Section 9.5: Inferences Concerning Two Population Variances
The F-distribution. Assume $X \sim \chi^2_m$ and $Y \sim \chi^2_n$ independently. Then,

$$F = \frac{X/m}{Y/n} \sim F_{m,n},$$

where $F_{m,n}$ represents the F-distribution with $m$ degree of freedom on the numerator and $n$ degree of freedom on the denominator.

F-distribution has the following properties:

- If $F \sim F_{m,n}$, then

$$E(F) = \frac{n}{n-2}; n > 2$$

and

$$V(F) = \frac{2n^2(m + n - 2)}{m(n-2)^2(n-4)}; n > 4.$$ 

- Therefore, if $F \sim F_{m,n}$ and $n$ is large, we expected to see $F$ the observed value of $F$ around 1.
• If $F \sim F_{m,n}$, then $1/F \sim F_{n,m}$. This implies that

$$P(F_{m,n} < c) = P(F_{n,m} > 1/c)$$

which gives

$$F_{\alpha,m,n} = 1/F_{1-\alpha,n,m}$$

where $F_{\alpha,m,n}$ represents the upper $\alpha$ quantile of the F-distribution with $m$ and $n$ degrees of freedom respectively. For example, if we know

$$F_{0.05,10,8} = 3.347$$

then we have

$$F_{0.95,8,10} = \frac{1}{3.347} = 0.2988.$$  

• The values of $F_{\alpha,m,n}$ for selected $\alpha$, $m$ and $n$ can be found in Table A.9 on textbook.
The method. Suppose

\[ X_1, X_2, \ldots, X_m \sim \text{iid } N(\mu_1, \sigma_1^2) \]

and

\[ Y_1, Y_2, \ldots, Y_n \sim \text{iid } N(\mu_2, \sigma_2^2) \]

Then,

\[ F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{m-1,n-1}. \]

Thus, the \((1 - \alpha)100\%\) confidence interval for \(\frac{\sigma_1^2}{\sigma_2^2}\) is

\[ \left[ \frac{s_1^2/s_2^2}{F_{\alpha/2,m-1,n-1}}, \frac{s_1^2/s_2^2}{F_{1-\alpha/2,m-1,n-1}} \right]. \]

To test

\[ H_0 : \sigma_1^2 = \sigma_2^2 \leftrightarrow H_a : \sigma_1^2 \neq \sigma_2^2, \]

We reject \(H_0\) and conclude \(H_a\) if

\[ \frac{s_1^2}{s_2^2} > F_{\alpha/2,m-1,n-1} \]

or

\[ \frac{s_1^2}{s_2^2} < F_{1-\alpha/2,m-1,n-1}. \]
Example 1 of Section 9.5: example 9.14 on textbook. In this example, we observe $s_1 = 52.6$, $m = 28$, $s_2 = 84.2$ and $n = 26$. Suppose we use $\alpha = 0.1$.

Then, the 90% confidence interval is

- Then, we need $F_{0.95,27,25}$ and $F_{0.05,27,25}$.
- However, $DF_1 = m - 1 = 27$ is not available. We can either roughly choose $m - 1 = 25$ or $m - 1 = 30$.
- Suppose we use $m - 1 = 30$. Then we use $F_{0.95,30,25}$ and $F_{0.05,30,25}$.
- Table A.9 directly gives

\[
F_{0.05,30,25} = 1.92
\]

However, it does not directly gives $F_{0.95,30,25}$.

- We can derive it by

\[
F_{0.95,30,25} = \frac{1}{F_{0.05,25,30}} = \frac{1}{1.89} = 0.529.
\]
• The 90% confidence interval of $\sigma_1^2/\sigma_2^2$ is

$$\left[\frac{52.6^2/84.2^2}{1.92}, \frac{52.6^2/84.2^2}{0.529}\right] = [0.2033, 0.7377].$$

• Since

$$\frac{s_1^2}{s_2^2} = 0.3903 < F_{0.95,30,25} = 0.529,$$

we reject $H_0: \sigma_1^2 = \sigma_2^2$ and conclude $\sigma_1^2 \neq \sigma_2^2$. 