STAT111 Homework 6 Solution

Total: 33

1. (6 points, 1 each)
   a. For example, Subject 1’s weight change is 61.7 - 55.7 = 6 kg.
   b. The mean change is $x = 4.73125$ kg and the standard deviation is $s = 1.7457$ kg.
   c. $SE = \frac{s}{\sqrt{16}} = 0.4364$ kg; for df = 15, $t^* = 2.131$, so the margin of error for 95% confidence is $\pm 0.9300$ (software: $\pm 0.9302$). Based on a method that gives correct results 95% of the time, the mean weight change is 3.801 to 5.661 kg.
   d. $\bar{x} = 10.40875$ lb, $s = 3.84054$ lb, and the 95% confidence interval is $8.363$ to $12.455$ lb.
   e. We test $H_0$: $\mu = 16$ lb. versus $H_a$: $\mu \neq 16$ lb. The test statistic is $t = \frac{4.73125 - 16}{1.7457/\sqrt{16}} = -25.851$ with df = 15, which is highly significant evidence ($P < 0.0001$) against $H_0$.
   f. The data suggest that the excess calories were not converted into weight; the subjects must have used this energy some other way.

2. (4 points, 1 each)
   a. For the differences, $x = $114 and $s = $114.402.
   b. We wish to test $H_0$: $\mu = 0$ versus $H_a$: $\mu > 0$, where $\mu$ is the mean difference between Jocko’s estimates and those of the other garage. (The alternative hypothesis is one-sided because the insurance adjusters suspect that Jocko’s estimates may be too high.) For this test, we find $t = \frac{114 - 0}{114.402/\sqrt{10}} = 3.151$ with df = 9, for which $P = 0.0059$. This is significant evidence against $H_0$. That is, we have good reason to believe that Jocko’s estimates are higher.
   c. The 95% confidence interval with df = 9 is $\bar{x} \pm 2.262s/\sqrt{10} = $114 ± $81.83 = $32.17 to $195.83$.
   d. Student answers may vary; based on the confidence interval, one could justify any answer in the range $32.17$ to $195.83$.

3. (1 point) The 90% confidence interval is $3.92 \pm t^*(1.02/\sqrt{2368})$. With Table D, take df = 1000 and $t^* = 1.646$; with software, take df = 2367 and $t^* = 1.6455$. Either way, the confidence interval is 3.886 to 3.955.
4. (2 points) With the given standard deviations, $SE_D = 0.2653$; regardless of how we choose df, a 95% confidence interval for the difference in means rounds to 4.37 to 5.43 mg/y/m^3. With the null hypothesis $H_0: \mu_1 = \mu_2$ (and either a one- or two-sided alternative, as in the previous exercise), we find $t = 18.47$, for which $P < 0.0001$ regardless of df and the chosen alternative. We have strong evidence that outdoor concrete workers have lower respirable dust exposure than the tunnel workers.

<table>
<thead>
<tr>
<th>df</th>
<th>$t^*$</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>121.5</td>
<td>1.9797</td>
<td>4.3747 to 5.4253</td>
</tr>
<tr>
<td>114</td>
<td>1.9810</td>
<td>4.3744 to 5.4256</td>
</tr>
<tr>
<td>100</td>
<td>1.984</td>
<td>4.3736 to 5.4264</td>
</tr>
</tbody>
</table>

5. (5 points, 1 each)

a. The 68-95-99.7 rule suggests that the distributions are not Normal: If they were Normal, then (for example) 95% of 7- to 10-year-olds drink between -13.2 and 29.6 oz of sweetened drinks per day. As negative numbers do not make sense, the distributions must be right-skewed.

b. We find $SE_D = 4.3786$ and $t = -1.439$, with either df = 7.8 ($P = 0.1890$) or df = 4 ($P = 0.2236$). We do not have enough evidence to reject $H_0$. There is insufficient evidence to say that one age group on average drinks more sweetened drinks than the other.

c. The possible 95% confidence intervals are given in the table.

d. Because the distributions are not Normal and the samples are small, the procedures are very questionable for these data.

e. Because this group is not an SRS (and indeed might not be random in any way) we would have to be very cautious about extending these results to other children.

<table>
<thead>
<tr>
<th>df</th>
<th>$t^*$</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.8</td>
<td>2.3159</td>
<td>-16.4404 to 3.8404</td>
</tr>
<tr>
<td>4</td>
<td>2.776</td>
<td>-18.4551 to 5.8551</td>
</tr>
</tbody>
</table>

6. (3 points, 1 each)

a. The null hypothesis is $H_0: \mu_1 = \mu_2$; the alternative can be either two- or one-sided. (It might be a reasonable expectation that $\mu_1 > \mu_2$.) We find $SE_D = 0.2796$ and $t = 8.369$. Regardless of the df and $H_a$, the conclusion is the same: $P$ is very small, and we conclude that WSJ ads are more trustworthy.

b. Possible 95% confidence intervals are given in the table; both place the difference in trustworthiness at between about 1.8 and 2.9 points.

c. Advertising in WSJ is seen as more reliable than advertising in the National Enquirer: a conclusion that probably comes as a surprise to no one.

<table>
<thead>
<tr>
<th>df</th>
<th>$t^*$</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>121.5</td>
<td>1.9797</td>
<td>1.79 to 2.89</td>
</tr>
<tr>
<td>60</td>
<td>2.000</td>
<td>1.78 to 2.90</td>
</tr>
</tbody>
</table>
This is a question about two-sample t test because the Domestic and International samples are independent. Let ‘Domestic’ be group 1 with mean $\mu_1$ and SD $\sigma_1$ and ‘International’ be group 2 with mean $\mu_2$ and SD $\sigma_2$. We want to test if $\mu_1 = \mu_2$. Therefore, the null hypothesis is $H_0 : \mu_1 - \mu_2 = 0$ and the alternative is $H_a : \mu_1 - \mu_2 \neq 0$. From the table, we have summary statistics for group 1 and group 2.

\[
\bar{x}_1 = 7.15, s_1 = 2.54, n_1 = 8 \quad \text{and} \quad \bar{x}_2 = 8.38, s_2 = 3.51, n_2 = 8
\]

The test statistic $T_0$ is given by

\[
T_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{7.15 - 8.38}{\sqrt{\frac{2.54^2}{8} + \frac{3.51^2}{8}}} = -0.80.
\]

The P-value is $2 \cdot P(T \geq |T_0|) = 2 \cdot 0.225 = 0.450$ (from software). Also, we can use t table ($df = 7$) and find the range of the P-value. This is $0.4 < P\text{-value} < 0.5$. Either way, we don’t reject the null hypothesis. There is not enough evidence to reject the null.

Now, the sample is a matched pairs design. We need to do one-sample t test using a new variable $D$ that is ‘Domestic−International’. The new variable $D$ is $-0.9, -1.5, -3.3, -2.7, 0.2, -2.5, 0.4, 0.5$

And, the summary statistics are

\[
\bar{x}_d = -1.23, s_d = 1.51, n_d = 8.
\]

The hypotheses are

\[
H_0 : \mu_d = 0, \quad H_a : \mu_d \neq 0.
\]

The test statistic $T_0$ is given by

\[
T_0 = \frac{\bar{x}_d - \mu_d}{s_d/\sqrt{n_d}} = \frac{-1.23}{1.51/\sqrt{8}} = -2.30.
\]

The P-value is $2 \cdot P(T \geq |T_0|) = 0.055$ (from software). Therefore, we don’t reject the null with a level 0.05.

We use a matched pairs test to see if there is a difference between average pretest and posttest scores. The hypotheses are

\[
H_0 : \mu_{post} - \mu_{pre} = 0, \quad H_a : \mu_{post} - \mu_{pre} \neq 0
\]

The test statistic $T_0$ is given by

\[
T_0 = \frac{\bar{x}_d - \mu_d}{s_d/\sqrt{n_d}} = \frac{1.45}{3.2032/\sqrt{20}} = 2.024.
\]
$T_0$ has a t distribution with $df = 19$. The P-value is $2 \cdot P(T \geq |T_0|) = 0.057$. Therefore, we don’t reject the null with a level 0.05.

The confidence interval is

$$(1.45 - 2.093 \times 3.2032/\sqrt{20}, 1.45 + 2.093 \times 3.2032/\sqrt{20}) = (-0.05, 2.95)$$

It contains zero, so we have the same conclusion of not rejecting the null with a level 0.05.

10. (2 points, 1 each)

a. This is a two-sample t test problem. We want to test if $\mu_f < \mu_h$.
Therefore, the null hypothesis is $H_0 : \mu_f - \mu_h \geq 0$ and the alternative is $H_a : \mu_f - \mu_h < 0$. We calculate the following summary statistics using the data.

$\bar{x}_f = 0.82364, s_f = 0.481072, n_f = 33$ and $\bar{x}_h = 1.72559, s_h = 0.639273, n_h = 68$

The test statistic $T_0$ is given by

$$T_0 = \frac{(\bar{x}_f - \bar{x}_h) - (\mu_f - \mu_h)}{\sqrt{\frac{s_f^2}{n_f} + \frac{s_h^2}{n_h}}} = -7.90.$$  

The P-value is $P(T < T_0) = 0$. So we reject the null and conclude that failed firms have a lower ratio with level 0.05.

b. We cannot do a randomized experiment for this kind of question because the data is observational. We cannot assign a healthy firm to the failed group.