in diameter. From past experience, the manager knows that the pearls found in his oyster bed have diameters that are normally distributed with a mean of 8 millimeters and a standard deviation of 2 millimeters. Assume that every oyster contains one pearl. The manager wants to know how many oysters he should expect to open to find two pearls of the appropriate size for this special order.
(a) Determine the probability of finding a pearl of the appropriate size in an oyster selected at random.
(b) In average, how many oysters does he need to open to find two pearls between

7 millimeters and 9 millimeters in diameter?

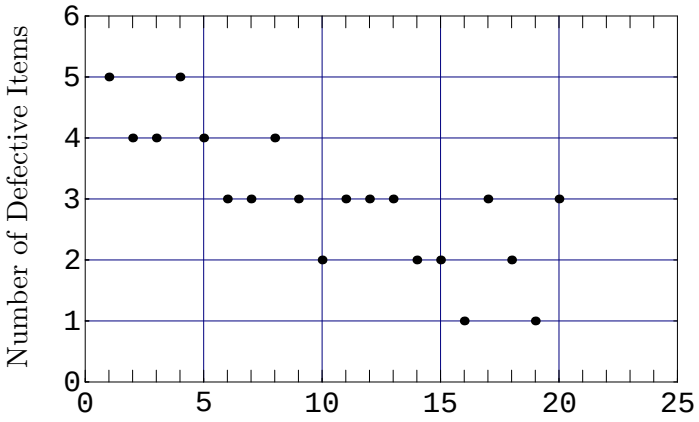
7. Researchers want toknow whether training increases the capability of people to correctly predict outcomes of coin tosses. Each of twenty people is asked to predict the outcome (heads or tails) of 100 independent tosses of a fair coin. After training, they are retested with a new set of 100 tosses. (All 40 sets of 100 tosses are independently generated.) Since the coin is fair, the probability of a correct guess by chance is 0.5 on each toss. The number of correct for each of the 20 people were as follows. The first number denotes the Score Before Training (number correct) and the second number denotes the Score After Training (number correct):

$$(46, 61), (48, 62), (50, 53), (54, 46), (54, 50), (54, 52), (54, 53), (54, 59), (54, 60),$$

$$(54, 61), (55, 55), (56, 59), (57, 55), (58, 50), (58, 56), (61, 58), (61, 64), (63, 67),$$

$$(64, 61), (65, 54)$$

(a) Do the data suggest that after training people can correctly predict coin toss outcomes better than the 50 percent expected by chance guessing alone? Give appropriate statistical evidence to support your conclusion.
(b) Does the statistical test that you completed in part (a) provide evidence that this training is effective in improving a person's ability to predict coin toss outcomes? If yes, justify your answer. If no, conduct an appropriate analysis that would allow you to determine whether or not the training effective.
(c) Would knowing a person's score before training be helpful in predicting his or her score after training? Justify your answer.


Number of Defective Items
0
1
2
3
4
5
6
0
5
10
15
20
25