As we discovered in studying confidence intervals, the analysis of differences in means presupposes the analysis of ratios of variances. Thus, we begin this module with the study of hypothesis tests for ratios of variances.

Suppose we have two independent random variables $X_1$ and $X_2$ with means $\mu_1$ and $\mu_2$, and variances $\sigma_1^2$ and $\sigma_2^2$, and suppose we want to test the hypothesis that the variances are equal. For example, suppose we want to test the hypothesis that $\sigma_1^2 = \sigma_2^2$. We can formulate this hypothesis as a ratio of variances, as follows:

$$H_0: \sigma_1^2/\sigma_2^2 = (\sigma_1^2/\sigma_2^2)_0$$

versus one of

- $H_A: \sigma_1^2/\sigma_2^2 \neq (\sigma_1^2/\sigma_2^2)_0$
- $H_A: \sigma_1^2/\sigma_2^2 > (\sigma_1^2/\sigma_2^2)_0$
- $H_A: \sigma_1^2/\sigma_2^2 < (\sigma_1^2/\sigma_2^2)_0$

Following the logic used for confidence intervals for ratios of variances (in Lecture 2.2.2), we note that the random variable

$$F = \frac{s_1^2}{s_2^2} \left(\frac{\sigma_1^2}{\sigma_2^2}\right)_0$$

is $F$ with $n_1-1$ numerator degrees of freedom and $n_2-1$ denominator degrees of freedom.

The procedure is best seen via an example. Suppose we have the hypothesis that the variances are equal, i.e., that $(\sigma_1^2/\sigma_2^2)_0 = 1$, and suppose we have the samples $\{1, 2, 3, 4, 5\}$ and $\{-3, -1, 5, 7, 12\}$. The procedure is as follows:

1. State the null hypothesis, the alternative hypothesis, and the significance level:

   $$H_0: \sigma_1^2/\sigma_2^2 = 1$$
   $$H_A: \sigma_1^2/\sigma_2^2 \neq 1$$
   $$\alpha = .05$$

2. Determine the critical region

   Given $H_A: \sigma_1^2/\sigma_2^2 \neq 1$, the critical region is $CR = \{F \mid F < F_{1-\alpha/2}\} \cup \{F \mid F > F_{\alpha/2}\}$

   For our example $n_1 = 5$ and $n_2 = 5$, so the $F$ variable has 5-1 = 4 numerator degrees of freedom and 5-1 = 4 denominator degrees of freedom, and the critical region is $\{F \mid F < (1/9.60)\} \cup \{F \mid F > 9.60\}$. Graphically, the critical region is as follows:
(3) Determine the value of the test statistic
The samples \{1, 2, 3, 4, 5\} and \{-3, -1, 5, 7, 12\} yield
\[
\bar{x}_1 = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3,
\]
\[
s_1^2 = \frac{(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2}{5-1} = \frac{10}{4} = 2.5,
\]
\[
\bar{x}_2 = \frac{-3+(-1)+5+7+12}{5} = \frac{20}{5} = 4,
\]
\[
s_2^2 = \frac{(-3-4)^2 + (-1-4)^2 + (5-4)^2 + (7-4)^2 + (12-4)^2}{5-1} = \frac{148}{4} = 37,
\]
so the value of the test statistic is
\[
F^* = \frac{s_1^2 / s_2^2}{\left(\sigma_1^2 / \sigma_2^2\right)_0} = \frac{2.5}{37} = 0.0675.
\]

(4) Locate the value of the test statistics relative to the critical region
Since 0.0675 < 0.104, \(F^*\) is in the critical region.

Graphically, we have the following:
(5) Make the decision

Since $F^*$ is in the critical region, we reject the null hypothesis $H_0$: $\frac{\sigma_1^2}{\sigma_2^2} = 1$.

As noted, testing the equality of variances is prior to testing the equality of means. We have found that the variances are not equal. Clearly, the sample sizes are equal. Thus, our samples will lead us to test for the equality of means as an instance of Case 2, the case where $\sigma_1^2 \neq \sigma_2^2$ and $n_1 = n_2$.

PROBLEMS:
Ch. 4, #21, 23, 25, 26