Two Sample Problem: the ratio of two population variances

Suppose that we have two **independent** random samples:

$$\{X_1, \ldots, X_n\} \text{ from } N(\mu_X, \sigma_X^2), \quad \text{and} \quad \{Y_1, \ldots, Y_m\} \text{ from } N(\mu_Y, \sigma_Y^2),$$

where μ_X, μ_Y, σ_X^2 and σ_Y^2 are unknown, n and m are respective sample sizes of X and Y.

Now we want to compare the two population variances by considering a $100(1-\alpha)\%$ C.I. for σ_X^2/σ_Y^2.

First, we consider the intuitive point estimator S_X^2/S_Y^2 for σ_X^2/σ_Y^2, where $S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ and $S_Y^2 = \frac{1}{m-1}\sum_{j=1}^m (Y_j - \overline{Y})^2$.

Since the two random samples are from the normal distribution, we can establish the confidence interval of σ_X^2/σ_Y^2 by using the result that

$$\frac{\sigma_Y^2 S_X^2}{\sigma_X^2 S_Y^2}$$

has a F distribution with $n-1$ and $m-1$ degrees of freedom. Notationally, we have

$$F \overset{def}{=} \frac{\sigma_Y^2 S_X^2}{\sigma_X^2 S_Y^2} \sim F(n-1, m-1)$$

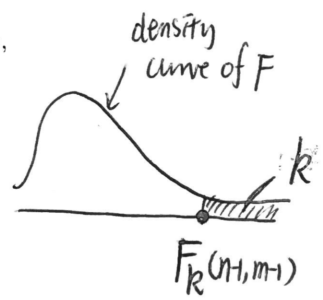
How to construct a confidence interval for σ_X^2/σ_Y^2?

Starting from

$$P(F_{1-\alpha/2}(n-1, m-1) < F < F_{\alpha/2}(n-1, m-1)) = 1 - \alpha,$$

where $F_{1-\alpha/2}(n-1, m-1)$ and $F_{\alpha/2}(n-1, m-1)$ are quantities such that

$$P(F > F_{1-\alpha/2}(n-1, m-1)) = 1 - \alpha/2 \quad \text{and}$$
$$P(F > F_{\alpha/2}(n-1, m-1)) = \alpha/2,$$



respectively. Thus, substituting for $F \overset{def}{=} \frac{\sigma_Y^2 S_X^2}{\sigma_X^2 S_Y^2}$, we have

$$P(F_{1-\alpha/2}(n-1, m-1) < \frac{\sigma_Y^2 S_X^2}{\sigma_X^2 S_Y^2} < F_{\alpha/2}(n-1, m-1)) = 1 - \alpha.$$

After some algebra, we obtain

$$P(\frac{1}{F_{\alpha/2}(n-1, m-1)} \frac{S_X^2}{S_Y^2} < \frac{\sigma_X^2}{\sigma_Y^2} < \frac{1}{F_{1-\alpha/2}(n-1, m-1)} \frac{S_X^2}{S_Y^2}) = 1 - \alpha.$$

Using a property of the F distribution that

$$F_{\alpha/2}(m-1, n-1) = \frac{1}{F_{1-\alpha/2}(n-1, m-1)}$$

finally we have that after sampling, a $100(1-\alpha)\%$ C.I. for $\frac{\sigma_X^2}{\sigma_Y^2}$ is given by

$$\left[\frac{1}{F_{\alpha/2}(n-1,\, m-1)}\frac{s_X^2}{s_Y^2},\quad F_{\alpha/2}(m-1,\, n-1)\frac{s_X^2}{s_Y^2}\right].$$

Example: On page 18, we had data sets of the weight of two independent random samples of tomatoes grown using each of the two fertilizers (in ounces):

Fertilizer A (X): 12, 11, 7, 13, 8, 9, 10, 13

Fertilizer B (Y): 13, 11, 10, 6, 7, 4, 10

Based on the assumption of same population (unknown) variances, we got a 95% confidence interval, [-1.366, 4.688], for the mean difference $\mu_X - \mu_Y$. Here we want to justify the assumption of the same population variances by constructing a 98% confidence interval of σ_X^2/σ_Y^2.

From the results we found before, we have $n = 8$, $s_X^2 = 5.125$, $m = 7$, and $s_Y^2 = 9.905$. Thus, a 98% confidence interval of σ_X^2/σ_Y^2 is given by

$$\left[\frac{1}{F_{0.01}(7,\, 6)}\frac{5.125}{9.905},\quad F_{0.01}(6,\, 7)\frac{5.125}{9.905}\right] = [0.0626,\ 3.7202],$$

where $F_{0.01}(7,6) = 8.26$ and $F_{0.01}(6,7) = 7.19$. Note that 1 is inside the interval, so our result allows for the possibility of σ_X^2/σ_Y^2 being equal to 1, i.e., we were correct in assuming that $\sigma_X^2 = \sigma_Y^2$ before.