1. It has long been reported that normal human body temperature follows a normal distribution with a mean of 98.6 degrees Fahrenheit. There is some debate in the medical field, however, about this claim. A group of medical researchers measured the body temperature of a random sample of 18 healthy individuals. After the researchers collected their data, they entered the data into a statistical software package and ran the appropriate statistical test. The output generated from this test is printed below. Based on this output, what should they conclude and why?

**T-Test of the Mean**

Test of mu = 98.60 vs mu not = 98.60

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEMP</td>
<td>18</td>
<td>98.2167</td>
<td>0.68363</td>
<td>0.16113</td>
<td>-2.39</td>
<td>0.029</td>
</tr>
</tbody>
</table>

Since the p-value is 2.9%, we have moderately strong evidence against H₀. That is, we have moderately strong evidence against the claim that the mean normal body temperature is 98.6 degrees of Fahrenheit. At the 5% level, we would reject the claim that the mean body temperature is 98.6 °F, but at the 1% level we would not reject it.

2. A recent article reported that college students tend to watch an average of 12 hours of television per week. Cathy, a college senior, decides to see if perhaps her school is different. She randomly selects 79 students from her college and asks each student to record the number of hours he/she spends watching television (on average) each week. Cathy finds that her sample of students watches an average of 9.49 hours of television per week, with a standard deviation of 8.42 hours.

a. Compute a 95% confidence interval for the mean number of hours of television watched per week at Cathy’s college.

Conditions: random sample, checked. n > 30, checked.

95% CI: with \( \bar{x} = 9.49 \), \( s = 8.42 \), \( t^* = 1.99 \)

\[
\bar{x} \pm t^* \frac{s}{\sqrt{n}} = 9.49 \pm 1.99 \frac{8.42}{\sqrt{79}} = (7.60, 11.38)
\]

Cathy should be 95% that the mean number of hours of television watched per week at Cathy’s college is between 7.60 hours and 11.38 hours.

b. Should Cathy reject the hypothesis that students at her school watch an average of 12 hours of television per week? Set up appropriate hypotheses, determine the value of the test statistic, compute the p-value, and draw the appropriate conclusion.

\[
H_0: \mu = 12 \quad Ha: \mu \neq 12
\]

Test statistic: \( t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{9.49 - 12}{8.42/\sqrt{79}} = -2.65 \)

[Calculator: STAT→TESTS→8: TInterval]

[Calculator: STAT→TESTS→2: T-Test]
p-value: 0.0097

Since the p-value is less than 5% (I use $\alpha = 5\%$ because we calculated a 95% CI in part $\alpha = 5\% + 95\% = 100\%$), we can reject the null hypothesis. That is, we can reject the claim that at Cathy’s school students watch an average of 12 hours of television per week.

c. Explain the connection between your answers to parts (a) and (b).

Same conclusion in both parts. In part (a) we can see that 12 is not in the 95% confidence interval, so it’s not a plausible value. Thus, we can reject the null hypothesis.

3. Ecologists periodically measure the concentration of coliform bacteria at an urban reservoir. Each time, water samples are taken from approximately 25 different locations within the reservoir. The distribution of bacteria counts in the past has followed a normal distribution with a mean of 2.8 ppm (parts per million) and an s.d. of 0.45 ppm. In the latest set of readings, the mean count for the 27 water samples was 2.93 ppm. Thus there is some evidence that the bacteria concentration in the reservoir has increased.

a. If the distribution of bacteria counts is still normal with mean 2.8 ppm and sd 0.45 ppm, show that the chance of a sample mean reading of 27 new measurements being as high as 2.93 (or higher) is 6.7%.

$$ z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.93 - 2.8}{0.45/\sqrt{27}} = 1.5 $$

The probability that that mean reading of the 27 new measurements is 2.93 or higher is 0.067. (You can get this from the z-table, or using your calculator: DISTR $\bar{x}$:normalcdf(1.5,999999)

b. Set up hypotheses for testing whether the bacteria concentration in the reservoir has increased. What does the 6.7% from part (a) represent?

$H_0$: $\mu = 2.8$  $H_a$: $\mu > 2.8$

Test statistic: $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{2.93 - 2.8}{0.45/\sqrt{27}} = 1.5$

p-value: 0.067

c. What is the strength of the evidence that the reservoir is more contaminated than before?

(i) very little   (ii) moderate, but not conclusive   (iii) fairly strong   (iv) very strong
d. What would you conclude at significance level 5%?

At the 5% level, since the p-value is greater than 5%, we cannot reject the null hypothesis. That is, we cannot reject the claim that mean bacteria is 2.8 ppm (parts per million).

e. What would you conclude at significance level 10%?

At the 10% level, since the p-value is less than 10%, we can reject the null hypothesis. That is, we have enough evidence to reject the claim that mean bacteria is 2.8 ppm (parts per million). We believe that the bacteria concentration has increased, i.e. the reservoir is more contaminated than before.

4. An experimental program similar to Head Start is interested in knowing if their program has an impact on the reading readiness of first graders. Assume that the mean Reading Readiness score for all first graders with the existing program is 100 and the population standard deviation is 25. A sample of 40 first graders who have been through the new program had a mean of 104.

a. Draw and label a graph of the sampling distribution of $\bar{x}$, where $\bar{x}$ is the mean score of a sample of 40 first graders under the existing program.

\[
\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{25}{\sqrt{40}} = 3.95 \approx 4
\]

\[
\mu_{\bar{x}} = \mu = 100
\]

b. Mark the location of $\bar{x} = 104$ and shade in the region corresponding to the p-value of a test of $H_0$: $\mu = 100$ vs. $H_a$: $\mu > 100$, where $\mu$ is the mean Reading Readiness score that would be achieved if all first graders used the new experimental program.

The upper tail should be shaded.

c. Based only on your sketch, does there appear to be statistically significant evidence that the new program is better than the existing one? Explain briefly.

No, it does not appear to be statistically significant evident that the new program is better than the existing one because our sample result, the sample mean, is just about one standard deviation above the mean.

d. Determine the p-value and determine whether the improvement in mean score for the sample group is statistically significant at level 5%.

The p-value is the area of the shaded part: since the sample mean is about one standard deviation above the mean, $z = 1$. Then using either the z-table, or the calculator, we get the probability: 0.16.
Conclusion: At the 5% level the mean score for the sample group is not statistically significant. Since the p-value > 5%, we cannot reject the null hypothesis. That is, we don’t have enough evidence to reject the claim that the mean score is 100.

5. The Food and Drug Administration (FDA) has a maximum upper limit of 12 mg for the mean nicotine content in cigarettes. An FDA evaluator took a random sample of 10 cigarettes from a new brand and found the mean nicotine content of the sample to be 13 mg, with a standard deviation of 2 mg.

a. Based on this sample, should the FDA conclude that the average nicotine level of the new brand is not acceptable? Perform the appropriate hypothesis test—set up hypotheses, determine the p-value, and draw an appropriate conclusion.

\[ H_0: \mu = 12mg \quad H_a: \mu > 12mg \]

Conditions: random sample checked. Since the sample size is 10, and the population distribution is unknown, we need to assume that it’s normally distributed.

Test statistic: \( t = 1.58 \)

p-value: 0.074

Since the p-value is about 7.4%, we have some evidence against the claim that the mean nicotine level is 12mg, but the evidence is not very strong.

b. Do the results of part (a) indicate that the mean nicotine level of the new cigarette brand is acceptable, that is, not higher than 12 mg?

The results show some evidence, but not strong evidence. At the 5% level we cannot reject the null hypothesis, but at the 10% level we can.

c. Suppose the FDA decided to take a larger sample, and still found the mean nicotine content of the sample to be 13 mg. How would the p-value compare to the one you found in part (a)? How would the FDA’s conclusion be likely to change?

For a larger sample, the test statistic will be higher, so the p-value will be smaller. Then, the FDA probably could reject the null hypothesis and conclude that the mean nicotine level of the new cigarettes is unacceptable.