

Homework AssignmentsHomework #6Assigned: 10/3/14Due: 10/14/14

#1. A random sample of 16 adults is taken from a large population and their weights are recorded.

(a) The sample mean is 130 lbs and the sample standard deviation is 9 lbs. Give the least squares estimate of the population mean. Find the estimated standard error of the estimator of the population mean and the estimated margin of error.

(b) Ten of these 16 adults have weight smaller than 125 lbs. Let p denote the proportion of people in the population with weight smaller than 125 lbs. Find the least squares estimate of p . What is the estimated standard error of your estimator? Give the estimated margin of error.

#2. A random sample of 25 adults is taken from a large population and their weights are recorded. The sample mean is 120 lbs. The weights of all adults in the population follow a normal probability model.

- (a) Suppose the population standard deviation is 9 lbs. Find a 95% confidence interval of the population mean.
- (b) Suppose the population standard deviation is unknown, and the sample standard deviation is 16. Find a 95% confidence interval for the population mean.
- (c) What are the approximate answers in (a) and (b) if the population has a probability model skewed to the right? How can this approximation be improved?
- (d) Suppose a scientist wants to guess the weight of another adult drawn at random from the population. Find a 95% prediction interval for this value assuming the population standard deviation is unknown, the population can be described by a normal probability model, and the sample standard deviation is 16.

#3. Exercise 5.8 in the PNC text book.

#4. Exercise 5.19 in the PNC text book.

5. (a) The time Professor Bal takes to get from his home to work at WPI varies according to a normal distribution. Professor Bal recorded 10 times (minutes) as follows:

12, 13, 11.5, 9, 12.2, 15.6, 10, 8, 11, 12.1.

Professor Bal wants to quantify the evidence that his time to work is generally more than 11 minutes. State the appropriate statistical hypotheses, the test statistic, present its p-value and discuss it. What would you tell Professor Bal if the distribution of the time to work is not normal?

#6. A coin was tossed 500 times, and the following sequence of heads (H) and tails (T) was recorded:

H H T T T T T T T T T T T T H H H H H H H H T H
H H H H T T T H H T H T H T T H H T T H H H H T H
T H H T H T T T T T T T T T H H H H T T H H T H
T T T T T T T H H T T T T T T H H H T T T H T H
H H T T T T T T T T T T T T H H H H H H H H T H
H H H H T T T H H T H T H T T H H T T H H H H T H
T H H T H T T T T T T T T T H H H H T T H H T H
T T T T T T T H H T T T T T T H H H T T T H T H
H H T T T T T T T T T T T T H H H H H H H H T T
H H H H T T T H H T H T H T T H H T T H H H H T H
T H H T H T T T T T T T T T H H H H T T H H T H
T T H H H H T H H T T T T T T H H H T T T T H T H

Is the coin fair? Explain your answer. Use both a confidence interval and a hypothesis test. [This is the same problem assigned in HW 2, Problem 1; so you have the information required to solve this problem already.]

Homework #5

Assigned: 9/26/14

Due: 10/3/14

#1. In a certain population blood pressures (mmHg) of adults follow the normal distribution model with mean 125 and variance 36. Answer the following questions about the population. [Please use a standard normal table, and show your standardization.]

- (a) What proportion of the adults have blood pressures more than 130?
- (b) What proportion of the adults have blood pressures less than 120?
- (c) What proportion of the adults have blood pressures between

#2. Find the quartiles of the standard normal random variable (Z). Thus, find the interquartile range of the standard normal random variable. What is IQRSTD for the standard normal?

#3. (a) Describe the binomial probability model and the normal probability model. Discuss their differences and their similarities.

(b) Discuss the meanings of the parameter p in the binomial probability model and the parameters μ and σ^2 in the normal probability model.

#4. (a) A random sample of 25 people is drawn from a population and the quantity of interest is weight. The population of weights can be described by a normal distribution with mean 120 pounds and variance 36 pounds. Find the probability that the sample mean is at least 121 pounds. What is the probability that a single person weighs at least 121 pounds?

(b) A random sample is drawn from a large population and it is noted whether the people belong to the city or not. The probability that someone belongs to the city is .25. If 20 people are selected, find the probability that there are at least two persons who do not belong to the city.

#5. About 15% of the families in a US town are poor (income below the poverty limit). For a random sample of 50 adults approximate

- (a) The probability that the proportion of poor families in the sample is at least 0.145;
- (b) The probability that the proportion of poor families in the sample is at most 0.155;
- (c) The probability that the proportion of poor families in the sample is at least 0.145 and at most 0.155.

#6. Please see MyWPI for the data in Lab3 (diff.sas) on the differences between strong and left hands. Use all the data to draw a normal quantile plot to assess normality.

Homework #4

Assigned: 9/19/14

Due: 9/26/14

#1. A population is stratified into 5 strata; the sizes of the strata are 120, 125, 280, 325 and 150. Suppose you determined that a sample of 100 is adequate, how many would you take from each strata? Why might these allocations be reasonable?

#2. A study was conducted to investigate the effects of two factors on the fragmentation of an explosive device. The two factors are pipe diameter (PD) and a restraint material (RM). A completely randomized experiment was performed with 3 replicates for each treatment, and the results are shown in the Table below.

	PD	
RM	0.75 in	1.75 in
Sand	.0698	.0625
	.0698	.0615
	.0686	.0619
Earth	.0659	.0699
	.0651	.0620
	.0676	.0602

Find the effects of PD for sand and earth. Find the effects of RM for PD= .75 and PD= 1.75.

#3. (a) In an observational study of elementary school children, ages 6 to 11, a scatter plot of shoe size and test score shows roughly a straight line with positive slope. Do you think that larger shoe size causes higher test score? Are there confounding factors? Explain.

(b) Consider Exercise 3.29 in the PNC text book. Discuss the two possible designs that can be used in this experiment.

#4. (a) What are the different types of errors in sample surveys? Which ones can probability theory handle?

(b) What are the different types of errors in controlled experiments? Which ones can probability theory handle?

#5. (a) Two events on the same sample space have probabilities .35 and .75. Are they mutually exclusive? Explain. Can two events with positive probabilities be both

mutually exclusive and independent? Explain.

(b) A committee of 4 people is to be formed by drawing at random from a single list of 6 men and 3 women including Mary and Tom.

(i) Find the probability that they are all men.

(ii) Find the probability that there are at least one woman.

(iii) Find the probability that Tom is selected but Mary is not.

#6. (a) An urn contains 10,000 balls; some balls are red and the others are blue. A random sample of 100 balls is drawn, and 20 blue balls are observed. Approximate the number of blue balls in the urn.

(b) The number of police offices in a US district is 100. A simple random sample of 200 people is taken from the district, and 20 police officers are found in the sample. Approximate the population size of the district.

(c) In a pond there are N fish (unknown). A biologist took a simple random sample of 25 fish, tagged their tails and returned them to the pond. A week later the biologist took another sample of size 45 and observed 20 fish with tagged tails. Approximate N .

Homework #3

Assigned: 9/12/14

Due: 9/19/14

#1. (a) In a large data set with 10,000 numbers. About how many numbers are at most the .25 quantile and how many numbers are at least the 95th percentile?

(b) Eight judges rated Domino Pizza on a 9-point hedonic scale, and gave scores 1, 8, 9, 6, 5, 4, 8, 9. Obtain the 1-time trimmed mean and the 1-time winsorized mean for these data.

#2. The following data represent the systolic pressures (mmHg) of 10 patients who visited the doctor yesterday:

100, 130, 101, 116, 145, 121, 161, 129, 132, 136

As a measure of spread, compare the standard deviation, interquartile range (IQR), IQRSTD and the mean absolute deviation.

#3. Compare and contrast designed experiments, observational studies and sample surveys. [Hint: Consider randomization.]

#4. (a) Randomize 10 subjects into four groups of 2, 3, 4, 1. [Please show your work.]

(b) In a double-blind randomized controlled experiment, a placebo and a drug are compared, with 10 people in the control and 6 people in the drug group. At the end of the study a score (a measure of pain) is given to each subject and the scores are

Placebo	20, 30, 40, 25, 18, 21, 26, 28, 29, 30
Drug	5, 20, 25, 21, 18, 15

Calculate the effect of the drug over the placebo. Is the drug helpful?

Homework #2

Assigned: 9/5/14

Due: 9/12/14

#1. A coin was tossed 500 times, and the following sequence of heads (H) and tails (T) was recorded:

H H T T T T T T T T T T T T H H H H H H H H T H
H H H H T T T H H T H T H T T H H T T H H H H T H
T H H T H T T T T T T T T T H H H H T T H H T H
T T T T T T T H H T T T T T T H H H T T T T H T H
H H T T T T T T T T T T T T H H H H H H H H T H
H H H H T T T H H T H T H T T H H T T H H H H T H
T H H T H T T T T T T T T T H H H H T T H H T H
T T T T T T T H H T T T T T T H H H T T T T H T H
H H T T T T T T T T T T T T H H H H H H H H T T
H H H H T T T H H T H T H T T H H T T H H H H T H
T H H T H T T T T T T T T T H H H H T T H H T H
T T T T T T T H H T T T T T T H H H T T T T H T H
H H T T T T T T T T T T T T H H H H H H H H T H
H H H H T T T H H T H T H T T H H T T H H H H T H
T H H T H T T T T T T T T T H H H H T T H H T H
T T H H H H T H H T T T T T H H H T T T T H T H

Use a frequency table and a bar chart to summarize the data. [You can do this by hand or use a software.] Is the coin fair? Explain your answer.

#2. The dataset, length.dat, on MyWPI - MA2611A14 - contains the lengths of very carefully measured 2-inch nails. Draw a frequency histogram of these data and comment on its shape. Then, summarize the data in an efficient manner.

#3. Let y_1, \dots, y_n denote a set of n observations. Let

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \text{ and } s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

denote respectively the mean and variance of the data. Explain why

$$\frac{y_i - \bar{y}}{s}, i = 1, \dots, n,$$

are unaffected by change of location and scale of the y_i .

#4. (a) A sample of students in our class was asked about the amount of pocket they spent weekly. The amounts (dollars) are given in the table below.

10, 20, 15, 12, 13, 11, 11, 12, 12, 15, 10, 10, 15, 15

Obtain the five-number [i.e., quartiles and the two adjacent values (fences)] summary of these data. Compare the (mean, standard deviation) and (median,interquartile range). [Use the method discussed in class to find the quartiles.]

Homework #1

Assigned: 8/28/14

Due: 9/5/14

#1. (a) Discuss how you can turn the following numbers into data

20, 25, 30, 22, 26, 27, 22, 28, 23

(b) Discuss how you can turn the following letters into data

a, a, a, a, c, t, g, g, g, c, a, a, t, t

#2. The following data were on the number of accidents on US 95 during 2005 for different segments of the highway,

10, 20, 21, 22, 20, 30, 50, 20, 25, 25, 30, 25, 25, 30, 31, 33,
8, 10, 16, 10, 20, 15, 16, 17, 21, 16, 22, 23, 18, 15, 14, 21,
40, 50, 39, 5, 4, 3, 2, 1, 0, 10, 3, 11, 15, 16, 20, 25,
20, 20, 20, 21, 18, 18, 18, 18, 18, 22, 26, 28, 28, 27, 29, 30,
10, 30, 20, 25, 25, 15, 10, 3, 2, 16, 20.

Draw a histogram of these data. What does the histogram say? [Please use a computer software e.g. SAS.]

#3. The following data were collected for the yield (number of apples per year) of Jim's apple farm over the past decade, starting from the earliest, are

600, 625, 620, 630, 700, 720, 750, 755, 800, 790.

Is the yield stationary? Explain using a time series plot. Obtain the smoothed series of 2-term moving averages and 4-term moving averages. Make a sensible comparison of these two filters.

#4. Exercise 1.6 in the PNC text book.

#5. Exercise 1.8 in the PNC text book.