

## ESTIMATING VARIANCE COMPONENTS FOR RANDOM EFFECTS (Sections 17.6.4 and 17.6.5)

**Unbiased estimators:** We will illustrate with a couple of models; analogous methods work with other models.

I. In the two-way complete model:

1. We have already seen that  $E(\text{MSE}) = \sigma^2$ , so MSE is an unbiased estimator for  $\sigma^2$ .

2. From  $E(\text{MSAB}) = r\sigma_{AB}^2 + \sigma^2$  and  $E(\text{MSE}) = \sigma^2$ , we get

$$E\left(\frac{\text{MSAB} - \text{MSE}}{r}\right) = \sigma_{AB}^2$$

3. From the above and  $E(\text{MSA}) = br\sigma_A^2 + r\sigma_{AB}^2 + \sigma^2$ , we get

$$E\left(\frac{\text{MSA} - \text{MSAB}}{br}\right) = \sigma_A^2.$$

4. Similarly,  $E\left(\frac{\text{MSB} - \text{MSAB}}{ar}\right) = \sigma_B^2$

II. In the model with model equation  $Y_{ijkl} = \mu + A_i + B_j + C_k + (AB)_{ij} + (BC)_{jk} + \varepsilon_{ijl}$ :  
How would we find an unbiased estimator of  $\sigma_B^2$ ?

**Confidence intervals:** First, use the method above to obtain an unbiased estimator  $U$  of  $\sigma_*^2$  (where  $*$  = A, B, AB, etc.).  $U$  will be some linear combination of mean squares – say  $U = \sum k_i(\text{MS})_i$ . Let  $(ms)_i$  be the value of  $(\text{MS})_i$  from the sample obtained in the experiment and let  $x_i$  be the degrees of freedom of  $(\text{MS})_i$ . Then it is known that  $xU/\sigma_*^2 \approx \chi^2(x)$ , where

$$x = \frac{\left(\sum k_i (ms)_i\right)^2}{\sum k_i^2 (ms)_i^2 / x_i}.$$

We can use this, following the method we used to find an approximate CI for  $\sigma_T^2$  in the one-random effects model, to get an approximate CI for  $\sigma_*^2$ .

*Example:* With the fiber breaking strength data, find an approximate 95% CI for  $\sigma_A^2$ , where A is operator. We have  $E([\text{MSA} - \text{MSAB}]/(4 \cdot 2)) = \sigma_A^2$ , so  $U = (1/8)\text{MSA} - (1/8)\text{MSAB}$  and ...

Questions: Comment on the size of this CI. Is it symmetric about the point estimate?

**Variance components on Minitab:** Run the fiber data, but clicking “Storage” (or “Options”) and “Display expected mean squares.”

Source	Variance component	Error term	Expected Mean Square (using unrestricted model)
1 Operator	9.0903	3	$(4) + 2(3) + 8(1)$
2 Machine	-0.5486	3	$(4) + 2(3) + 6(2)$
3 Operator*Machine	1.8264	4	$(4) + 2(3)$
4 Error	3.7917		(4)

a. Note “Variance components” – does the value for Operator agree with the above calculation?

b. Note the variance component for Machine. Is this good or bad? What does this have to do with random variability? How does this relate to the p-value in the output?

c. Where does most of the variance come from?

d. Recall the hypothesis tests:

For Operator and Machine – denominator was MSAB – item 3.

For Operator\*Machine – denominator was MSE – item 4.

e. The last column contains a shorthand for the expected mean squares. For example, recall that in this example,  $E(MSA) = b r \sigma_A^2 + r \sigma_{AB}^2 + \sigma^2$ ,  $r = 2$ ,  $a = 3$ , and  $b = 4$ . Read off  $E(MSB)$  from the shorthand.