Testing for Equality of Two Variances

In previous test, assumed variances from two populations were the same. Sometimes we want to test this

H_0: $\sigma_1^2 = \sigma_2^2$
H_1: $\sigma_1^2 \neq \sigma_2^2$

Assume the 2 samples are independent and come from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ distributions.

Usually, we don't know σ_1^2 or σ_2^2. We will base test on sample estimates, s_1^2, s_2^2.

Best test is based on ratio s_1^2 / s_2^2. We reject H_0 if the ratio is very small or very large. To do this, we first have to describe the sampling distribution of s_1^2 / s_2^2.

This ratio follows an F-distribution under H_0. This is a family of distributions dependent on 2 parameters called the <u>numerator degrees of freedom</u> and the <u>denominator degrees of freedom</u>.

If samples are of size n_1 and n_2, then
$s_1^2 / s_2^2 \sim F(n_1-1, n_2-1)$.

$$F_{n_1-1, n_2-1}$$

Most *F* distributions are positively skewed and the amount of skew depends on the relative magnitudes of the two degrees of freedom. The only exceptions are when the numerator degrees of freedom are either 1 or 2. The mode is then at 0. The *F* distribution takes positive values only.

The *F* distribution is well tabled (See Table 9 for selected degrees of freedom).

This table only gives the upper percentage points

$$\Pr\left(\frac{s_1^2}{s_2^2} \geq F_{n_1-1, n_2-1, 1-p}\right) = p$$

We can also show that

$$F_{n_1-1, n_2-1, p} = \frac{1}{F_{n_2-1, n_1-1, 1-p}}$$

Therefore, we can also compute the lower percentiles because of this symmetry property.

Example: Suppose $n_1 = 9$, $n_2 = 13$.

$$F_{8,12,0.975} = 3.51$$
$$F_{12,8,0.975} = 4.20$$
$$F_{8,12,.025} = \frac{1}{F_{12,8,0.975}} = \frac{1}{4.20} = 0.2381$$
$$F_{12,8,.025} = \frac{1}{F_{8,12,0.975}} = \frac{1}{3.51} = 0.2849$$

Practice using Table 9.

num df = 8, denom df = 16

Critical values: $F_{.95} = 2.59$, $F_{.975} = 3.12$, $F_{.995} = 4.52$

num df = 6, denom df = 16

Critical values: $F_{.95} = 2.74$, $F_{.975} = 3.34$, $F_{.995} = 4.91$

To return to the *F*-test:

H_0: $\sigma_1^2 = \sigma_2^2$
H_1: $\sigma_1^2 \neq \sigma_2^2$

We will simplify test. We will always conduct test by putting larger sample variance in the numerator. Therefore, we will only reject H_0 for large values of *F*. Reject H_0 if

$$F > F_{n_1-1,n_2-1,1-\alpha/2}$$

For example, if $n_1 = 13$ and $n_2 = 21$, $\alpha = 0.05$

$$F_{n_1-1,n_2-1,1-\alpha/2} = F_{12,20,.975} = 2.68$$

If s_1^2 is the larger variance, reject H_0 if

$$\frac{s_1^2}{s_2^2} > 2.68$$

Note if $\alpha = 0.01$, critical value is $F_{12,20,.995} = 3.68$

As numerator degrees of freedom increases, the critical value decreases

As denominator degrees of freedom increases, the critical value decreases

Example: Behr investigated alterations of thermoregulation in patients with certain pituitary adenomas. The standard deviation of the weights of a sample of 13 patients was 21.4 kg. The weights of a sample of 5 control subjects yielded a standard deviation of 12.4 kg. We wish to know if we may conclude that the weights of the two populations have the same variability.

H_0: $\sigma_1^2 = \sigma_2^2$
H_1: $\sigma_1^2 \neq \sigma_2^2$

Let $\alpha = 0.05$

Reject H_0 if $F > F_{12,4,0.975} = 8.75$

Test Statistic:

$$F = \frac{s_1^2}{s_2^2}$$
$$= \frac{21.4^2}{12.4^2}$$
$$= 2.98$$

Decision: Do not reject H_0. The weights of the two populations do not appear to have different variability.

One-tailed alternatives:

H_0: $\sigma_1^2 = \sigma_2^2$	H_0: $\sigma_1^2 = \sigma_2^2$
H_1: $\sigma_1^2 > \sigma_2^2$	H_1: $\sigma_1^2 < \sigma_2^2$
Ratio: s_1^2 / s_2^2	Ratio: s_2^2 / s_1^2

Note that the larger variance is always in the numerator. For a one-tailed alternative, the critical value for F is obtained for $1-\alpha$, not $1-\alpha/2$.