HOMEWORK 19  Due: next class  5/3

1. Tell in each of the following instances whether the study uses an independent samples or a matched pairs design.

a. A survey is conducted of teens from inner city schools to estimate the proportion who have tried drugs. A similar survey is conducted of teens from suburban schools.

   Independent samples

b. A psychologist measures the response times of subjects under two stimuli; each subject is observed under both of the stimuli, in a random order.

   Matched pair design

c. An agronomist compares the yields of two varieties of soybean by planting each variety in 10 separate plots of land (a total of 20 plots).

   Independent samples

d. Lung cancer patients admitted in a hospital over a 12 month period are each matched with a non-cancer patient by age, sex, and race. To determine whether or not smoking is a risk factor for lung cancer, it is noted for each patient if he or she is a smoker.

   Matched pair design

e. An advertising agency has come up with two different TV commercials for a household detergent. To determine which one is more effective, a test is conducted in which a sample of 100 adults is randomly divided into two groups. Each group is shown a different commercial, and the people in the group are asked to score the commercial.

   Matched pair design

f. Twenty-five people have their cholesterol measured before eating a Big Mac and again after eating it. On average, does eating a Big Mac increase cholesterol?

   Matched pair design

f. What is the difference in average salaries for high school graduates and college graduates?

   Independent samples

h. In fifty married couples, the husband and wife each separately take the same test of marital satisfaction. Is there a difference, on average, between the scores of husbands and wives?

   Matched pair design
2. A mathematics test was given to a random sample of 1000 17-year-old students in 1978, and again to another random sample of 1000 students in 1992. The mean score in 1978 was 300.4. In 1992 it was 306.7. Is this 6.3 point difference evidence of a real difference, or likely just a chance variation?

a. State the null and alternative hypotheses.

\[ H_0: \mu_{1978} - \mu_{1992} = 0 \quad H_a: \mu_{1978} - \mu_{1992} \neq 0 \]

b. Carry out the appropriate hypothesis test. Use the following additional information: s.d.(1978) = 34.9; s.d.(1992) = 30.1

\[
t = \frac{(\bar{x}_{1978} - \bar{x}_{1992}) - 0}{\sqrt{\frac{s_{1978}^2}{n_{1978}} + \frac{s_{1992}^2}{n_{1992}}}} = \frac{(300.4 - 306.7) - 0}{\sqrt{\frac{34.9^2}{1000} + \frac{30.1^2}{1000}}} = -4.32
\]

Or you can use your calculator: TESTS \rightarrow 2-SampTTest

p-value: \(1.618 \times 10^{-5}\)

c. Based on your results in part b, what can you conclude? Is there a significant difference between average math scores in 1978 vs. 1992?

Since the p-value is very small, we have strong evidence against the null hypothesis. We can reject the claim that the difference in the mean scores of 1978 and 1992 is zero. There is a significant difference between average math scores in 1978 vs. 1992.

d. The 95% confidence interval for the difference between the mean scores is (-9.158, -3.442).

Based on only this information, what can you conclude? Is there a significant difference between average math scores in 1978 vs. 1992?

Since the claimed value, zero, is not inside the confidence interval, it’s not a plausible value. We can reject the null hypothesis. Same conclusion as in part c.

3. A study was carried out to investigate the effectiveness of a treatment. 1000 subjects participated in the study, with 500 being randomly assigned to the “treatment group” and the other 500 to the “control (or placebo) group”. A statistically significant difference was reported between the responses of the two groups (\(P < .005\)). Thus, we can conclude that

a. there is a large difference between the effects of the treatment and the placebo.

b. there is strong evidence that the treatment is very effective.

c. there is strong evidence that there is some difference in effect between the treatment and the placebo.

d. there is little evidence that the treatment has any effect.

e. there is evidence of a strong treatment effect.
4. The Excellent Drug Company claims its aspirin tablets will relieve headaches faster than any other aspirin on the market. To determine whether Excellent’s claim is valid, random samples of size 15 are chosen from aspirins made by Excellent and the Simple Drug Company. An aspirin is given to each of the 30 randomly selected persons suffering from headaches and the number of minutes required for each to recover from the headache is recorded.

A 5% significance level test is performed to determine whether Excellent’s (E) aspirin cures headaches significantly faster than Simple’s (S) aspirin.

The appropriate set of hypotheses to be tested is:

(a) \( H_0: \mu_E - \mu_S = 0 \)  \( H_a: \mu_E - \mu_S > 0 \)
(b) \( H_0: \mu_E - \mu_S = 0 \)  \( H_a: \mu_E - \mu_S \neq 0 \)
(c) \( H_0: \mu_E - \mu_S = 0 \)  \( H_a: \mu_E - \mu_S < 0 \)
(d) \( H_0: \mu_E - \mu_S < 0 \)  \( H_a: \mu_E - \mu_S = 0 \)
(e) \( H_0: \mu_E - \mu_S > 0 \)  \( H_a: \mu_E - \mu_S = 0 \)

5. The superintendent of the local public school district is concerned that boys in grades 1 through 5 are less mathematically competent than girls in grades 1 through 5. If this is the case, changes might need to be made in the mathematics curriculum delivered in these grades. To determine if such changes need to be made, the superintendent takes a random sample of boys and a random sample of girls in grades 1 through 5 from his district. She then compares the mean score for boys and the mean score for girls on a standardized math competency exam. Here are the results:

<table>
<thead>
<tr>
<th>Gender</th>
<th>N</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>60</td>
<td>84.05</td>
<td>12.96</td>
</tr>
<tr>
<td>Female</td>
<td>34</td>
<td>90.21</td>
<td>14.42</td>
</tr>
</tbody>
</table>

Does the data above give good evidence that, if all boys and all girls in the district took the exam, the mean score for boys would be lower than the mean score for girls? Perform a hypothesis test to find out:

a. Give the null and alternative hypotheses using symbols:

\( H_0: \mu_{boys} - \mu_{girls} = 0 \)  \( H_a: \mu_{boys} - \mu_{girls} < 0 \)

b. Define each of your parameters in words:

\( \mu_{boys} \): the mean score of ALL boys in the district
\( \mu_{girls} \): the mean score of ALL girls in the district

c. Check the technical conditions.

Observations are from two independent simple random samples of size 60 and 34. The sample sizes are fairly large, both are greater than 30.
d. Compute the test statistic.

\[ t = \frac{\bar{X}_{boys} - \bar{X}_{girls}}{\sqrt{\frac{s^2_{boys}}{n_{boys}} + \frac{s^2_{girls}}{n_{girls}}}} = \frac{(84.05 - 90.21) - 0}{\sqrt{\frac{12.96^2}{60} + \frac{14.42^2}{34}}} = -2.06 \]

Or you can use your calculator: TESTS \( \rightarrow \) 2-SampTTest

e. Determine the p-value

p-value: 0.022

f. Based on the p-value, how strong is the evidence that population of boys would achieve a lower mean than the population of girls?

Little or no evidence    Some evidence \( \bigcirc \) Good evidence    Very strong evidence

g. What conclusion would you make at significance level 5%?

Since the p-value is less than 5%, we have enough evidence to reject the null hypothesis. We can conclude that the mean scores for the boys and the girls in the district is not the same.

h. What conclusion would you make at significance level 1%?

Since the p-value is greater than 5%, at the 1% significance level we don’t have enough evidence to reject the null hypothesis. We can’t conclude that the mean scores for the boys and the girls in the district is not the same.

j. The 95% confidence interval for the difference between the mean score that all boys would achieve and the mean score that all girls would achieve is (-12.13, -0.1929), and the 99% confidence interval is (-14.09, 1.7717)

Explain the connection between these results and what you found in g & h.

Since the claimed value, zero, is not in the 95% confidence interval, we can reject the null hypothesis at the 5% significance level. But zero is inside the 99% confidence interval, so at the 1% significance level we cannot reject the null hypothesis. Same conclusions as above.

6. Do women have a better vocabulary than men? Ten married couples are chosen at random and a vocabulary test is given to each of the husbands and each of the wives. Here are the scores:

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
<td>76</td>
<td>73</td>
</tr>
<tr>
<td>74</td>
<td>84</td>
<td>83</td>
</tr>
<tr>
<td>91</td>
<td>84</td>
<td>88</td>
</tr>
<tr>
<td>84</td>
<td>77</td>
<td>79</td>
</tr>
<tr>
<td>70</td>
<td>82</td>
<td>81</td>
</tr>
<tr>
<td>77</td>
<td>92</td>
<td>88</td>
</tr>
<tr>
<td>92</td>
<td>69</td>
<td>80</td>
</tr>
</tbody>
</table>

Here is the summary statistics for the difference, define by Difference = Women – Men
mean of the differences = 4.1  standard deviation of the differences = 4.17

Perform a test of significance to help answer the question above. Give the null and alternative hypotheses, and carry out the test at the 10% significance level. Write your conclusion in context.

H₀: μₐ = 0   Hₐ: μₐ > 0

Using the t-test with the given mean and s.d.:
Test statistic =3.109  
p-value = 0.00627

Conclusion: since the p-value is less than 10%, we can reject the null hypothesis. We have enough evidence to reject the claim that there is no difference between the vocabulary scores of women and men.

7. A study of asthmatics measured the peak expiratory flow rate (basically, a person’s maximum ability to exhale) before and after a walk on a cold winter's day for a random sample of nine asthmatics. Here are the results:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Before</th>
<th>After</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>312</td>
<td>300</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>242</td>
<td>201</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>340</td>
<td>232</td>
<td>108</td>
</tr>
<tr>
<td>4</td>
<td>388</td>
<td>312</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>296</td>
<td>220</td>
<td>76</td>
</tr>
<tr>
<td>6</td>
<td>254</td>
<td>256</td>
<td>-2</td>
</tr>
<tr>
<td>7</td>
<td>391</td>
<td>328</td>
<td>63</td>
</tr>
<tr>
<td>8</td>
<td>402</td>
<td>330</td>
<td>72</td>
</tr>
<tr>
<td>9</td>
<td>290</td>
<td>231</td>
<td>59</td>
</tr>
<tr>
<td>mean</td>
<td>323.89</td>
<td>267.78</td>
<td>56.11</td>
</tr>
<tr>
<td>s.d.</td>
<td>59.83</td>
<td>50.01</td>
<td>34.17</td>
</tr>
</tbody>
</table>

Use the data to test whether there is a difference between the peak expiratory flow rate before vs. after a walk on a cold winter's day for asthmatics. Give null and alternative hypotheses, compute the test statistic (hint: there is more information above than you need), determine the p-value, and make a conclusion based on a significance level of 5%.

H₀: μₐ = 0   Ha: μₐ ≠ 0

Using the t-test with the given mean (56.11) and s.d.(34.17):

Test statistic =4.93
p-value = 0.0012

Conclusion: since the p-value is less than 5%, we can reject the null hypothesis. We have enough evidence to reject the claim that there is no difference between the peak expiratory flow rate before vs. after walk on a cold winter’s day for asthmatics.
8. In order to test the effect of Prozac on the well-being of depressed individuals, a questionnaire was administered to nine patients to gauge their level of well-being. The questionnaire was given to each subject both before and after treatment with Prozac. The data is shown below. Higher scores indicate greater well-being (that is, Prozac is having a positive effect).

<table>
<thead>
<tr>
<th></th>
<th>moodpre</th>
<th>moodpost</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.00</td>
<td>5.00</td>
<td>2.00</td>
</tr>
<tr>
<td>2</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>6.00</td>
<td>5.00</td>
<td>-1.00</td>
</tr>
<tr>
<td>4</td>
<td>7.00</td>
<td>7.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>4.00</td>
<td>10.00</td>
<td>6.00</td>
</tr>
<tr>
<td>6</td>
<td>3.00</td>
<td>9.00</td>
<td>6.00</td>
</tr>
<tr>
<td>7</td>
<td>2.00</td>
<td>7.00</td>
<td>5.00</td>
</tr>
<tr>
<td>8</td>
<td>1.00</td>
<td>11.00</td>
<td>10.00</td>
</tr>
<tr>
<td>9</td>
<td>4.00</td>
<td>8.00</td>
<td>4.00</td>
</tr>
</tbody>
</table>

Perform a test of significance to determine whether this study gives good evidence that Prozac has a positive effect. Follow the steps below:

a. Define the parameter we wish to test. (Hint: It will involve the population of differences, of which the data above is a sample.)

The mean of differences between the “moodpre” and “moodpost” scores for all depressed individuals.

b. Give the null and alternative hypotheses using symbols:

\[ H_0 : \mu_d = 0 \quad H_a : \mu_d > 0 \]

c. What two things do we need to assume in order for the test to be valid?

We need to assume that the sample is random, and that the differences between the scores are normally distributed.

d. The sample mean of the differences is 3.67, and the sample standard deviation of the differences is 3.5. Compute the test statistic.

Using the t-test with the given mean and s.d.:
Test statistic = 3.15

e. One of the values below is the correct p-value for the test. Which one must it be? Explain.

1.62 -0.016 0.007
f. Based on the p-value you selected in the previous part, how strong is the evidence that Prozac has a positive effect?

   We have strong evidence that Prozac has a positive effect.

g. Make a decision based on a 1% significance level.

   Conclusion: since the p-value is less than 1%, we can reject the null hypothesis. We have enough evidence to reject the claim that there is no difference between the mood scores before and after taking Zantac.