Objectives

1. Conduct a hypothesis test of a claim about two independent population means, or
2. Construct a confidence interval estimate of the difference between two independent population means.

Notation

For population 1 we let

μ_1 = *population* mean

σ_1 = *population* standard deviation

n_1 = size of the first sample

$\bar{x}_1$ = *sample* mean

s_1 = *sample* standard deviation

The corresponding notations μ_2, σ_2, $\bar{x}_2$, s_2, and n_2 apply to population 2.

Requirements

1. The values of σ_1 and σ_2 are unknown and we do not assume that they are equal.
2. The two samples are *independent*.
3. Both samples are *simple random samples*.
4. Either or both of these conditions is satisfied: The two sample sizes are both *large* (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions. (The methods used here are *robust* against departures from normality, so for small samples, the normality requirement is loose in the sense that the procedures perform well as long as there are no outliers and departures from normality are not too extreme.)

Hypothesis Test Statistic for Two Means: Independent Samples

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{(where } \mu_1 - \mu_2 \text{ is often assumed to be 0)}$$

Degrees of freedom: When finding critical values or *P*-values, use the following for determining the number of degrees of freedom, denoted by df. (Although these two methods typically result in different numbers of degrees of freedom, the conclusion of a hypothesis test is rarely affected by the choice.)

1. In this book we use this simple and conservative estimate:

   df = **smaller of** $n_1 - 1$ **and** $n_2 - 1$

2. Statistical software packages typically use the more accurate but more difficult estimate given in Formula 9-1. (We will not use Formula 9-1 for the examples and exercises in this book.)

Formula 9-1

$$\text{df} = \frac{(A + B)^2}{\frac{A^2}{n_1 - 1} + \frac{B^2}{n_2 - 1}}$$

where $A = \frac{s_1^2}{n_1}$ and $B = \frac{s_2^2}{n_2}$

***P*-values:** *P*-values are automatically provided by technology. If technology is not available, refer to the *t* distribution in Table A-3. Use the procedure summarized in Figure 8-4 from Section 8-2.

Critical values: Refer to the *t* distribution in Table A-3.

continued