Warm Up

A random sample of 13 men and 19 women in a college class reported their grade point averages (GPAs). Histograms were created for both men and women and were approximately symmetric and unimodal. A woman in the class says that she believes that college women tend to have higher GPAs than do college men. Her summary statistics are below.

Does this sample support her claim?

Test an appropriate hypothesis and state your conclusion.

<table>
<thead>
<tr>
<th></th>
<th>$\bar{y}$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>2.898</td>
<td>0.583</td>
</tr>
<tr>
<td>Women</td>
<td>3.330</td>
<td>0.395</td>
</tr>
</tbody>
</table>
**Warm Up**

**Think**

$H_0: \mu_W = \mu_m$

$H_A: \mu_W > \mu_m$

**Conditions:**

✓ SRS? Stated!

✓ Approx. Normal: The samples are pretty small, but the histograms are stated to be symmetric and unimodal

✓ Independence: It’s safe to say the women and men were independent groups. Both are certainly less than 10% of the total men and women taking the course.

**Show**

Method: Two-Sample t-test

$$t = \frac{\bar{y}_W - \bar{y}_M}{SE(\bar{y}_W - \bar{y}_M)} = \frac{0.432 - 0}{0.185} = 2.335$$

$$P(t_{19.44} > 2.335) = 0.0152$$

**Tell**

The P-value is low, so we will reject the null hypothesis. There is evidence that college women have a higher mean GPA than do college men.
On the AP Statistics exam you have 6 Free-Response Questions. #1–5 tend to be very straightforward, and will be material that you’ve seen. #1 tends to be the easiest 😊

The last FRQ (#6) is called an **INVESTIGATIVE TASK**. This is usually a question that you’ve never seen and it takes the material to the next level. It usually takes ideas you’ve seen, and puts a spin on it. It gets those creative juices flowing!

This ONE problem is **worth 25% of your score** on the AP exam. **Always attempt it!**
A consumer organization was concerned that an automobile manufacturer was misleading customers by overstating the average fuel efficiency (measured in miles per gallon, or mpg) of a particular car model. The model was advertised to get 27 mpg. To investigate, researchers selected a random sample of 10 cars of that model. Each car was then randomly assigned a different driver. Each car was driven for 5,000 miles, and the total fuel consumption was used to compute mpg for that car.

(a) Define the **parameter of interest** and state the **null and alternative hypotheses** the consumer organization is interested in testing.

**Part (a):**

The parameter of interest is \( \mu = \) population mean miles per gallon (mpg) of a particular car model.

The null and alternative hypotheses are as follows:

\[ H_0: \mu = 27 \]
\[ H_\alpha: \mu < 27 \]
One condition for conducting a one-sample t-test in this situation is that the mpg measurements for the population of cars of this model should be normally distributed. However, the boxplot and histogram shown below indicate that the distribution of the 10 sample values is skewed to the right.

(b) One possible statistic that measures skewness is the ratio \( \frac{\text{sample mean}}{\text{sample median}} \)

What values of that statistic (small, large, close to one) might indicate that the population distribution of mpg values is skewed to the right? **Explain.**

**Part (b):**

If the distribution is right-skewed, one would expect the mean to be greater than the median. Therefore the ratio \( \frac{\text{sample mean}}{\text{sample median}} \) should be large (greater than 1).
Part (c):

Because we are testing for right-skewness, the estimated $p$-value will be the proportion of the simulated statistics that are greater than or equal to the observed value of 1.03. The dotplot shows that 14 of the 100 values are more than 1.03. Because this simulated $p$-value (0.14) is larger than any reasonable significance level, we do not have convincing evidence that the original population is skewed to the right and conclude that it is plausible that the original sample came from a normal population.
FRAPPY 09  
**Average:** 1.32 (out of 4)  
**St. Dev:** 1.09

<table>
<thead>
<tr>
<th>Minimum</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>24</td>
<td>25.5</td>
<td>28</td>
<td>32</td>
</tr>
</tbody>
</table>

Choosing only from the summary statistics in the table, define a formula for a different statistic that measures skewness.

What values of that statistic might indicate that the distribution is skewed to the right? Explain.

**Part (d):**

One possible statistic is \( \frac{\text{maximum} - \text{median}}{\text{median} - \text{minimum}} \)

If the distribution is right-skewed, one would expect the distance from the median to the maximum to be larger than the distance from the median to the minimum; thus the ratio should be greater than 1.
MATCHED PAIRS VS. TWO SAMPLE

ESSENTIAL QUESTION
How can I discern between Two-Samples and Matched Pairs?

Coming Up

Project Due – Fri Mar 10

Chapter 23-25 Test
FRIDAY!
You may either turn in your REVIEW or Socrative!

It's a rather interesting phenomenon. Every time I press this lever, that post-graduate student breathes a sigh of relief.
RECALL:

If you have a matched-pair design (or a _______ experiment) you find the _______ difference in the two samples, and then treat those as _______ sample.
Before you took this course, you probably heard many stories about Statistics courses. Oftentimes parents of students have had bad experiences with Statistics courses and pass on their anxieties to their children.

**To test whether actually taking AP Statistics decreases students' anxieties about Statistics**, an AP Statistics instructor (yours truly) gave a test to rate student anxiety at the beginning and end of my 2015-2016 course. Anxiety levels were measured on a scale of 0-10. Here are the data for 16 randomly chosen students:

<table>
<thead>
<tr>
<th>Pre-course anxiety level</th>
<th>7</th>
<th>6</th>
<th>9</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>5</th>
<th>7</th>
<th>6</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>3</th>
<th>4</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-course anxiety level</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Difference (Post – Pre)</td>
<td>-3</td>
<td>-3</td>
<td>-2</td>
<td>-2</td>
<td>-2</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
null: $H_0: \mu_d = 0$  \\* $\mu_d$ — True mean difference (Post-Pre) in anxiety scores \\* $H_A: \mu_d < 0$

1. **Random?** Stated randomly selected students

2. **Approx Normal?** Differences appear approx. symmetric

\[ t_{15} = \frac{\bar{d} - 0}{SE(\bar{d})} = \frac{-1.125 - 0}{0.287} = -3.92 \]

\[ P\text{-value is } P(t_{15} < -3.82) = 0.0007 \]

Since $0.0007 < 0.05$, I reject the null. There is enough evidence to suggest the mean difference in anxiety levels is below 0. We have strong evidence that taking AP Statistics class reduces the anxiety level of students.
Scoring Guide

\[ H_0 : \mu_d = 0 \]
\[ H_A : \mu_d < 0 \]

\( \mu_d \) — True mean difference (Post-Pre) in anxiety scores

**Matched-Pair t-Interval**

I’m 90% confident that the true difference in anxiety levels after taking AP Statistics is between -1.628 and -0.622 points.
Investigators at the U.S. Department of Agriculture wished to compare methods of determining the level of *E. coli* bacteria contamination in beef. Two different methods (A and B) of determining the level of contamination were used on each of the ten randomly selected specimens of a certain type of beef. The data obtained, in millimicrobes/liter of ground beef, for each of the methods are shown in the table below. *(FRAPPY 2007)*

<table>
<thead>
<tr>
<th>Specimen</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>22.7</td>
<td>23.6</td>
<td>24.0</td>
<td>27.1</td>
<td>27.4</td>
<td>27.8</td>
<td>34.4</td>
<td>35.2</td>
<td>40.4</td>
<td>46.8</td>
</tr>
<tr>
<td>B</td>
<td>23.0</td>
<td>23.1</td>
<td>23.7</td>
<td>26.5</td>
<td>26.6</td>
<td>27.1</td>
<td>33.2</td>
<td>35.0</td>
<td>40.5</td>
<td>47.8</td>
</tr>
</tbody>
</table>

Is there a **significant difference** in the mean amount of *E. coli* Bacteria detected by the two methods for this type of beef? *Provide a statistical justification to support your answer.*
Since $0.1793 > 0.05$, I fail to reject the null. There is NOT enough evidence to suggest there is a difference in the mean amount of E. Coli in beef detected by the two methods.

\[ H_0 : \mu_d = 0 \]
\[ H_A : \mu_d \neq 0 \]

\[ \mu_d \] — True mean difference (A-B) between Method A and B

**CONDITIONS**

1. **Random?** Stated
   
   10 beef samples were random

2. **Approx Normal?**
   
   Differences appear approx. symmetric

**METHOD: MATCHED-PAIR T-TEST**

\[ \overline{x}_d = 0.29 \]
\[ s_d = 0.629727 \]

\[ t = \frac{0.29 - 0}{0.629727/\sqrt{10}} = \frac{0.29}{0.199137} = 1.46 \]

\[ P\text{-value} = 0.1793 \]
Because most walkers have tray tables in front that block babies' views of their feet, child psychologists have begun to question whether walks affect infants' cognitive development. One study compared mental skills of a random sample of those who used walkers with a random sample of those who never used walkers. Mental skill scores averaged 113 for 54 babies who used walkers (standard deviation of 12) and 123 for 55 babies who did not use walkers (standard deviation of 15).

a) Is there evidence that the mean mental skill of babies who use walkers is different from the mean mental skill score of babies who do not use walkers? Explain your answer.
a. **part 1**: States a correct pair of hypotheses

\[ H_0: \mu_W = \mu_N \quad \quad \quad H_0: \mu_W - \mu_N = 0 \]

**OR**

\[ H_a: \mu_W \neq \mu_N \quad \quad \quad H_a: \mu_W - \mu_N \neq 0 \]

where \( \mu_W \) is the mean mental skill score for babies who used walkers and \( \mu_N \) is the mean for those who did not. Nonstandard notation must be explained. Hypotheses about statistics (e.g. \( \bar{x} \) or \( \hat{p} \)) are unacceptable.

**part 2**: Identifies a correct test (by name or by formula), and **checks** appropriate assumptions.

Note: Problem states that samples are random samples, so this does not need to be addressed in the assumptions.

Independent samples t test. Assumptions: large sample or normal population distributions. Check: OK, because, for example, \( n_1 \& n_2 \geq 30 \).
**Example 3**

**Part 4:** Stating a correct conclusion in the context of the problem, using the result of the statistical test (i.e., linking the conclusion to the result of the hypothesis test).

Reject the null hypothesis because P-value is less than stated $\alpha$ (or because P-value is very small, or because test statistic falls in the rejection region). There is convincing evidence that the mean mental score of babies who used walkers is different from the mean score for babies who did not use walkers.

If both an $\alpha$ and a P-value are given, the linkage is implied. If no $\alpha$ is given, the solution must be explicit about the linkage by giving a correct interpretation of the P-value or explaining how the conclusion follows from the P-value.

If the P-value in part 3 is incorrect but the conclusion is consistent with the computed P-value, part 4 can be considered as correct.
Because most walkers have tray tables in front that block babies' views of their feet, child psychologists have begun to question whether walks affect infants' cognitive development. One study compared mental skills of a random sample of those who used walkers with a random sample of those who never used walkers. Mental skill scores averaged 113 for 54 babies who used walkers (standard deviation of 12) and 123 for 55 babies who did not use walkers.

b) Suppose that a study using this design found a statistically significant result. Would it be reasonable to conclude that using a walker causes a change in mean mental skill score? Explain your answer.

No. This was an observational study, and a causal relationship can not be inferred from an observational study.

- It is sufficient to say any of:
  - "no; observational study" (or "no; not controlled experiment").
  - "no; no randomization in grouping" or "no; parents choose which babies use walkers".
  - “no” and then cite a plausible confounding variable and indicate how it is confounded with the formation of the groups.
- It is not sufficient to either:
  - merely mention lurking and/or confounding variables without indicating how they are confounded with the formation of the groups.
  - mention a causal factor which is a treatment “side effect”, e.g. that walkers may contain plastics which are toxic to children.
Homework

time remaining in class