UNEQUAL SAMPLE SIZES

Model:

$$y_{ij} = \mu_i + \epsilon_{ij} \text{ for } i = 1, 2, \ldots, k;\ j = 1, 2, \ldots, n_i$$

(where the sample sizes n_i of each treatment group may vary).

Now let $N = \sum_{i=1}^{k} n_i$ be the total number of observations.

Then

$$\bar{y}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

$$\bar{y}_{\cdot\cdot} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}}{N} = \frac{\sum_{i=1}^{k} n_i \bar{y}_{i\cdot}}{N}$$

where the latter is a weighted average of the $y_{i\cdot}$

ANOVA decomposition

$$y_{ij} - \bar{y}_{\cdot\cdot} = (y_{ij} - \bar{y}_{i\cdot}) + (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})$$

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{\cdot\cdot})^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2 + \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

or

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{\cdot\cdot})^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2 + \sum_{i=1}^{k} n_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$$

which is

$$\text{SST} = \text{SSE} + \text{SSA}$$

with degrees of freedom

SST: $N - 1$

SSE: $\sum_{i=1}^{k} (n_i - 1) = N - k$

SSA: $(N - 1) - (N - k) = k - 1$

$$F = \frac{\frac{\text{SSA}}{k-1}}{\frac{\text{SSE}}{N-k}} = \frac{\text{MSA}}{\text{MSE}}$$

CHECK THAT SUM OF DOUBLE PRODUCTS IS ZERO

$$2\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i\cdot})(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot}) = 2\sum_{i=1}^{k}\left[(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i\cdot})\right] = 0$$

since

$$\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i\cdot}) = \sum_{j=1}^{n_i} y_{ij} - \sum_{j=1}^{n_i}\bar{y}_{i\cdot} = n_i\bar{y}_{i\cdot} - n_i\bar{y}_{i\cdot} = 0$$

ESTIMATION OF σ^2

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i\cdot})^2 = \sum_{i=1}^{k}(n_i-1)s_i^2$$

where

$$s_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i\cdot})^2$$

is usual estimator of σ^2 from the ith treatment group.

Now

$$s_p^2 = \frac{\sum_{i=1}^{k}(n_i-1)s_i^2}{\sum_{i=1}^{k}(n_i-1)} = \frac{\sum_{i=1}^{k}(n_i-1)s_i^2}{N-k} = \frac{SSE}{N-k}$$

is the pooled estimator of σ^2 in the case of different sample sizes.

13.9 Randomized Complete Block Designs

TABLE 13.10 k × b Array for the RCB Design

Treatment	Block: 1	2	···	j	···	b	Total	Mean
1	y_{11}	y_{12}	···	y_{1j}	···	y_{1b}	$T_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21}	y_{22}	···	y_{2j}	···	y_{2b}	$T_{2\cdot}$	$\bar{y}_{2\cdot}$
⋮	⋮	⋮		⋮		⋮	⋮	⋮
i	y_{i1}	y_{i2}	···	y_{ij}	···	y_{ib}	$T_{i\cdot}$	$\bar{y}_{i\cdot}$
⋮	⋮	⋮		⋮		⋮	⋮	
k	y_{k1}	y_{k2}	···	y_{kj}	···	y_{kb}	$T_{k\cdot}$	$\bar{y}_{k\cdot}$
Total	$T_{.1}$	$T_{.2}$	···	$T_{.j}$	···	$T_{.b}$	$T_{..}$	
Mean	$\bar{y}_{.1}$	$\bar{y}_{.2}$	···	$\bar{y}_{.j}$	···	$\bar{y}_{.b}$		$\bar{y}_{..}$

The sum-of-squares identity may be presented symbolically by the equation

$$SST = SSA + SSB + SSE,$$

where

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{..})^2 = \text{total sum of squares},$$

$$SSA = b \sum_{i=1}^{k} (\bar{y}_{i.} - \bar{y}_{..})^2 = \text{treatment sum of squares},$$

$$SSB = k \sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})^2 = \text{block sum of squares},$$

$$SSE = \sum_{i=1}^{k} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 = \text{error sum of squares}.$$

TABLE 13.11 Analysis of Variance for the Randomized Complete Block Design

Source of variation	Sum of squares	Degrees of freedom	Mean square	Computed f
Treatments	SSA	$k-1$	$s_1^2 = \frac{SSA}{k-1}$	$f_1 = \frac{s_1^2}{s^2}$
Blocks	SSB	$b-1$	$s_2^2 = \frac{SSB}{b-1}$	
Error	SSE	$(k-1)(b-1)$	$s^2 = \frac{SSE}{(b-1)(k-1)}$	
Total	SST	$bk-1$		

Example 13.6 Four different machines, M_1, M_2, M_3, and M_4, are being considered for the assembling of a particular product. It is decided that 6 different operators are to be used in a randomized block experiment to compare the machines. The machines are assigned in a random order to each operator. The operation of the machines requires physical dexterity, and it is anticipated that there will be a difference among the operators in the speed with which they operate the machines (Table 13.12). The amount of time (in seconds) were recorded for assembling the product:

TABLE 13.12 Time, in Seconds, to Assemble Product

Machine	Operator: 1	2	3	4	5	6	Total
1	42.5	39.3	39.6	39.9	42.9	43.6	247.8
2	39.8	40.1	40.5	42.3	42.5	43.1	248.3
3	40.2	40.5	41.3	43.4	44.9	45.1	255.4
4	41.3	42.2	43.5	44.2	45.9	42.3	259.4
Total	163.8	162.1	164.9	169.8	176.2	174.1	1010.9

TABLE 13.13 Analysis of Variance for the Data of Table 13.12

Source of variation	Sum of squares	Degrees of freedom	Mean square	Computed f
Machines	15.93	3	5.31	3.34
Operators	42.09	5	8.42	
Error	23.84	15	1.59	
Total	81.86	23		

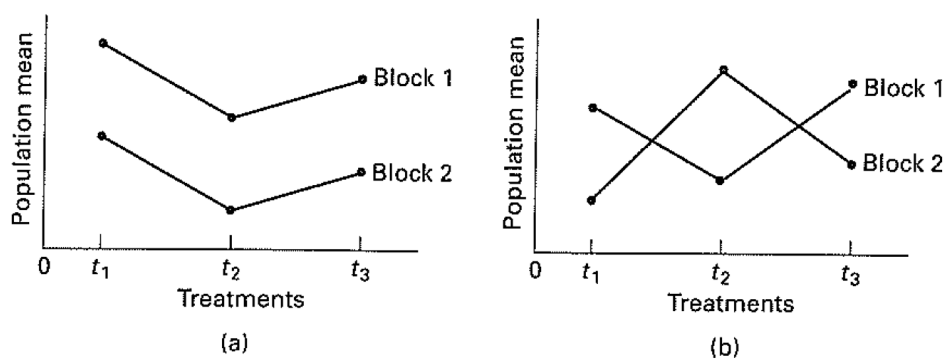

Figure 13.1 Population means for (a) additive results, and (b) interacting effects.

Model with interactions between treatment and block:

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ij},$$

on which we impose the additional restrictions

$$\sum_{i=1}^{k} (\alpha\beta)_{ij} = \sum_{j=1}^{b} (\alpha\beta)_{ij} = 0.$$

One can now readily verify that

$$E\left[\frac{SSE}{(b-1)(k-1)}\right] = \sigma^2 + \frac{\sum_{i=1}^{k}\sum_{j=1}^{b}(\alpha\beta)_{ij}^2}{(b-1)(k-1)}.$$