

## 6.1: Random Effects

When a treatment (or factor) is a random effect, the model specifications together with relevant null and alternative hypotheses will have to be changed. Recall the cell means model defined in Chapter 4 for the fixed effect case, which has the model equation:

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad (6.1.1)$$

where  $\mu_i$  are parameters for the treatment means.

For the *single factor random effects model* we have:

$$Y_{ij} = \mu_i + \epsilon_{ij} \quad (6.1.2)$$

where  $\mu_i$  and  $\epsilon_{ij}$  are independent random variables such that  $\mu_i \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma_\mu^2)$  and  $\epsilon_{ij} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$ . Here,  $i = 1, 2, \dots, T$  and  $j = 1, 2, \dots, n_i$ , where  $n_i \equiv n$  if balanced.

Notice that the random effects ANOVA model is similar in appearance to the fixed effects ANOVA model. However, the treatment mean  $\mu_i$ 's are constant in the fixed-effect ANOVA model, whereas in the random-effects ANOVA model the treatment mean  $\mu_i$ 's are random variables.

Note that the expected mean response, in the random effects model stated above, is the same at every treatment level and equals  $\mu$ .

$$E(Y_{ij}) = E(\mu_i + \epsilon_{ij}) = E(\mu_i) + E(\epsilon_{ij}) = \mu \quad (6.1.3)$$

The variance of the response variable (say  $\sigma_Y^2$ ) in this case can be partitioned as:

$$\sigma_Y^2 = V(Y_{ij}) = V(\mu_i + \epsilon_{ij}) = V(\mu_i) + V(\epsilon_{ij}) = \sigma_\mu^2 + \sigma_\epsilon^2 \quad (6.1.4)$$

as  $\mu_i$  and  $\epsilon_{ij}$  are independent random variables.

Similar to fixed effects ANOVA model, we can express the random effects ANOVA model using the factor effect representation, using  $\tau_i = \mu_i - \mu$ . Therefore the factor effects representation of the random effects ANOVA model would be:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij} \quad (6.1.5)$$

where  $\mu$  is a constant overall mean, and  $\tau_i$  and  $\epsilon_{ij}$  are independent random variables such that  $\tau_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_\mu^2)$  and  $\epsilon_{ij} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$ . Here,  $i = 1, 2, \dots, T$  and  $j = 1, 2, \dots, n_i$ , where  $n_i \equiv n$  if balanced. Here,  $\tau_i$  is the effect of the randomly selected  $i^{th}$  level.

The terms  $\sigma_\mu^2$  and  $\sigma_\epsilon^2$  are referred to as **variance components**. In general, as will be seen later in more complex models, there will be a variance component associated with each effect involving at least one random factor.

Variance components play an important role in analyzing random effects data. They can be used to verify the significant contribution of each random effect to the variability of the response. For the single factor random-effects model stated above, the appropriate null and alternative hypothesis for this purpose is:

$$H_0 : \sigma_\mu^2 = 0 \text{ vs. } H_A : \sigma_\mu^2 > 0 \quad (6.1.6)$$

Similar to the fixed effects model, an ANOVA analysis can then be carried out to determine if  $H_0$  can be rejected.

The MS and the df computations of the ANOVA table are the same for both the fixed and random-effects models. However, the computations of the F-statistics needed for hypothesis testing require some modification.

Specifically, the F statistics denominator will no longer always be the mean squared error (MSE or MSERROR) and will vary according to the effect of interest (listed in the Source column of the ANOVA table). For a random-effects model, the quantities known as **Expected Means Squares (EMS)**, shown in the ANOVA table below, can be used to identify the appropriate F-statistic denominator for a given source in the ANOVA table. These EMS quantities will also be useful in estimating the variance components associated with a given random effect. Note that the EMS quantities are in fact the population counterparts of the mean sums of squares (MS) that we are already familiar with. In SAS the `proc mixed, method=type3` option will generate the EMS column in the ANOVA table output.

Source	df	SS	MS	F	P	EMS (Expected Means Squares)
Trt						$\sigma_{\epsilon}^2 + n\sigma_{\mu}^2$
Error						$\sigma_{\epsilon}^2$
Total						

#### Note

Variance components are NOT synonymous with mean sums of squares. Variance components are usually estimated by using the Method of Moments where algebraic equations, created by setting the mean sums of squares (MS) equal to the EMS for the relevant effects, are solved for the unknown variance components. For example, the variance component for the treatment in the single-factor random effects discussed above can be solved as:

$$s_{\text{among trts}}^2 = \frac{MS_{\text{trt}} - MS_{\text{error}}}{n} \quad (6.1.7)$$

This is by using the two equations:

$$MS_{\text{error}} = \sigma_{\epsilon}^2$$

$$MS_{\text{trt}} = \sigma_{\epsilon}^2 + n\sigma_{\mu}^2$$

#### More about variance components...

Often the variance component of a specific effect in the model is expressed as a percent of the total variation of the variation in the response variable.

Another common application of variance components is when researchers are interested in the relative size of the treatment effect compared to the within-treatment level variation. This leads to a quantity called the **intraclass correlation coefficient (ICC)**, defined as:  $ICC = \frac{\sigma_{\text{among trts}}^2}{\sigma_{\text{among trts}}^2 + \sigma_{\text{within trts}}^2}$

For single random factor studies,  $ICC = \frac{\sigma_{\mu}^2}{\sigma_{\mu}^2 + \sigma_{\epsilon}^2}$ . ICC can also be thought of as the correlation between the observations within the group (i.e.  $\text{corr}(Y_{ij}, Y_{ij'})$ , where  $j \neq j'$ ). Small values of ICC indicate a large spread of values at each level of the treatment, whereas large values of ICC indicate relatively little spread at each level of the treatment:

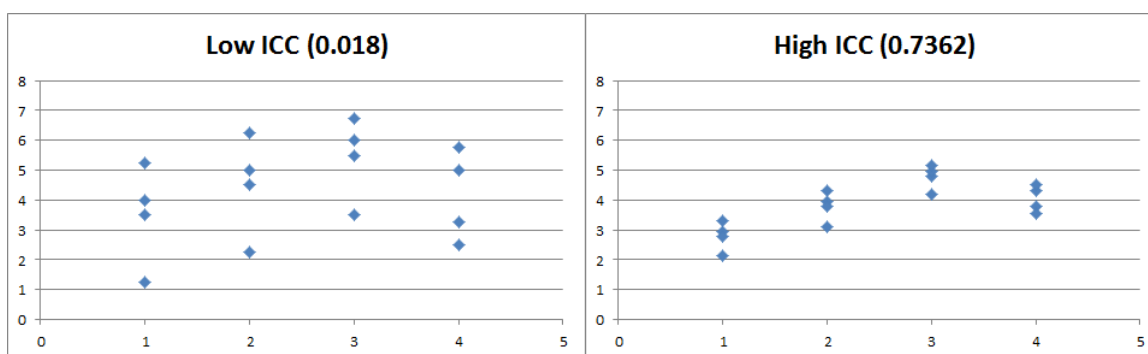


Figure 6.1.1: Dot plots for data sets with low and high ICC values.

This page titled [6.1: Random Effects](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Penn State's Department of Statistics](#).