Solutions 8

(1) In this question we investigate whether there is a difference between the mean number of blue and yellow M&Ms in a fun size bag. To investigate this I am not sure whether to use matched paired method or two-sample independent t methods to compare the two samples.

Matched pair methods are used and the output is give in Figure 1 (you can replicate this by inputting the data into Statcrunch).

(a) By considering how the data was collected, which procedure should be used to analyse the data, a matched paired method or a two-sample independent t.

A matched paired method should be use, because both the variables (number of yellows and number of blues) are taken from the same bag. Since they share they same bag it is highly likely there is some dependence between the number of yellows and the number of blues.

(b) By identifying the correct output give the margin of error for difference in mean number of M&Ms.

The top output should be used. The margin of error is simply half the length of the CI, which in this case is \((0.612 + 0.17)/2\). Alternatively, one could use the formula \(t_{0.04}(2.5) \times 0.1969\), using Statcrunch \(t_{0.04}(2.5) = 1.9855\) thus the margin of error = 1.9885 \times 0.1969.

(c) If we want to see whether there is a difference between the mean number of blues and yellows and test \(H_0: \mu_{\text{blue}} - \mu_{\text{yellow}} = 0\) against \(H_A: \mu_{\text{blue}} - \mu_{\text{yellow}} \neq 0\), and do the test at the 5% level. By identifying the output what are the conclusions of the test?

As this is a two-sided test (at the 5% level) and 0 is inside the 95% confidence interval \([-0.1699, 0.612]\) we CANNOT reject the null hypothesis. There is not enough evidence in the data to say that the mean number of blues and yellows are different.

(d) Suppose we want to see if there is evidence that the number of blues is greater than the number of yellows and test \(H_0: \mu_{\text{blue}} - \mu_{\text{yellow}} = 0\) against \(H_A: \mu_{\text{blue}} - \mu_{\text{yellow}} > 0\).
(i) Identify the correct output calculate the t-transform for the test. 
**Use the top output.**

(ii) By using the critical values from the tables below determine a bound for the p-value for the above test and whether we can reject the null hypothesis at the 5% level.

<table>
<thead>
<tr>
<th>probability</th>
<th>0.15</th>
<th>0.10</th>
<th>0.05</th>
<th>0.025</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t^*$</td>
<td>1.04</td>
<td>1.29</td>
<td>1.66</td>
<td>1.98</td>
<td>2.36</td>
</tr>
</tbody>
</table>

The critical values for the t distribution with 94 degrees of freedom.

<table>
<thead>
<tr>
<th>probability</th>
<th>0.15</th>
<th>0.10</th>
<th>0.05</th>
<th>0.025</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t^*$</td>
<td>1.03</td>
<td>1.28</td>
<td>1.65</td>
<td>1.97</td>
<td>2.34</td>
</tr>
</tbody>
</table>

The critical values for the t distribution with 187.2 degrees of freedom.

Since this is a matched paired t-test we need to use the critical values for the t-distribution with 94df. The t-transform = $\frac{0.2215}{0.196} = 1.13$. As 1.13 lies between 1.04 and 1.29, the area to the right of 1.13 (since the alternative hypothesis is $H_A : \mu_{\text{blue}} - \mu_{\text{yellow}} > 0$) lies between $10 - 15\%$. The p-value for the test is between 10 – 15%. As the p-value is greater than 5%, at the 5% level we are UNABLE to reject the null, there is not enough evidence to suggest that the mean number of yellow M&Ms is less than the mean number of blues.

(e) The distribution of M&Ms is clearly NOT normal, because the number of M&Ms can only take integer (whole) numbers, whereas the normal distribution describes the distribution of (some) numerical continuous observations. Explain why we can still use the normal (or t, does not really matter which) to construct confidence intervals and do the tests given above.

This is because of the central limit theorem, which says that if the sample size is large enough the distribution of the sample mean (average) will be close to normal, even if the distribution of the variables (such as the number of M&Ms in a bag) is not at all normal.

(2) Fred wants to use statistics to ‘prove’ his research. Following what we learn in his statistics class, he stated his research hypothesis as the alternative and the opposite of that as his null hypothesis. After collecting the data, the test at the 5% level and is unable to reject the null. Disappointed with this result, he goes and collects another sample does the test at the 5% level and again is unable to reject the null. He does this 5 times and on the fifth attempt rejects the null. He then writes a paper saying that there is evidence to reject the null hypothesis, because his p-value was less than 5%.
Are you convinced by his results? Comment on the validity of the result?

The significance level of 5% means that if the null is true there is a 5% chance the data will be such that the p-value is less than 5%. This means if the null is true and the experiment is repeated several times over, then eventually (in 5 times out of 100) we will be able to reject the null. Therefore, by repeating an experiment several times until eventually you are able to get a p-value less than 5% does not mean the alternative is true (as this will eventually happen if the null is true).

Detecting the alternative when the null is true is called a false positive. By repeating an experiment several times over, we increase the chance of a false positive.

(3) You are planning an evaluation of a semester long alcohol awareness campaign. You intend to make a survey where you ask the question ‘Did the campaign alter your behaviour towards alcohol consumption?’ the reply can either be yes or no. Prior to taking the survey you need to decide the number of people you will survey.

(assume a 95% CI).

(a) How large a sample size should you take if you want the margin of error for the true proportion of students (who were influence by the campaign) to be about 10%?

To answer this question we need to know that the 95% CI for the population proportion p is \[ \hat{p} \pm 1.96 \times \sqrt{\frac{p(1-p)}{n}} \], where we usually replace the p in the square root by the estimated \( \hat{p} \).

Therefore the MoE = 1.96 \times \sqrt{\frac{p(1-p)}{n}} = 0.1. However, we cannot directly solve for this since \( p \) is unknown (this is what we trying to estimate!). However we do know that \( \sqrt{p(1-p)} \) is maximised when \( p = 1/2 \). This means a conservative sample size would be the solution of 1.96 \times \sqrt{\frac{0.5^2}{n}} = 0.1. Solving this we get \( n = (0.5)^2 \times \left(\frac{1.96}{0.1}\right)^2 = 96 \). Since this is a conservative estimate, \( n = 96 \) will lead to a MoE which is 0.1 or less.

(b) Suppose that previous campaigns indicate that about 25% of the students will respond ‘yes’. Based on this information how large a sample size should you take if you want the margin of error for the true proportion of students (who were influence by the campaign) to be about 10%?

Here we have some prior information on a possible value for \( p \), that is \( p = 0.25 \). Using this we solve \( n = 0.25 \times 0.75 \times \left(\frac{1.96}{0.1}\right)^2 = 72 \).

(4) Castaneda v Partida is an important court case in which statistical methods were used as part of the legal argument.

In Castaneda the plaintiffs alleged that the method for selecting juries in a county in Texas as biased against Mexican Americans. For the period of time of issue, there were 181,535 persons eligible for jury duty, of whom 143,611 were Mexican Americans. Of the 870 people selected for jury duty, 339 were Mexican American.
(a) We can do a one-sample proportion test, by using \( p = \frac{143,611}{181,535} = 0.79 \) as the overall proportion of mexicans in the county and comparing the proportion of Mexican american on jury to this.

**Hypothesis test results:**

\( p \) : proportion of successes for population  
\( H_0 : p = 0.79 \)  
\( H_A : p < 0.79 \)

<table>
<thead>
<tr>
<th>Proportion</th>
<th>Count</th>
<th>Total</th>
<th>Sample Prop.</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p )</td>
<td>339</td>
<td>870</td>
<td>0.38965517</td>
<td>0.013809042</td>
</tr>
</tbody>
</table>

Figure 2:

(i) Explain why we are testing \( H_0 : p = 0.79 \) against \( H_A : p < 0.79 \)?

The plaintiff alleges that the method of selection is biased against Mexican Americans. This means they want to see if there is evidence to suggest that Mexican Americans are under-represented in Juries. We articulate this in the alternative as the proportion of Mexican Americans doing jury service being less than the total proportion in the Mexican Americans in the county, ie. \( H_A : p < 0.79 \).

(ii) Using the output in Figure 2 calculate the z-transform and the corresponding p-value.

\[ z = \frac{0.38 - 0.79}{0.138} = -29.7 \]. The alternative is pointing to the left, so the p-value is the area to the left of \( z = -29.7 \). If we use the tables we see that it only goes up to \( z = -3.49 \), which corresponds to the area of 0.02%. Therefore the p-value is a lot less than 0.02% and there is evidence to suggest that the proportion of Mexican Americans who did jury service is less than 0.79.

(iii) Is there any evidence to support the view that Mexican Americans are being underrepresented in jury duty?

Yes, since the p-value is a lot smaller than 5%.

(b) An alternative method is to do a two-sided test and compare proportions of Mexican Americans who did jury duty with the proportion of non-Mexican American who did jury duty.

**Hypothesis test results:**

\( \pi_1 \) : proportion of successes for population 1  
\( \pi_2 \) : proportion of successes for population 2  
\( \pi_1 - \pi_2 \) : difference in proportions  
\( H_0 : \pi_1 - \pi_2 = 0 \)  
\( H_A : \pi_1 - \pi_2 > 0 \)

<table>
<thead>
<tr>
<th>Difference</th>
<th>Count1</th>
<th>Total1</th>
<th>Count2</th>
<th>Total2</th>
<th>Sample Diff.</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi_1 - \pi_2 )</td>
<td>531</td>
<td>37924</td>
<td>339</td>
<td>143611</td>
<td>0.011641144</td>
<td>3.987177E-4</td>
</tr>
</tbody>
</table>

Figure 3:
(i) What hypothesis are you testing using this method?

$H_0 : p_A - p_M = 0$ against $H_A : p_A - p_M > 0$ (where $p_M$ = proportion of Mexican Americans and $p_A$ = proportion of non-Mexican Americans).

(ii) Using the output in Figure 3 calculate the z-transform and the corresponding p-value.

The z-transform is $z = 0.0116/(3.98 \times 10^{-4}) = 29.1$. Since the alternative is pointing to the right, the p-value is the area to the right of 29.1. Using the same arguments as those in (aii) this corresponds to the a p-value which is less than 0.02%.

(iii) Is there any evidence to support the view that Mexican Americans are being underrepresented in jury duty?

As the p-value is less than 0.02%, which is less than the 5% significance level, there is evidence to suggest that the proportion of Mexican Americans who did jury service is less than the proportion of non-Mexican Americans who did jury service.