Inference for Population Mean

1. List all the assumptions that need met in order for a t distribution based inference to be valid.

Correct:
data must be from a random sample from a large population
observations within the sample must be independent
n small - population distribution must be approximately normal
n large - population distribution need not be approximately normal (CLT kicks in)

2. The blood pressure (average of systolic and diastolic measurements) of each of 38 randomly selected persons was measured. The average was 94.5 mm Hg and the standard deviation 8.0497 mm Hg.

a. Construct a 95% confidence interval for the true mean blood pressure.

Correct:
CI for population mean is
\[ \bar{x} \pm t_{a/2, df=n-1} \frac{s}{\sqrt{n}} \]
Find the t critical point using the “qt” function in R. In this case, confidence is 95% = 0.95, so \( \alpha = 0.05 \). Then,
\[ t_{a/2, df=n-1} = t_{0.05, df=38} \approx 2.026 \]
which says “give me the value in the t distribution with 37 degrees of freedom that has 0.025 above it”. The t value that has 0.025 above it has 0.975 below it (so find the 97.5th percentile).
\[ > qt(0.975, 37) \]
\[ [1] 2.026192 \]
Then, the interval becomes
\[ 94.5 \pm 2.026192 \frac{8.0497}{\sqrt{38}} \]
\[ (91.85, 97.15) \text{ or } 91.85 < \mu < 97.15 \text{ mm Hg} \]

b. Interpret the interval you just computed in part (a).

Correct:
We are 95% confident that true mean blood pressure for all people is at least 91.85 mm Hg and at most 97.15 mm Hg.

c. Use the QQplot to comment on whether it appears the assumption of normality has been violated.

Correct:
Goal: We need to check the assumption that \( \bar{x} \) has a normal distribution. We get that one of two ways:

1) Either the data that created \( \bar{x} \) is from a normal distribution
2) Or the sample size is large enough for the CLT to “kick in”

Step 1 – Check the QQ plot
I see no systematic departure from the line, except for that little blip in the upper end of the distribution.

Step 2 – Can we call the CLT?
Mainly, the points form a linear pattern and our sample size is 38, so the CLT should kick in anyway.

Conclusion: I do not see evidence against normality of \( \bar{x} \) - it should be just fine to proceed with this t distribution based procedure.
3. The wing lengths (in.) of \( n = 7 \) randomly chosen fledglings from an avian nesting site are
3.14 2.37 2.94 3.60 1.70 3.99 1.85

a. Find a 90% confidence interval for the true mean wing length.

```r
> winglengths<-c(3.14,2.37,2.94,3.60,1.70,3.99,1.85)
> t.test(winglengths,conf.level=0.90)

One Sample t-test
data:  winglengths
t = 8.5605, df = 6, p-value = 0.0001395
alternative hypothesis: true mean is not equal to 0
90 percent confidence interval:
 2.163316 3.433826
sample estimates:
  mean of x
2.798571
```

(2.16, 3.43) -or- 2.16 < \( \mu \) < 3.43 in.

b. Interpret the interval you just computed in part (a).

**Correct:**
We are 90% confident that the true mean wing length for all fledglings* is at least 2.16 inches and at most than 3.43 inches.  
*in the population to which the ones used in this study belong

c. Use the QQplot to comment on whether it appears the assumption of normality has been violated.

**Correct:**
Goal: We need to check the assumption that \( \bar{x} \) has a normal distribution. We get that one of two ways:

3) Either the data that created \( \bar{x} \) is from a normal distribution
4) Or the sample size is large enough for the CLT to “kick in”

**Step 1 – Check the QQ plot**
This QQplot is not perfectly linear. The second and third points are falling below the line and I see a hint of an "S" shape forming in this plot, which could be indicative of a “light-tailed” distribution. These data may not have been drawn from a normal distribution.

**Step 2 – Can we call the CLT?**
I only have 7 points, so the bottom line is - I need more data points to draw a conclusion here.

**Conclusion about \( t \) distribution based inference:**
With only 7 data points on our QQplot, we cannot tell whether this departure from the line is systematic and we certainly don't have enough data points to invoke the CLT, \( \bar{x} \) might not be normally distributed, so proceed with caution.
4. For each of the following situations, suppose $H_0: \mu = 0$ is being tested against $H_A: \mu \neq 0$. Find the P-value and state whether or not $H_0$ would be rejected.

a. $t_s = 2.2$ with $df = 23; \alpha = 0.02$

Draw a picture! Then, use R to get the P-value.

$$P = Pr\{tdf=23 < -2.2\} + Pr\{tdf=23 > 2.2\}$$

$$> \text{pt}(-2.2, 23) + (1- \text{pt}(2.2, 23))$$

$$[1] 0.03812092$$

Or double the lower tail area (the two tail areas are equal)

$$> 2*\text{pt}(-2.2, 23)$$

$$[1] 0.03812092$$

Or double the upper tail area

$$> 2*(1- \text{pt}(2.2, 23))$$

$$[1] 0.03812092$$

P > $\alpha$ (0.0381 > 0.02), so fail to reject $H_0$ (do not find significant evidence for $H_A$)

b. $t_s = -2.7$ with $df = 27; \alpha = 0.05$

Correct:

TI-84

$$P = Pr\{tdf=27 < -2.7\}$$

$$> \text{pt}(-2.7, 27)$$

$$[1] 0.01182146$$

P < $\alpha$ (0.0118 < 0.05), so reject $H_0$ (find significant evidence for $H_A$)

5. For each of the following situations, suppose $H_0: \mu = 0$ is being tested against $H_A: \mu < 0$. Find the P-value and state whether or not $H_0$ would be rejected.

a. $t_s = -2.9$ with $df=15 \alpha = 0.05$

Correct:

TI-84

$$P = Pr\{tdf=15 < -2.9\}$$

$$> \text{pt}(-2.9, 15)$$

$$[1] 0.00549734$$

P < $\alpha$ (0.0055 < 0.05), so reject $H_0$ (find significant evidence for $H_A$)

b. $t_s = 2.9$ with $df=15 \alpha = 0.05$

Correct:

TI-84

$$P = Pr\{tdf=15 < 2.9\}$$

$$> \text{pt}(2.9, 15)$$

$$[1] 0.9945027$$

P > $\alpha$ (0.9945 > 0.05), so fail to reject $H_0$ (do not find significant evidence for $H_A$)
6. In recent years, female athletes have been identified as a population particularly at risk for developing eating disorders. The article "Disordered Eating in Female Collegiate Gymnasts" (J. of Sport and Exercise Psychology (1993): 424-436) reported that for a sample of 17 gymnasts who were classified as binge eaters, the sample mean difference between current weight and ideal weight was 6.65 pounds with the sample standard deviation of this difference as 2.97 pounds. Suppose that the true mean weight difference for female gymnasts with normal eating habits is 4.5 pounds.

Is there evidence to suggest that the mean difference between current and ideal weight is significantly more than 4.5 pounds for female gymnasts who are binge eaters? You may proceed at the 0.05 significance level as though the assumptions have been checked and deemed acceptable.

Correct:

(1) Set $\alpha = 0.05$.
(2) $H_0: \mu = 4.5$
$H_A: \mu > 4.5$
(3) $t_s = \frac{6.65 - 4.5}{\frac{2.97}{\sqrt{17}}} = 2.985$
(4) $P = Pr\{T_{df=16} \geq 2.985\} = 1 - pt(2.985,16) = 0.0044$
(5) Since $P \leq \alpha$, we reject $H_0$ (find significant evidence to conclude $H_A$)
(6) There is significant evidence (at the 0.05 significance level) to conclude that the true mean difference between current and ideal weight for female gymnasts who are binge eaters is more than 4.5 pounds.

7. Human beta-endorphin (HBE) is a hormone secreted by the pituitary gland under conditions of stress. A researcher conducted a study to investigate whether a program of regular exercise might affect the resting (unstressed) concentration of HBE in the blood. He measured blood HBE levels, in January and again in May, in ten participants in a physical fitness program. The results were as shown in the table.

<table>
<thead>
<tr>
<th>Participant</th>
<th>HBE Level (pg/mL)</th>
<th>January</th>
<th>May</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>42</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>47</td>
<td>29</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>37</td>
<td>9</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>33</td>
<td>26</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>70</td>
<td>36</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>54</td>
<td>38</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>27</td>
<td>32</td>
<td>-5</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>41</td>
<td>33</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>18</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>37.8</td>
<td>24.8</td>
<td>13.0</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td>17.6</td>
<td>10.9</td>
<td>12.4</td>
</tr>
</tbody>
</table>

a. Conduct a test of hypothesis, using $\alpha = 0.05$, to investigate whether there is an effect of exercise on HBE levels (i.e. Is there a difference in resting HBE levels between January and May (resting HBE level possibly perturbed by exercise)). (Hint: You need to use only the values in the right-hand column and your $H_0$ will be $\mu = 0.$)
(1) Set $\alpha = 0.05$.
(2) $H_0: \mu = 0$
$H_A: \mu \neq 0$

```r
> t.test(diff,alternative="two.sided")
```

One Sample t-test

data:  diff
t = 3.3151, df = 9, p-value = 0.00901
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  4.129062 21.870938
sample estimates:
mean of x
  13

(3) $t_s = 3.3151$
(4) $P = 0.0090$

(5) Since $P \leq \alpha$, we reject $H_0$ (find significant evidence for $H_A$)
(6) There is significant evidence (at the 0.05 level) to conclude that the true mean resting HBE level is different after five months of participating in a physical fitness program (like the one used in this study).

b. Use the QQplot to comment on whether it appears the assumption of normality has been violated.

**Correct:**

I see no systematic departure from the line on this QQplot. It's not perfectly linear, so normally we'd like to invoke the CLT at this point, which may be feasible in this case since we have 10 observations that aren't too far from looking like they are from a normal distribution.

**Conclusion:** I see no glaring evidence against normality and the possible slight deviation in the x’s from normality will likely be corrected for (when thinking about the distribution of $\bar{x}$) by the CLT - it is probably OK to proceed.

8. Jose is studying a population of hospital patients. He rejects the null hypothesis that the mean age of the patients is 40 years in favor of the alternative hypothesis that the mean age is greater than 40 years. Which of the following statements is correct?
Jose may have made a Type I error.
Jose has definitely made a Type I error.
Jose may have made a Type II error.
Jose has definitely made a Type II error.
Jose has definitely not made either a Type I or a Type II error.