CONFIDENCE INTERVALS FOR VARIANCE COMPONENTS (Section 17.3.5)

In practice, these play the role for random effects that confidence intervals for contrasts play for fixed effects.

Confidence intervals for $\boldsymbol{\sigma^2}$: These are constructed just as for fixed effects; see Section 3.4.6 or the class notes *Choosing Sample Sizes*.

Confidence intervals involving $\boldsymbol{\sigma_T^2}$: Three types of confidence intervals are of interest: for $\boldsymbol{\sigma_T^2}$, for $\boldsymbol{\sigma_T^2/\sigma^2}$, and for $\boldsymbol{\sigma_T^2/(\sigma_T^2 + \sigma^2)}$. The first cannot be done exactly, so we'll take that last.

Confidence intervals for $\boldsymbol{\sigma_T^2/\sigma^2}$: We use the fact (see notes *Testing for Treatment Effect as a Proportion of Error Variance)* that

$$\frac{MST/(c\sigma_T^2+\sigma^2)}{MSE/(\sigma^2)} \sim F(v-1, n-v),$$

where c is a certain constant defined in terms of n,v, and the r_i; c = r if the design is balanced) (See notes *Random Effects Models* or Section 17.3)

If we want a (1-α)100% CI for $\boldsymbol{\sigma_T^2/\sigma^2}$, take

f_1 = F(v-1, n-v, 1- α/2) (so that there is area α/2 to the left of f_1 in the F(v-1, n-v) distribution), and

f_2 = F(v-1, n-v, α/2) (so that there is area α/2 to the right of f_2 in the F(v-1, n-v) distribution). [Draw a picture!]

Then

$$\text{Prob}\left(f_1 \le \frac{MST/(c\sigma_T^2+\sigma^2)}{MSE/(\sigma^2)} \le f_2\right) = 1-\alpha,$$

or equivalently,

$$\text{Prob}\left(f_1 \le [MST/MSE]\,[\boldsymbol{\sigma^2/(\sigma_T^2+\sigma^2)}] \le f_2\right) = 1-\alpha,$$

The left inequality is equivalent to

$$(c\boldsymbol{\sigma_T^2} + \boldsymbol{\sigma^2})/\boldsymbol{\sigma^2} \le (MST/MSE)(1/f_1), \text{ or}$$

$$c(\boldsymbol{\sigma_T^2/\sigma^2}) + 1 \le (MST/MSE)(1/f_1),$$

which is equivalent to

$$c(\boldsymbol{\sigma_T^2/\sigma^2}) \le (MST/MSE)(1/f_1) - 1$$

The right inequality is equivalent to

$(MST/MSE)(1/f_2) \le (c\boldsymbol{\sigma_T^2} + \boldsymbol{\sigma^2})/\boldsymbol{\sigma^2} = c(\boldsymbol{\sigma_T^2}/\boldsymbol{\sigma^2}) + 1$,

which is equivalent to

$(MST/MSE)(1/f_2) - 1 \le c(\boldsymbol{\sigma_T^2}/\boldsymbol{\sigma^2})$

So

Prob $((1/c)[\,(MST/MSE)(1/f_2) - 1] \le \boldsymbol{\sigma_T^2}/\boldsymbol{\sigma^2} \le (1/c)[\,(MST/MSE)(1/f_1) - 1]) = 1 - \alpha$.

Thus $((1/c)[\,(msT/msE)(1/f_2) - 1]\,,\,(1/c)\,(msT/msE)(1/f_1) - 1)$ is the desired confidence interval. (This means ____)

Note: Conceivably the left hand endpoint could be less than 0, which is unrealistic, If it is < 0, do *not* give in to the temptation to replace it by zero; that would give the false impression of a smaller confidence interval than warranted.

Example: Use the loom data to find a 95% confidence interval for $\boldsymbol{\sigma_T^2}/\boldsymbol{\sigma^2}$.

Confidence intervals for $\boldsymbol{\sigma_T^2/(\sigma_T^2 + \sigma^2)}$ *= the proportion of the total variance if the response attributable to the treatment level*: Such confidence intervals are readily obtained from confidence intervals for $\boldsymbol{\sigma_T^2}/\boldsymbol{\sigma^2}$ as follows. Divide both numerator and denominator of $\boldsymbol{\sigma_T^2}/(\boldsymbol{\sigma_T^2} + \boldsymbol{\sigma^2})$ by $\boldsymbol{\sigma^2}$ to obtain

$$\boldsymbol{\sigma_T^2}/(\boldsymbol{\sigma_T^2} + \boldsymbol{\sigma^2}) = \frac{\sigma_T^2/\sigma^2}{\left(\sigma_T^2/\sigma^2\right)+1} = f(\sigma_T^2/\sigma^2), \text{ where } f(x) = x/(x + 1)) = \frac{1}{1+\frac{1}{x}}.$$

From the last formula for f(x), we can see that f(x) is an increasing function of x. Thus if (a,b) is a $(1-\alpha)100\%$ confidence interval for σ_T^2/σ^2, then $(f(a), f(b)) = (a/(a + 1), b/(b + 1))$ is a $(1-\alpha)100\%$ confidence interval for $\boldsymbol{\sigma_T^2}/(\boldsymbol{\sigma_T^2} + \boldsymbol{\sigma^2})$.

Note: $\boldsymbol{\sigma_T^2/(\sigma_T^2 + \sigma^2)}$ is sometimes called the "population intraclass correlation coefficient" (*Caution*: The phrase "intraclass correlation coefficient" is also used to refer to other things.)

Example: With the loom data, find a 95% confidence interval for $\boldsymbol{\sigma_T^2}/(\boldsymbol{\sigma_T^2} + \boldsymbol{\sigma^2})$.

Confidence intervals for $\boldsymbol{\sigma_T^2}$: There is no exact method. There are several approximate methods. Here is one. It is useful if $\boldsymbol{\sigma_T^2}$ is not too small, and is adaptable to more complicated models.

Recall that $U = (1/c)(MST - MSE)$ is an unbiased estimator of $\boldsymbol{\sigma_T^2}$. If we knew its distribution, we could use that to get confidence intervals for $\boldsymbol{\sigma_T^2}$ in the usual way. However, it does not have a tractable distribution. But it *is* true that

$U/\boldsymbol{\sigma_T^2} \approx \chi^2(x)/x$, where

$$x \approx \frac{(msT - msE)^2}{\frac{(msT)^2}{(v-1)} + \frac{(msE)^2}{(n-v)}}$$

(Note: This formula is given correctly on p. 605 of the text, but incorrectly on p. 600.) x is not usually an integer, so we need to interpret degrees of freedom in the χ^2 distribution as the parameter in a formula for the pdf. (This is analogous to the two-sample, unequal variance t-test.)

Thus (Draw a picture!)

$$P(\chi^2(x, 1-\alpha/2) \lesssim xU/\boldsymbol{\sigma_T^2} \lesssim \chi^2(x, \alpha/2)) \approx 1-\alpha,$$

where $\lesssim$ means "is less than or approximately equal to", and $\chi^2(x,\beta)$ is the value with proportion β of the $\chi^2(x)$ distribution to its *right*.

The left and right approximate inequalities are, respectively, equivalent to

$$\boldsymbol{\sigma_T^2} \lesssim xU/\chi^2(x, 1-\alpha/2) \qquad \text{and} \qquad \boldsymbol{\sigma_T^2} \lesssim xU/\chi^2(x, \alpha/2).$$

Thus if

u = (1/c)(msT – msE) (which is our estimate for $\boldsymbol{\sigma_T^2}$), then

$(xu/\chi^2(x, \alpha/2), xu/\chi^2(x, 1-\alpha/2))$ is an approximate $(1-\alpha)100\%$ confidence interval for $\boldsymbol{\sigma_T^2}$.

<u>Example</u>: With the loom data, find a 95% confidence interval for $\boldsymbol{\sigma_T^2}$.