Completely Randomized (One Way) Design — Random Effects

Up to this point in time we have examined the model $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ where the $\alpha_i$ were assumed to represent fixed effects, such as arise when we want to compare specific treatments or groups. We now consider the one-way random effects model, which is used when we want to make inferences to a population from which the $k$ groups represent a random sample.

Model

$Y_{ij} = \mu + A_i + e_{ij}$

where

- $i = 1, \ldots, k$ indexes treatment (or group)
- $j = 1, \ldots, n$ indexes experimental unit (plot) number for each treatment
- $A_i \sim N(0, \sigma^2_A)$ corresponds to the random effect (varying from group to group)
- $e_{ij} \sim N(0, \sigma^2_e)$ corresponds to error within each group

ANOVA Table

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trts (Groups)</td>
<td>$k - 1$</td>
<td>SSTrt</td>
<td>MSTrt</td>
<td>$\sigma^2_A + n\sigma^2_e$</td>
</tr>
<tr>
<td>Error</td>
<td>$k(n - 1)$</td>
<td>SSErr</td>
<td>MSErr</td>
<td>$\sigma^2_e$</td>
</tr>
<tr>
<td>Total</td>
<td>$kn - 1$</td>
<td>SSTot</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes

- The sums of squares are exactly the same as for the one-way fixed effects design.
- The model is sometimes described as one where we are not interested in the specific group effects, but rather in the population of groups which is represented by the sample that we have. For example, the $A_i$ might represent a sample of different sixth grade classes in a county, and the $e_{ij}$ might represent a sample of students within each class. $Y_{ij}$ could represent the reading scores of the students. In that case, our focus might be on $\mu$, the average reading score, or on $\sigma^2_A$, the variation from school to school.
- This model can also be viewed as a model for subsampling: the $A_i$s represent variation among samples, and the $e_{ij}$ represent variation among subsamples. For example, we may be interested in the mean blood sugar level ($\mu$) for a population. We sample $k$ persons (represented by the $A_i$s) and, for each person, take $n$ blood readings (whose variation is captured by the $e_{ij}$). The persons represent a sample from the population, and the individual readings are subsamples from each person.
- A $1 - \alpha$ confidence interval for $\mu$ in this model is given by $\bar{Y} \pm T_{k-1,\alpha/2} \sqrt{\frac{\text{MSTrt}}{nk}}$.
- We can estimate $\sigma^2_e$ using MSErr, and estimate $\sigma^2_A$ using:

$$\hat{\sigma}^2_A = \max \left\{ \frac{\text{MSTrt} - \text{MSErr}}{n}, 0 \right\}.$$  

- We test $H_0: \sigma^2_A = 0$ vs $H_A: \sigma^2_A > 0$ using $F = \text{MSTrt}/\text{MSErr}$, analogous to the fixed effects case.
- When the focus is on subsampling (as opposed to a one-way random effects situation) the model is sometimes written $Y_{ij} = \mu + \epsilon_i + \delta_{ij}$ where $\epsilon_i \sim N(0, \sigma^2_e)$ and $\delta_{ij} \sim N(0, \sigma^2_A)$.  

1