

# Statistics 101 L – Homework 10

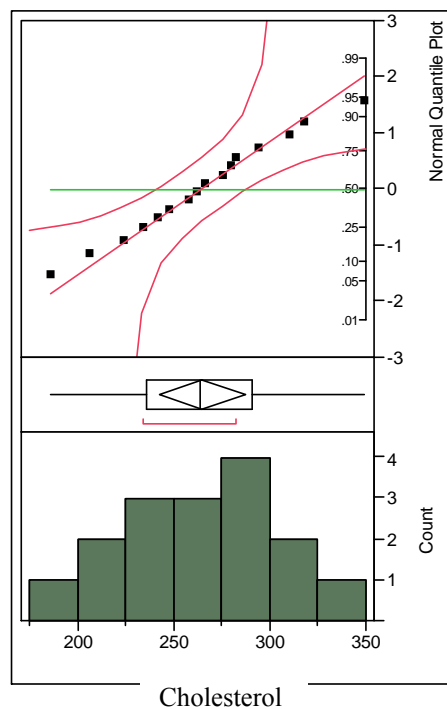
## Solution

### Assignment:

1. A major medical center in the Northeastern U.S. conducted a study looking at blood cholesterol levels and incidence of heart attack. Below are cholesterol levels for 16 randomly selected people who have had a heart attack.

242	318	224	310	186	294	276	280
266	282	262	248	206	234	349	258

$$n = 16, \bar{y} = 264.7, s = 42.12$$



- a) Verify that the conditions for statistical inference are met.

**Random sample condition:** The problem states that this is a random sample of 16 people who have had a heart attack.

**10% condition:** 16 people who have had a heart attack are less than 10% of all the people who have had a heart attack.

**Nearly normal condition:** The histogram is fairly symmetric (a bit skewed left) and mounded in the middle. The box plot is symmetric. The normal quantile plot has all of the data points falling close to the diagonal line representing a normal model. All of this indicates that the sample could have come from a population whose distribution follows a normal model.

- b) Construct a 95% confidence interval for the mean cholesterol for all people who have had a heart attack.

$$\bar{y} - t^* \left( \frac{s}{\sqrt{n}} \right) \text{ to } \bar{y} + t^* \left( \frac{s}{\sqrt{n}} \right) \text{ for } df = n - 1 = 15, t^* = 2.131$$

$$264.7 - 2.131 \left( \frac{42.12}{\sqrt{16}} \right) \text{ to } 264.7 + 2.131 \left( \frac{42.12}{\sqrt{16}} \right)$$

$$264.7 - 22.44 \text{ to } 264.7 + 22.44$$

$$242.26 \text{ to } 287.14$$

- c) Give an interpretation of the confidence interval in b).

**We are 95% confident that the mean cholesterol level for all people who have had a heart attack is between 242.26 and 287.14.**

- d) Is 240 a plausible value for the mean cholesterol level for all people who have had a heart attack? Support your answer by referring to the confidence interval you constructed in b).

**No, 240 is not a plausible value for the mean cholesterol level for all people who have had a heart attack. The value of 240 is not in the confidence interval.**

- e) Perform a test of hypothesis that the mean cholesterol level for all people who have had a heart attack is 240 against the alternative that it is greater than 240. Be sure to include the value of the test statistic, P-value, decision, reason for the decision and a conclusion.

$$H_0 : \mu = 240$$

$$H_A : \mu > 240$$

**where  $\mu$  is the mean cholesterol level of all people who have had a heart attack.**

$$t = \frac{(\bar{y} - \mu_0)}{\left( \frac{s}{\sqrt{n}} \right)} = \frac{(264.7 - 240)}{\left( \frac{42.12}{\sqrt{16}} \right)} = \frac{24.7}{10.53} = 2.346$$

One tail probability	0.10	0.05	0.025	P-value	0.01	0.005
df				↑		
15	1.341	1.753	2.131	▶ 2.346	2.602	2.947

**P-value is between 0.01 and 0.025.**

**Because the P-value is less than 0.05, we should reject the null hypothesis.**

**The mean cholesterol level of all people who have had a heart attack is greater than 240.**

2. The local grocery store packages different grades (based on fat content) of hamburger. Usually the leaner (less fat) grades are more expensive. A simple random sample of 12 one-pound packages of hamburger is obtained from the local store and tested for fat content. Below are the percentages of lean meat for each package.

90.7 94.1 91.7 92.3 91.6 93.2 91.4 92.4 92.1 93.4 92.9 93.0

- a) Calculate the sample mean percentage of lean meat and the sample standard deviation.

**The sample mean,  $\bar{y} = 92.4$ , the sample standard deviation,  $s = 0.9658$ .**

- b) The packages are labeled 93% lean. Use the summaries in a) to test the hypothesis  $H_0: \mu = 93$  against  $H_A: \mu < 93$ , where  $\mu$  is the population mean percentage lean. Be sure to calculate the appropriate test statistic and P-value, reach a decision and justify that decision, and state a conclusion.

$$t = \frac{(\bar{y} - \mu_0)}{\frac{s}{\sqrt{n}}} = \frac{(92.4 - 93)}{\frac{0.9658}{\sqrt{12}}} = \frac{-0.6}{0.2788} = -2.152$$

**The  $df = n - 1 = 11$ .**

One tail probability	0.10	0.05	P-value	0.025	0.01	0.005
			↑			
<b>11</b>	1.363	1.796	→ <b>2.152</b>	2.201	2.718	3.106

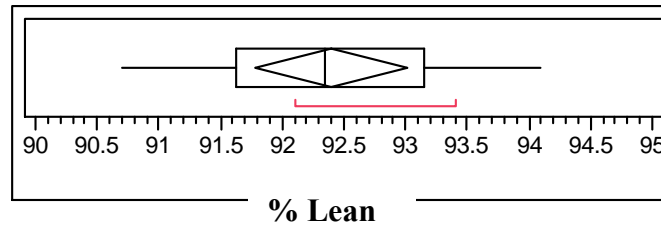
**The P-value is between 0.025 and 0.05. This P-value is small (less than 0.05), therefore we should reject the null hypothesis. The population mean percentage of lean is less than 93%.**

- c) Based on the test in b), are the packages labeled correctly? Explain briefly.

**The packages are labeled incorrectly. From the test of hypothesis we conclude that the mean percentage of lean is less than 93% so the packages should not be labeled as saying they contain 93% lean meat.**

- d) One of the conditions for the test in b) is that the random sample could have come from a population described by a normal model (the nearly normal condition). Use the sample data to construct a box plot. Does the shape of the box plot support the nearly normal condition? Explain briefly.

**The box plot (see next page) is symmetric which agrees with the nearly normal condition because the normal model is a symmetric distribution.**



<b>100.0%</b>	<b>maximum</b>	<b>94.100</b>
<b>75.0%</b>	<b>quartile</b>	<b>93.150</b>
<b>50.0%</b>	<b>median</b>	<b>92.350</b>
<b>25.0%</b>	<b>quartile</b>	<b>91.625</b>
<b>0.0%</b>	<b>minimum</b>	<b>90.700</b>

3. Earlier this semester (Homework #2) we looked at the body temperature ( $^{\circ}$  F) for 130 randomly selected adults. The data appear in the JMP file BodyTemp.JMP on the course web page. Use JMP to analyze the distribution of body temperatures and answer the following questions. Be sure to turn in a copy of the JMP output with your homework.
- a) Test the hypothesis that the population mean body temperature of adults is  $98.6^{\circ}$  F against the alternative that the population mean body temperature is less than  $98.6^{\circ}$  F. Be sure to include all steps for the test of hypothesis including verifying that the conditions for performing the test are met.

**Step 1:**  $H_0 : \mu = 98.6$  where  $\mu$  is the mean body temperature of all adults.  
 $H_A : \mu < 98.6$

**Step 2:**

**Random sample condition:** The problem states that we have a random sample.

**10% condition:** 130 adults are less than 10% of all adults.

**Nearly normal condition:** The histogram of body temperature is fairly symmetric and mounded in the middle. The box plot of body temperature is symmetric with two low and two high potential outliers. The normal quantile plot of body temperature has all of the values falling very close to the diagonal line representing a normal model.

All of these observations support that the sample could have come from a distribution described by a normal model.

**Step 3:**

Value of the test statistics is  $t = -5.4548$ , with an associated P-value of  $< 0.0001$ .

**Step 4:**

Because the P-value is very small (less than 0.05), we reject the null hypothesis.

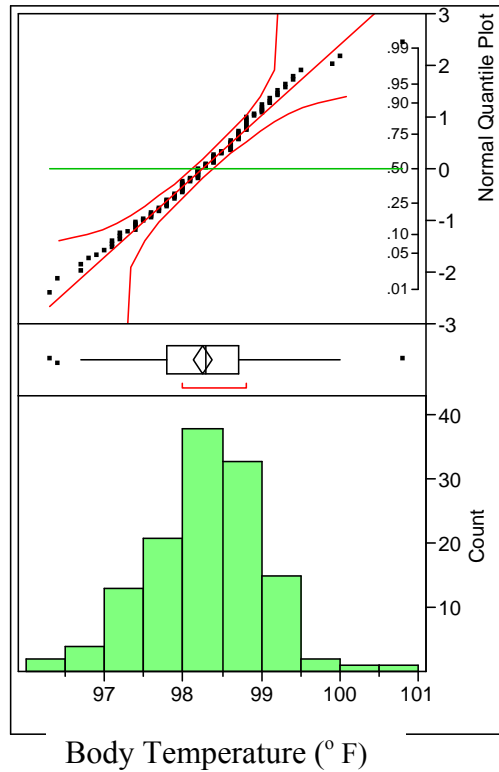
**Step 5:**

The mean pulse of all adults is less than  $98.6^{\circ}$  F.

- b) Construct a 90% confidence interval for the population mean body temperature of adults.

**A 90% confidence interval for the population mean body temperature of adults goes from 98.14 ° F to 98.36 ° F.**

**JMP Output for the Analysis of Body Temperature Data**



**Quantiles**

100.0%	maximum	100.80	Mean	98.249231
75.0%	quartile	98.70	Std Dev	0.7331832
50.0%	median	98.30	Std Err Mean	0.0643044
25.0%	quartile	97.80	upper 95% Mean	98.376459
0.0%	minimum	96.30	lower 95% Mean	98.122003
			N	130

**Confidence Intervals**

Parameter	Estimate	Lower CI	Upper CI	1-Alpha
Mean	98.24923	98.14269	98.35577	0.900

**Test Mean=value**

		t Test
Hypothesized Value	98.6	Test Statistic
Actual Estimate	98.2492	Prob >  t
df	129	Prob > t
Std Dev	0.73318	Prob < t