Analysis with Multiple Observations Per Experimental Unit
(Balanced Data Case)

A soil scientist studied the growth of barley plants under three different levels of salinity in a controlled growth medium. Each salinity treatment was assigned to two containers using a completely randomized design. Three plants were measured in each container. The data on the dry weight of the plants in grams follow.

Salinity	Container	Weight in grams	i	j	k
control	1	11.29, 11.08, 11.10	1	1	1, 2, 3
control	2	7.37, 6.55, 8.50	1	2	1, 2, 3
6 bars	3	5.64, 5.98, 5.69	2	1	1, 2, 3
6 bars	4	4.20, 3.34, 4.21	2	2	1, 2, 3
12 bars	5	4.83, 4.77, 5.66	3	1	1, 2, 3
12 bars	6	3.28, 2.61, 2.69	3	2	1, 2, 3

1. What are the experimental units and observational units in this example?

Consider the following model for the data.

$$y_{ijk} = \mu + \tau_i + e_{ij} + d_{ijk} \quad \text{where...}$$

y_{ijk} is dry weight of plant k in container j of treatment i,

μ denotes the average of the treatment means,

τ_i denotes the effect of treatment i (mean for treatment i minus average of all treatmentmeans, i.e., $\tau_i = \mu_i - \mu$),

e_{ij} is the random effect of container j of treatment i,

d_{ijk} is the random effect of plant k in container j of treatment i,

$e_{ij} \sim N(0, \sigma_e^2)$, $d_{ijk} \sim N(0, \sigma_d^2)$, and all e_{ij} and d_{ijk} are independent.

2. In general we use t to denote the number of treatments, r to denote the number of experimental units per treatment, and n to denote the number of observations per experimental unit. What are t, r, and n in this example?

We can partition the total variation as follows

$$\begin{aligned}\sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{\cdot\cdot\cdot})^2 &= \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2 + \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})^2 + \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij\cdot})^2 \\ &= rn\sum_{i=1}^{t}(\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2 + n\sum_{i=1}^{t}\sum_{j=1}^{r}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot})^2 + \sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij\cdot})^2\end{aligned}$$

where

$$\bar{y}_{\cdot\cdot\cdot} = \frac{1}{trn}\sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{i\cdot\cdot} = \frac{1}{rn}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk} \qquad \bar{y}_{ij\cdot} = \frac{1}{n}\sum_{k=1}^{n} y_{ijk}.$$

We can organize the sums of squares in an ANOVA table:

Source	d.f.	Sum of Squares	Mean Squares	Expect Mean Squares
TRT	$t-1$	$rn\sum_{i=1}^{t}(\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot\cdot\cdot})^2$	$\frac{SST}{t-1}$	$\sigma_d^2+n\sigma_e^2+rn\theta_t^2$
EU(TRT)	$t(r-1)$	$n\sum_{i=1}^{t}\sum_{j=1}^{r}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot})^2$	$\frac{SSE}{t(r-1)}$	$\sigma_d^2+n\sigma_e^2$
OU(EU TRT)	$tr(n-1)$	$\sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij\cdot})^2$	$\frac{SSO}{tr(n-1)}$	σ_d^2
Total	$trn-1$	$\sum_{i=1}^{t}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{\cdot\cdot\cdot})^2$		

where $\theta_t^2=\frac{1}{t-1}\sum_{i=1}^{t}\tau_i^2$.

3. By examining the expected mean squares, can you suggest how to test for differences among treatment means?

4. $\bar{y}_{i\cdot\cdot}$ is the estimated mean for treatment i. It can be show that

$$\mathrm{Var}(\bar{y}_{i\cdot\cdot})=\frac{\sigma_e^2}{r}+\frac{\sigma_d^2}{rn}.$$

   Give the formula for the standard error of $\bar{y}_{i\cdot\cdot}$.

5. It can be show that the variance of a linear combination of treatment means $\sum_{i=1}^{t}c_i\bar{y}_{i\cdot\cdot}$ is

$$\mathrm{Var}\left(\sum_{i=1}^{t}c_i\bar{y}_{i\cdot\cdot}\right)=\left(\frac{\sigma_e^2}{r}+\frac{\sigma_d^2}{rn}\right)\sum_{i=1}^{t}c_i^2.$$

   Give the formula for the standard error of a linear combination of estimated treatment means.