Inferences for the difference between two means based on two independent samples

Notation: Assume two independent samples from two populations with the following notations:

Sample I: Mean = $\bar{x}_1$, Standard Deviation = $s_1$, sample size = $n_1$.

Population I: Mean = $\mu_1$, variance = $\sigma_1^2$.

Sample II: Mean = $\bar{x}_2$, Standard Deviation = $s_2$, sample size = $n_2$.

Population II: Mean = $\mu_2$, variance = $\sigma_2^2$.

Sampling distribution of $\bar{x}_1 - \bar{x}_2$:

(i) The mean of the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is $\mu(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$.

(ii) Standard deviation of the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is

$$\sigma(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$  

This quantity is referred to as standard error of $(\bar{x}_1 - \bar{x}_2)$.

(iii) The sampling distribution of $(\bar{x}_1 - \bar{x}_2)$ is approximately normal for large sample sizes.

1. Large sample test and confidence interval estimation of $\mu_1 - \mu_2$ based on two independent samples.

(i) Test of $H_0 : \mu_1 - \mu_2 = D_0$ (known value):
\[ Z = \frac{\bar{x}_1 - \bar{x}_2 - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \]

(ii) 100(1 - \alpha)\% confidence interval for \( \mu_1 - \mu_2 \):

\[ \bar{x}_1 - \bar{x}_2 \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \]

Example: An experiment has been conducted at a university to compare the mean number of study hours expended per week by student athletes with the mean number of hours expended by nonathletes. A random sample of 55 athletes produced a mean equal to 20.6 hours studied per week and a standard deviation equal to 5.2 hours. A second random sample of 200 nonathletes produced a mean equal to 23.5 hours per week and a standard deviation equal to 3.8 hours. Do these samples provide sufficient evidence to conclude that the population means of nonathletes and athletes differ significantly? Find the observed significance probability.

\( H_0 : \mu_1 - \mu_2 = 0; H_a : \mu_1 - \mu_2 \neq 0; \) Rejection region: \( Z > 1.96, \) or \( Z < -1.96; \)

\( Z = -3.86; \) p-value = 0; Sufficient evidence to conclude that the population means of nonathletes and athletes differ significantly)

2. Small sample test and confidence interval estimation of \( \mu_1 - \mu_2 \) based on two independent samples. This test requires that the two populations are normal with the same variance \( \sigma^2 \).
(i) Test of $H_0 : \mu_1 - \mu_2 = D_0$:

$$t = \frac{\bar{x}_1 - \bar{x}_2 - D_0}{\sqrt{s_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

with $n_1 + n_2 - 2$ degrees of freedom, where

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

(ii) $100(1 - \alpha)\%$ confidence interval for $\mu_1 - \mu_2$:

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} \sqrt{s_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}.$$

Example 9.4; p-417

Suppose you wish to compare a new method of teaching to slow learners with the current standard method. You decide to base your comparison on the results of a reading test given at the end of a learning period of six months. Of a random sample of 22 slow learners, 10 are taught by the new method and 12 by the standard method. All 22 children are taught by qualified instructors under similar conditions for the designated six-month period. The results of the reading test at the end of this period are given below (assume that the assumptions stated above are satisfied):

Test scores (new method): 80, 76, 70, 80, 66, 85, 79, 71, 81, 76.
Test scores (standard method): 79, 73, 72, 62, 76, 68, 70, 86, 75, 68, 73, 66.

(mean = 72.333, std dev = 6.3437); (pooled variance = 37.45)

(i) Estimate the true mean difference between the test scores using a 95% confidence interval. (-1.4, 9.53).

(ii) Conduct a test to see if the new method is better than the standard method. Use $\alpha = 0.05$.

$(H_0 : \mu_1 - \mu_2 = 0; H_a : \mu_1 - \mu_2 > 0; \text{Rejection region: } t > 1.725; t = 1.55; \text{Not sufficient evidence to conclude that the new method is better than the standard method})$

3. Inferences for means based on two correlated samples: Paired Difference Experiments

Notation: Assume a random sample of $n_d$ differences $x_d = x_1 - x_2$ from a normal population of differences with mean $\mu_d$ and variance $\sigma_d^2$. Use $\bar{x}_d$ and $s_d$ to denote the mean and standard deviation of sample differences.

(i) Test of $H_0 : \mu_d = d_0$ (known value):

$$ t = \frac{\bar{x}_d - d_0}{\frac{s_d}{\sqrt{n_d}}} \text{ with } n_d - 1 \text{ degrees of freedom.} $$

(ii) 100$(1 - \alpha)$% confidence interval for $\mu_1 - \mu_2$: 4
\[ \bar{x}_d \pm t_{\alpha} \frac{s_d}{\sqrt{n_d}} \]

Example: Exercise: 9:46, page 441

(difference=(Standard- Huffman); \( \bar{x}_d = .13 \), \( s_d = 0.13928 \), 95% CI for mean difference is (0.03643, 0.22357), t-score = 3.096, df = 10; Reject \( H_0 \))

Example:

A new weight-reducing technique, consisting of a liquid protein diet, is currently undergoing tests by the Food and Drug Administration (FDA) before its introduction into the market. A typical test performed by the FDA is the following: The weights of a random sample of five people are recorded before they are introduced to the liquid protein diet. The five individuals are then instructed to follow the liquid protein diet for 3 weeks. At the end of this period, their weights (in pounds) are again recorded. The results are listed in the table. Let \( \mu_1 \) be the true mean weight of individuals before starting the diet and let \( \mu_2 \) be the true mean weight of individuals after 3 weeks on the diet.

<table>
<thead>
<tr>
<th>Person</th>
<th>Weight Before Diet</th>
<th>Weight After Diet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>161</td>
<td>154</td>
</tr>
<tr>
<td>2</td>
<td>206</td>
<td>201</td>
</tr>
<tr>
<td>3</td>
<td>199</td>
<td>196</td>
</tr>
<tr>
<td>4</td>
<td>208</td>
<td>202</td>
</tr>
<tr>
<td>5</td>
<td>215</td>
<td>211</td>
</tr>
</tbody>
</table>
Summary information

\[ x_d = \text{difference} = (\text{before} - \text{After}); \bar{x}_d = 5, s_d = 1.58114; \]

Calculate a 90% confidence interval for the difference between the mean weights before and after the diet is used.

(95% CI for the mean difference is (3.03676, 6.96324))

Test to determine if the diet is effective at reducing weight. Use alpha = .05.

\[ (H_0 : \mu_1 - \mu_2 = 0; H_a : \mu_1 - \mu_2 > 0; \ t\text{-score} = 7.071; \ df = 4; \text{Rejection region } t > 2.132; \text{ New diet is effective in reducing weight}) \]

Example

According to research published in *Science*, the mere belief that you are receiving an effective treatment for pain can reduce the pain you actually feel. Researchers from the University of Michigan and Princeton University tested this placebo effect on 24 volunteers as follows. Each volunteer was put inside a MRI for two consecutive sessions. During the first session, electric shocks were applied to their arms and the blood oxygen level-dependent (BOLD) signal (a measure related to neural activity in the brain) was recorded during pain. The second session was identical to the first, except that prior to applying the electric shocks the researchers smeared a cream on the volunteer’s arms. The volunteers were informed that the cream would block the pain, when, in fact, it was just a regular skin lotion (i.e. placebo). If the placebo is effective in reducing the pain experience, the BOLD
measurements should be higher, on average, in the first MRI session than in the second MRI session.

What type of design was used to collect the data? What is the appropriate null and alternative hypothesis for testing the placebo effect theory? The difference between the BOLD measurements in the first and second sessions were computed and summarized in the study as follows: \( n_d = 24, \bar{x}_d = 0.21, s_d = 0.47 \). Use this information to calculate the test statistic (\( t = 2.19 \)). The p-value of the test was reported as \( p\)-value = 0.02. Make the appropriate conclusion at \( \alpha = 0.05 \).

Comparing two proportions: Independent samples

Notation: Assume two independent samples from two binomial populations with the following notations:

Sample I: Proportion = \( \hat{p}_1 \), sample size = \( n_1 \). Population I: proportion = \( p_1 \).

Sample II: Proportion = \( \hat{p}_2 \), sample size = \( n_2 \). Population II: proportion = \( p_2 \).
Sampling distribution of $\hat{p}_1 - \hat{p}_2$:

(i) The mean of the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is $\mu(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$.

(ii) Standard deviation of the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is

$$
\sigma(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}.
$$

This quantity is referred to as standard error of $(\hat{p}_1 - \hat{p}_2)$.

(iii) The sampling distribution of $(\hat{p}_1 - \hat{p}_2)$ is approximately normal for large sample sizes.

4. Large sample test and confidence interval estimation of $p_1 - p_2$.

(i) 100$(1 - \alpha)$% confidence interval for $p_1 - p_2$:

$$
\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}
$$

(ii) Test of $H_0 : p_1 - p_2 = 0$:

$$
Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where } \hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}
$$

Example
Two surgical procedures are widely used to treat a certain type of cancer. To compare the success rates of the two procedures, random samples of the two types of surgical patients were obtained and the numbers of patients who showed no recurrence of the disease after a 1-year period were recorded. The data are shown in the table.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>n</th>
<th>Number of Successes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedure A</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>Procedure B</td>
<td>100</td>
<td>96</td>
</tr>
</tbody>
</table>

(i) Estimate the difference in the true success rates with a 95% confidence interval. 

(-0.13, 0.01)

(ii) Conduct a test to see if the success rate ($p_2$) for procedure B is significantly higher than that ($p_1$) of A. Use $\alpha = 0.05$. ($H_a : p_1 - p_2 < 0; z = -1.66; \text{RR: } z < -1.645; \text{Reject } H_0 \text{ at } \alpha = 0.05$)

(iii) Find the observed significance probability. (p-value = $p(Z < -1.66) = 0.0485)$