Assignment 1: Answers
January 9, 2017

1. The probability decreases. As sample size increases, sample means will become less variable, so it will become less likely that a sample mean very different from \( \mu \) will be obtained. With \( \sigma = 10 \), the standard error of the mean (\( \sigma_M \)), which is the standard deviation of the distribution of sample means would be \( \frac{10}{\sqrt{20}} = 2.236 \). A sample mean of 54 corresponds to a \( z \) score of \( \frac{54 - 50}{2.236} = 1.79 \). The area corresponding to values greater than or equal to a \( z \) value of 1.79 is the smaller portion. From the \( z \) table, that portion is .0367, so the probability of drawing a sample with \( M \geq 54 \) is .0367.

2. The sampling distribution of the mean will be normal because the population of raw scores is normal. The mean of this distribution, \( \mu_M \) will be equal to \( \mu \) and its standard deviation will be \( \frac{s}{\sqrt{N}} = \frac{4}{\sqrt{10}} = 1.26 \). In a normal distribution with \( \mu_M = 20 \) and \( \sigma_M = 1.26 \), a sample mean of 22 would correspond to a \( z \) score of \( \frac{22 - 20}{1.26} = 1.59 \). Using the standard normal distribution (Table E.10 in Howell’s book), the smaller portion of the normal distribution corresponding to a \( z \) value of 1.59 is .0559. We use the smaller portion because we are interested in sample means of 22 or larger and means in that range occupy the small portion of the distribution when it is cut along a \( z \) value of 1.59.

With a sample size of 30, the only thing that changes is the standard error of the mean: \( \sigma_M = \frac{4}{\sqrt{30}} = 0.73 \). For \( M = 22 \), \( z = \frac{22 - 20}{0.73} = 2.74 \); area in smaller portion = .0031. Notice how increasing the sample size makes it less likely to get a sample mean that is larger than \( \mu \) by 2 or more points.

3. A rejection criterion is the least extreme observed outcome that will allow rejection of the null hypothesis. A region of rejection represents the part(s) of a distribution of outcomes that will allow rejection of the null hypothesis (i.e., outcomes that are equal to, or more extreme than, the rejection criterion).

(a) For this particular problem, the rejection criterion is the value of the sample mean that corresponds to a \( z \) score that cuts off the lower .05 of the standard normal distribution. From the standard normal distribution (\( z \) table, the closest to .05 we can get (without exceeding .05) when examining the smaller portion column is \( z = 1.65 \). Because we are working with the lower end of the distribution (the alternative hypothesis claims the true value of \( \mu \) is less than 50), the relevant \( z \) score is \( -1.65 \). That is the cutoff for rejecting \( H_0 \).
The value of the sample mean corresponding to that cutoff is \( M = \mu + \sigma_M(z) \). \( \sigma_M = \sigma / \sqrt{N} = 10 / \sqrt{30} = 1.83 \), so we have \( M = 50 + (1.83)(-1.65) = 46.98 \). This is the rejection criterion; the least extreme value of \( M \) that will allow rejection of \( H_0 \). The region of rejection is any sample mean with a value equal to or less than (more extreme than) 46.98.

(b) With a nondirectional test, there are two rejection criteria and two corresponding regions of rejection, one at each end of the distribution. The relevant \( z \) values cutting off the lower and upper .025 of the standard normal distribution are \( z = \pm 1.96 \). Combined, these two regions account for the most extreme .05 of the normal distribution. The rejection criteria are as follows. At the lower end, \( M = 50 + (1.83)(-1.96) = 46.41 \). At the upper end, \( M = 50 + (1.83)(1.96) = 53.59 \). The regions of rejection are values of \( M \) that are equal to or less than 46.41 and values that are equal to or greater than 53.59.

4. Copying the data into \textbf{R} as suggested would produce:

\begin{verbatim}
> data = c(48, 45, 51, 41, 46, 48, 41, 58, 52, 51, 51, 53, 54, 55, 48, 57, 45, 47, 43, 48, 56, 52, 45, 48, 51, 32, 41, 50, 48, 51)
\end{verbatim}

The \texttt{t.test} function with the correct parameters would be set up as follows. The first parameter is the variable containing the scores, the second parameter specifies a directional alternative hypothesis with the true value of \( \mu \) expected to be less than the null value, and the third parameter specifies the null hypothesis value for \( \mu \).

\begin{verbatim}
> t.test(data, alternative = "less", mu = 50)
\end{verbatim}

\textbf{One Sample t-test}

\begin{verbatim}
data:  data
t = -1.4522, df = 29, p-value = 0.0786
alternative hypothesis: true mean is less than 50
95 percent confidence interval:  
  -Inf 50.24943
sample estimates:
mean of x
  48.53333
\end{verbatim}

The obtained \( t \) ratio is reported as \(-1.4522\), with the negative sign implying that the sample mean is less 50, which at least is in the direction expected by the alternative hypothesis. The last line of the output shows the sample mean (48.53). But is it small enough to reject the null hypothesis? The reported \( p \)-value of .0786 indicates that the obtained \( t \) ratio is located at a point below which the lower .0786 of the \( t \) distribution falls (see the diagram below). This means that the obtained \( t \) ratio is not among the .05 most extreme outcomes, and therefore we cannot reject the null hypothesis.
Using R to compute the mean and standard deviation of this sample, we have:

```r
> mean(data)
[1] 48.53333
> sd(data)
[1] 5.531934
```

So the standard error of the mean, based on a sample size of 30, is $\frac{5.532}{\sqrt{30}} = 1.01$. This means that $t = \frac{(48.533 - 50)}{1.01} = -1.452$.

5. The first step in computing the related-samples $t$ test is to obtain a difference score for each subject by subtracting the score in one condition from that person’s score in the other condition. In this example, it makes sense to subtract the post-test score from the pre-test score to assess the extent to which stress levels have been reduced. So we have the following set of difference scores.

<table>
<thead>
<tr>
<th>Subj</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

The mean of the difference scores is $M_D = 2.25$. The standard deviation of the difference scores is $s_D = 1.488$. The standard error of the mean of difference scores is $s_{M_D} = \frac{1.488}{\sqrt{8}} = .526$ so the obtained $t$ ratio is $t = \frac{(2.25 - 0)}{.526} = 4.28$. The critical $t$ ratio for a nondirectional test and $\alpha = .05$ with 7 degrees of freedom is $t_{\text{crit}}(7) = \pm 2.365$. The obtained $t$ ratio is more extreme than the critical $t$ ratio, so the null hypothesis can be rejected and we can concluded that the post-test scores are significantly lower than the pre-test scores (pre-test mean = 6.75, post-test mean = 4.50). But can we conclude that the naturopathic agent caused this difference? No, because we do not have an adequate control condition to rule out the possibility that stress might have been reduced just by the passing of time or that subjects might have experienced a placebo effect when receiving the naturopathic agent.
6. Begin by reading the data file into R. Download the file from the course web site and be sure it is named `infant.txt` in your directory. The following command can be used to open a dialogue box in which you can find and select this file. The file will be placed into the data variable called `look`.

```r
> look = read.table(file.choose(new = T), header = T)
```

Load the psych library so you can use the `describe` command to obtain the condition means.

```r
> library(psych)
> describe(look)
```

```
  vars n  mean   sd median trimmed mad min max range  skew kurtosis   se
old  1 24 2.94 1.67   2.90    2.90 2.30 0.6 5.9   5.3  0.20 -1.34 0.34
new  2 24 3.72 1.74   3.55    3.73 1.48 0.4 6.8   6.4 -0.01 -0.94 0.36
```

Notice that the mean looking time is greater for the new condition than for the old condition. Now we can apply the related-samples `t` test to test this difference for significance.

```r
> t.test(look$old, look$new, paired = T)
```

```
  Paired t-test

data:  look$old and look$new
  t = -2.7775, df = 23, p-value = 0.01071
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:  
 -1.3594791 -0.1988542
sample estimates:  
 mean of the differences  
 -0.7791667
```

The obtained `t` ratio is significant because its `p` value is less than `α` (which is .05). The result could be reported as `t(23) = −2.78, p < .05`. The conclusion is infants looked longer at the novel pattern, even after a day had elapsed since the old pattern had been seen, indicating that on average, their memory for the old pattern lasted at least one day.