MSA 8190 Statistical Foundations

Fall 2016

Assignment 5

Issued: October 12, 2016

Due: October 25, 2016

Note: Excel file for Problems 5, 7, 11, 12 and 13 can be found in iCollege. Use R for computations in the following questions and include the R code in your solution sheet. Moreover, upload your R code in iCollege

Problem 1

A confidence interval estimate is desired for the gain in a circuit on a semiconductor device. Assume that gain is normally distributed with standard deviation $\sigma = 20$.

(a) Find a 95% CI for μ when $n = 10$ and $\bar{x} = 1000$.

(b) Find a 95% CI for μ when $n = 25$ and $\bar{x} = 1000$.

(c) Find a 99% CI for μ when $n = 10$ and $\bar{x} = 1000$.

(d) Find a 99% CI for μ when $n = 25$ and $\bar{x} = 1000$.

Problem 2

A manufacturer produces piston rings for an automobile engine. It is known that ring diameter is normally distributed with $\sigma = 0.001$ millimeters. A random sample of 15 rings has a mean diameter of $\bar{x} = 74.036$ millimeters.

(a) Construct a 99% two-sided confidence interval on the mean piston ring diameter.

(b) Construct a 95% lower-confidence bound on the mean piston ring diameter.

Problem 3

A manufacturer produces crankshafts for an automobile engine. The wear of the crankshaft after 100,000 miles (0.0001 inch) is of interest because it is likely to have an impact on warranty claims. A random sample of $n = 15$ shafts is tested and $\bar{x} = 2.78$. It is known that $\sigma = 0.9$ and that wear is normally distributed.

(a) Test $H_0: \ \mu = 3$ versus $H_1: \ \mu \neq 3$ using $\alpha = 0.05$

(b) What is the power of this test if $\mu = 3.25$?

(c) What sample size would be required to detect a true mean of 3.75 if we wanted the power to be at least 0.9?

(d) Explain how the question in part (a) could be answered by using p-value and confidence interval.

Problem 4
Calculate the following probabilities.

(a) If $X_1, X_2, X_3 \sim NID(\mu = 1, \sigma^2 = 1)$, what is $\mathbb{P}(X_1 + X_2 > X_3)$?

(b) If $X_1, X_2 \sim NID(\mu = 0, \sigma^2)$, what is $\mathbb{P}(\bar{X} < S)$?

(c) If $X_1, X_2 \sim NID(\mu = 0, \sigma^2)$, what is $\mathbb{P}(X_1 + X_2 < \sqrt{2(X_1 - X_2)^2})$?

(d) If $X_1, X_2 \sim NID(\mu = 0, \sigma^2)$, $Y_1, Y_2 \sim NID(\mu = 0, 2\sigma^2)$, and X's and Y's are independent, what is

$$\mathbb{P}\left(2\sum_{i=1}^{2}(X_i - \bar{X})^2 + \sum_{i=1}^{2}(Y_i - \bar{Y})^2 > \sigma^2\right) \ ?$$

What about

$$\mathbb{P}\left((X_1 - X_2)^2 + 0.5(Y_1 - Y_2)^2 > \sigma^2\right) \ ?$$

Problem 5
The compressive strength of concrete is being tested by a civil engineer. He tests 12 specimens and obtains the following data.

2216	2237	2225	2301	2318	2255
2249	2204	2281	2263	2275	2295

(a) Is there evidence to support the assumption that compressive strength is normally distributed? Does this data set support your point of view? Include a graphical display in your answer.

(b) Test the normality of the data by Shapiro-Wilk test?

(c) Construct a 95% two-sided confidence interval on the mean strength.

(d) Construct a 95% lower-confidence bound on the mean strength.

Problem 6
The sugar content of the syrup in canned peaches is normally distributed. Suppose that the variance is thought to be $\sigma^2 = 18$ (milligrams)^2. A random sample of $n = 10$ cans yields a sample standard deviation of $s = 4.8$ milligrams.

(a) Test the hypothesis $H_0: \ \sigma^2 = 18$ versus $H_1: \ \sigma^2 \neq 18$ using $\alpha = 0.05$.

(b) What is the P-value for this test?

(c) Find a 95% two-sided confidence interval for σ.

(d) Find a 90% lower confidence bound for σ.

(e) Discuss how part (a) could be answered by constructing a 95% two-sided confidence interval for σ.

(f) Suppose that the true variance is $\sigma^2 = 40$. How large a sample would be required to detect this difference with probability at least 0.90?

Problem 7
The rainfall in acre-feet from 20 clouds that were selected at random and seeded with silver nitrate follows:

18.0	30.7	19.8	27.1	22.3	18.8	31.8	23.4	21.2	27.9
31.9	27.1	25.0	24.7	26.9	21.8	29.2	34.8	26.7	31.6

(a) Can you support a claim that mean rainfall from seeded clouds exceeds 25 acre-feet? Use $\alpha = 0.01$.

(b) Is there evidence that rainfall is normally distributed?

(c) Compute the power of the test if the true mean rainfall is 27 acre-feet.

(d) What sample size would be required to detect a true mean rainfall of 27.5 acre-feet if we wanted the power of the test to be at least 0.9?

(e) Explain how the question in part (a) could be answered by constructing a one-sided confidence bound on the mean diameter.

(f) Explain how the question in part (a) could be answered by using p-value.

Problem 8
The sodium content of thirty 300-gram boxes of organic corn flakes was determined. The data (in milligrams) are as follows:

131.15	130.69	130.91	129.54	129.64	128.77	130.72	128.33	128.24	129.65
130.14	129.29	128.71	129.00	129.39	130.42	129.53	130.12	129.78	130.92
131.15	130.69	130.91	129.54	129.64	128.77	130.72	128.33	128.24	129.65

(a) Can you support a claim that mean sodium content of this brand of cornflakes is 130 milligrams? Use $\alpha = 0.05$.

(b) Is there evidence that sodium content is normally distributed?

(c) Compute the power of the test if the true mean sodium content is 130.5 miligrams.

(d) What sample size would be required to detect a true mean sodium content of 130.1 milligrams if we wanted the power of the test to be at least 0.75?

(e) Explain how the question in part (a) could be answered by constructing a two-sided confidence interval on the mean sodium content.

(f) Explain how the question in part (a) could be answered by using p-value.

Problem 9
A researcher claims that at least 10% of all football helmets have manufacturing flaws that could potentially cause injury to the wearer. A sample of 200 helmets revealed that 16 helmets contained such defects.

(a) Does this finding support the researcher's claim? Use $\alpha = 0.01$.

(b) Find the P-value for this test.

Problem 10
The advertised claim for batteries for cell phones is set at 48 operating hours, with proper charging procedures. A study of 5000 batteries is carried out and 15 stop operating prior to 48 hours. Do these experimental results support the claim that less than 0.2 percent of the company's batteries will fail during the advertized time period, with proper charging procedures? Use a hypothesis-testing procedure with $\alpha = 0.01$.

Problem 11
Consider the following 75 observations.

2	3	1	0	0	0	1	1	1	0	1	1	2	1	1	2	0	0	1	1
2	4	0	1	1	2	3	1	0	0	0	0	1	2	0	0	2	2	2	2
2	1	0	0	2	0	0	1	0	0	2	3	1	2	0	4	1	1	0	2
3	1	1	0	0	2	1	3	0	2	1	1	1	0	0					

Based on this data, is a Poisson distribution an appropriate model?

Problem 12
Use the following 60 observations.

107.20	102.87	109.85	103.37	105.16	93.71	99.05	96.96	92.30	105.36
91.38	113.37	81.19	100.71	90.80	121.75	104.78	94.85	83.65	91.93
101.49	105.57	99.24	98.94	100.42	112.15	99.41	83.20	89.52	108.96
107.78	99.79	98.99	85.94	101.03	102.62	96.50	89.35	108.82	106.07
109.11	88.68	101.44	98.70	95.82	93.14	101.10	115.03	126.73	109.71
97.90	109.03	109.97	100.38	104.57	91.03	91.35	94.60	100.26	103.60

Based on this data, is a normal distribution an appropriate model? To check this do the following:

(a) Plot the normal probability plot for this data.

(b) Perform a chi square goodness of fit test?

(c) Use Shapiro-Wilk test.

Problem 13
Use the following 100 observations.

7.08	19.33	2.36	15.61	9.73	3.86	5.23	7.36	17.62	11.27
47.93	0.39	68.18	14.60	34.61	18.79	9.67	13.82	12.02	4.81
6.69	20.26	20.97	0.12	1.50	2.19	1.11	21.18	0.78	13.63
14.77	9.41	9.27	5.21	8.37	16.81	0.12	5.55	7.71	3.28
7.01	7.29	5.33	1.84	30.61	13.41	10.28	6.21	14.30	0.55
62.21	3.49	0.15	9.36	3.78	2.64	18.79	4.37	35.40	2.63
2.59	9.17	6.97	0.91	1.02	2.19	3.27	15.29	5.24	3.79
0.45	5.64	0.36	2.61	35.01	7.65	12.00	3.43	39.58	17.65
15.55	14.66	0.37	3.02	49.43	14.70	0.42	34.84	2.42	10.54
15.19	15.13	15.37	17.79	22.65	2.31	12.34	12.28	22.88	0.15

(a) Use normal probability plot, chi square goodness of fit test, and Shapiro-Wilk test to check normality of data.

(b) Does the exponential distribution seem to be a reasonable model for these data? Perform an appropriate goodness-of-fit test to answer this question.