1. Problem 1 from the text

For the following studies, indicate whether a repeated-measures \( t \) test is the appropriate analysis. Explain your answers.

a. A researcher is comparing the amount of time spent playing video games each week for college males versus college females.

Each participant has one score associated with them – the amount of time spent playing video games each week. Because each participant has only one score, this cannot be a repeated-measures design.

b. A researcher is comparing two new designs for cell phones by having a group of high school students send a scripted text message on each model and measuring the difference in speed for each student.

Each participant has two scores which are then compared – one for each of the two phones. This makes the study a repeated-measures design.

c. A researcher is evaluating the effects of fatigue by testing people in the morning when they are well rested and testing again at midnight when they have been awake for at least 14 hours.

Each person provides two scores which are then compared – one in the morning and one at midnight. This is a repeated measures design.

2. Problem 6 from the text

A repeated-measures study with a sample of \( n = 25 \) participants produces a mean difference of \( M_D = 3 \) with a standard deviation of \( s = 4 \).

a. Based on the mean and standard deviation, you should be able to visualize (or sketch) the sample distribution. Use a two-tailed hypothesis test with \( \alpha = .05 \) to determine whether it is likely that this sample came from a population with \( \mu_D = 0 \).
The sample distribution with a mean of 3. The inflection points of the distribution are at the mean plus or minus the standard deviation. (Normally, I would not use this scale on the X axis, but I want to be able to compare this distribution to the one in part b of this question, and the scale needs to be same in order to compare.)

Step 1: State the hypotheses and select an α level.

$H_0: \mu_D = 0$

$H_1: \mu_D \neq 0$

$\alpha = .05$

Step 2: Locate the critical region

$df = n - 1 = 25 - 1 = 24$

From a table of critical $t$ values with $df = 24$, $\alpha = .05$, two tailed, the critical $t = \pm 2.064$.

Step 3: Compute the test statistics

$$s_{MD} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{4^2}{25}} = \sqrt{\frac{16}{25}} = 0.8$$

$$t = \frac{M_D - \mu_D}{s_{MD}} = \frac{3 - 0}{0.8} = 3.75$$

Step 4: Make a decision about $H_0$ and state the conclusion

Because the calculated $t$ is in the tail cut off by the critical $t$, we can reject $H_0$ and conclude that it is unlikely that the sample came from a population with $\mu_D = 0$.

b. Now assume that the sample standard deviation is $s = 12$, and once again visualize the sample distribution. Use a two-tailed hypothesis test with $\alpha = .05$ to determine whether it is likely that this sample came from a population with $\mu_D = 0$. 
The sample distribution with a mean of 3. The inflection points of the distribution are at the mean plus or minus the standard deviation.

Step 1: State the hypotheses and select an α level.

H₀: μ₀ = 0

H₁: μ₀ ≠ 0

α = .05

Step 2: Locate the critical region

df = n - 1 = 25 - 1 = 24

From a table of critical t values with df = 24, α = .05, two tailed, the critical t = ±2.064.

Step 3: Compute the test statistics

\[ s_{M_D} = \frac{s^2}{n} = \frac{12^2}{25} = \frac{144}{25} = 2.4 \]

\[ t = \frac{M_D - \mu_D}{s_{M_D}} = \frac{3 - 0}{2.4} = 1.25 \]

Step 4: Make a decision about H₀ and state the conclusion

Because the calculated t is not in the tail cut off by the critical t, we fail to reject H₀ and conclude that there is insufficient information to state that the sample came from a population with μ₀ = 0.

c. Explain how the size of the sample standard deviation influences the likelihood of finding a significant mean difference.

Increasing the sample standard deviation decreases the statistical power of the repeated measures t test. Increasing the sample standard deviation increases the standard error of the difference of the means which decreases the size of the calculated t.
3. Problem 11 from the text

Strack, Martin, and Stepper (1988) reported that people rate cartoons as funnier when holding a pen in their teeth (which forced them to smile) than when holding a pen in their lips (which forced them to frown). A researcher attempted to replicate this result using a sample of \( n = 25 \) adults between the ages of 40 and 45. For each person, the researcher recorded the difference between the rating obtained while smiling and the rating obtained while frowning. On average the cartoons were rated as funnier when the participants were smiling, with an average difference of \( M_D = 1.6 \) with \( SS = 150 \).

a. Do the data indicate that the cartoons are rated significantly funnier when the participants are smiling? Use a one-tailed test with \( \alpha = .01 \).

Step 1:

\( H_0: \mu_D \leq 0 \)

\( H_1: \mu_D > 0 \)

\( \alpha = .01 \)

Step 2:

\( t_{\text{critical}} = 2.492 \) (\( \alpha = .01 \), one-tailed, \( df = n - 1 = 25 - 1 = 24 \))

Step 3:

\[
\begin{align*}
s^2 &= \frac{SS}{n - 1} = \frac{150}{25 - 1} = 6.25 \\
s_{MD} &= \sqrt{\frac{s^2}{n}} = \sqrt{\frac{6.25}{25}} = 0.5 \\
t &= \frac{M_D - \mu_D}{s_{MD}} = \frac{1.6 - 0}{0.5} = 3.2
\end{align*}
\]

Step 4:

Because the calculated \( t \) is in the tail cut off by the critical \( t \), we can reject \( H_0 \) and conclude that it is likely that the cartoons are rated as more funny when smiling.

Note: In a series of 17 studies, Wagenmakers, Beek, Dijkhoff and Gronau (2016) failed to replicate the results of Strack, Martin and Stepper (1988). At this point in time, I wouldn’t put great faith in the results of Strack et al.

b. Compute \( r^2 \) to measure the size of the treatment effect.

\[
\begin{align*}
r^2 &= \frac{t^2}{t^2 + df} = \frac{3.2^2}{3.2^2 + 24} = 0.299
\end{align*}
\]
This is a large effect (≥ .25).

c. Write a sentence describing the outcome of the hypothesis test and the measure of effect size as it would appear in a research report.

Holding a pencil in the mouth increased the rating of how funny a cartoon was compared to not holding a pencil in the mouth by an average of \( M = 1.6 \) points with \( s = 2.5 \). The effect was statistically significant, \( t(24) = 3.20, p < .01, r^2 = .299 \).

4. Problem 15 from the text

The following data are from a repeated-measures study examining the effect of a treatment by measuring a group of \( n = 4 \) participants before and after they receive the treatment.

a. Calculate the difference scores and \( M_D \).

<table>
<thead>
<tr>
<th>Participant</th>
<th>Before Treatment</th>
<th>After Treatment</th>
<th>( D )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7</td>
<td>10</td>
<td>10 - 7 = 3</td>
</tr>
<tr>
<td>B</td>
<td>6</td>
<td>13</td>
<td>13 - 6 = 7</td>
</tr>
<tr>
<td>C</td>
<td>9</td>
<td>12</td>
<td>12 - 9 = 3</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>8</td>
<td>8 - 5 = 3</td>
</tr>
</tbody>
</table>

\[ M_D = \frac{(3 + 7 + 3 + 3)}{4} = 4 \]

b. Compute \( SS \), sample variance, and estimated standard error.

\[ SS_D = \Sigma (X_D - M_D)^2 = (3 - 4)^2 + (7 - 4)^2 + (3 - 4)^2 + (3 - 4)^2 = 1^2 + 3^2 + 1^2 + 1^2 = 12 \]

\[ s^2_D = SS_D / df = 12 / (4 - 1) = 4 \]

\[ s_{MD} = \sqrt{\frac{s^2}{n}} = \sqrt{\frac{4}{4}} = 1 \]

c. Is there a significant treatment effect? Use \( \alpha = .05 \), two-tails.

Step 1: State the hypotheses and \( \alpha \)

\( H_0: \mu_D = 0 \)

\( H_1: \mu_D \neq 0 \)

\( \alpha = .05 \)

Step 2: Determine the critical region

\( df = n - 1 = 4 - 1 = 3 \)

The critical \( t \) for \( df = 3 \), \( \alpha = .05 \), two-tailed is \( t = \pm 3.182 \)

Step 3: Calculate the test statistic

\[ t = \frac{M_D - \mu_D}{s_{MD}} = \frac{4 - 0}{1} = 4 \]
Step 4: Make the decision

Because the calculated value of $t$ is in the tail cut off by the critical $t$, we can reject $H_0$ and conclude that the treatment likely was effective.

5. Enter the data from problem 21 into SPSS. Use SPSS to answer the following question: Do the data indicate a significant difference between the two conditions?

Yes – the $p$ value is .004 which is less than $\alpha$ (.05), so we can reject $H_0$ that the population means for changing answers is less than or equal to the population mean for not changing answers.

Give $H_0$, $H_1$ and $\alpha$.

$H_0$: $\mu_D \geq 0$ where $D =$ No change score – Change score
$H_1$: $\mu_D < 0$
$\alpha = .05$

Sketch a $t$ distribution and indicate the critical region(s) on the distribution and the critical $t$ value(s).

Is this a one-tailed or two-tailed test?

This is a one-tailed (directional) test.

From the SPSS output, report the following:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{\text{No Change}}$</td>
<td>76.11</td>
</tr>
<tr>
<td>$M_{\text{Change}}$</td>
<td>83.11</td>
</tr>
<tr>
<td>$df$</td>
<td>8</td>
</tr>
</tbody>
</table>
What is the estimated value of Cohen's $d$? Is this a small, medium, or large effect?

Estimated $d = M_D / s = 7 / 6 = 1.167$, large effect

There is some evidence suggesting that you are likely to improve your test score if you rethink and change answers on a multiple choice exam (Johnston, 1975). To examine this phenomenon, a teacher gave the same final exam to two sections of a psychology course. The students in one section were told to turn in their exams immediately after finishing, without changing any of their answers. In the other section, students were encouraged to reconsider each question and to change answers whenever they felt it was appropriate. Before the final exam, the teacher had matched 9 students in the first section with 9 students in the second section based on their midterm grades. For example, a student in the no-change section with an 90 on the midterm exam was matched with a student in the change section who also had an 89 on the midterm. The final exam grades for the 9 matched pairs of students are presented in the following table:

<table>
<thead>
<tr>
<th>Matched Pair</th>
<th>No-Change Section</th>
<th>Change Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>71</td>
<td>86</td>
</tr>
<tr>
<td>#2</td>
<td>68</td>
<td>80</td>
</tr>
<tr>
<td>#3</td>
<td>91</td>
<td>88</td>
</tr>
<tr>
<td>#4</td>
<td>65</td>
<td>74</td>
</tr>
<tr>
<td>#5</td>
<td>73</td>
<td>82</td>
</tr>
<tr>
<td>#6</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>#7</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>#8</td>
<td>86</td>
<td>88</td>
</tr>
<tr>
<td>#9</td>
<td>65</td>
<td>76</td>
</tr>
</tbody>
</table>

**Paired Samples Statistics**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NoChange</td>
<td>76.1111</td>
<td>9</td>
<td>9.81637</td>
<td>3.27212</td>
</tr>
<tr>
<td>Change</td>
<td>83.1111</td>
<td>9</td>
<td>5.46453</td>
<td>1.82151</td>
</tr>
</tbody>
</table>

**Paired Samples Correlations**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Correlation</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>9</td>
<td>.841</td>
<td>.005</td>
</tr>
</tbody>
</table>
### Paired Samples Test

<table>
<thead>
<tr>
<th>Paired Differences</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
</table>

6. Write a sentence or two in APA format that summarizes the results of the above question.

Not changing answers ($M = 76.11, s = 9.82$) and changing answers ($M = 83.11, s = 5.46$) have different accuracies, $t(8) = 3.50, p = .008, d = 1.167$ (or $r^2 = .605$).