Randomization restrictions

Stat 566
4-16-15
(1)

Example Paper manufacturing

Factor A: pulp preparation method
3 levels, fixed

Factor B: Cooking temperature
4 levels, fixed

Run 3 replications

There are $3 \times 4 \times 3 = 36$ runs

(2)

To run this as a completely randomized design, the 36 runs should be done in a random order.

But: They can only do 12 runs per day
To run it as a randomized complete block design, the 12 runs on each day must be done in a random order.

But: Each day, they randomly select a preparation method + prepare a batch of pulp

③

They divide the batch into 4 equal parts + cook each part at a different temperature.

They repeat this for the other 2 prep methods, giving 12 runs.

Repeat all of this on day 2 and day 3.

$$\text{Model: } y_{ijkl} = \mu + \underbrace{\tau_i}_{\text{day}} + \underbrace{\beta_j}_{\text{Factor A prep method}} + \underbrace{\gamma_k}_{\text{Factor B temp.}} + \tau\beta_{ij} + \tau\gamma_{ik} + \beta\gamma_{jk} + \tau\beta\gamma_{ijk} + \varepsilon_{(ijk)l}$$

④

	R 3 i	F 3 j	F 4 k	R 1 l	EMS
τ_i	1	3	4	1	$12\sigma^2_\tau + \sigma^2$
β_j	3	0	4	1	$12\sum\beta_j^2/2 + 4\sigma^2_{\tau\beta} + \sigma^2$
γ_k	3	3	0	1	$9\sum\gamma_k^2/3 + 3\sigma^2_{\tau\gamma} + \sigma^2$
$\tau\beta_{ij}$	1	0	4	1	$4\sigma^2_{\tau\beta} + \sigma^2$
$\tau\gamma_{ik}$	1	3	0	1	$3\sigma^2_{\tau\gamma} + \sigma^2$
$\beta\gamma_{jk}$	3	0	0	1	$3\frac{\sum\sum(\beta\gamma)^2_{jk}}{6} + \sigma^2_{\tau\beta\gamma} + \sigma^2$
$\tau\beta\gamma_{ijk}$	1	0	0	1	$\sigma^2_{\tau\beta\gamma} + \sigma^2$
$\varepsilon_{(ijk)l}$	1	1	1	1	σ^2

(5)

df	F tests	denom.
~~2~~	~~Days~~	~~MSE~~
2	Prep	Days x prep
3	Temp	Days x temp
~~4~~	~~Days x prep~~	~~MSE~~
~~6~~	~~Days x temp~~	~~MSE~~
6	Prep x temp	Days x prep x temp
~~12~~	~~Days x prep x temp~~	~~MSE~~
35		

So there are no df for error

(6)

Different layout of the ANOVA table

Split plot design

Source	df	
Days (blocks)	2	
Prep (A)	2	whole plot
Days x Prep	4	
Temp (B)	3	whole plot
Days x temp	6	
AB	6	sub plot
Days x prep x temp	12	
	35	

Revisit the RCBD

$i=1,\ldots,a \quad j=1,\ldots,b \quad k=1$

$$y = \mu + \tau_i + \beta_j + \tau\beta_{ij} + \varepsilon_{(ij)k}$$

	$\overset{a}{\underset{i}{F}}$	$\overset{b}{\underset{j}{R}}$	$\overset{1}{\underset{k}{R}}$	EMS
τ_i	0	b	1	$b\,\Sigma\tau_i^2/(a-1) + \sigma^2_{\tau\beta} + \sigma^2$
β_j	a	1	1	$a\sigma^2_\beta + \sigma^2$
$\tau\beta_{ij}$	0	1	1	$\sigma^2_{\tau\beta} + \sigma^2$
$\varepsilon_{(ij)k}$	1	1	1	σ^2

8

Source	df	denom
A	$a-1$	AB
~~B~~	$b-1$	~~E~~
~~AB~~	$(a-1)(b-1)$	~~E~~

$\cancel{a}-\cancel{1}+\cancel{b}-1+ab-\cancel{a}-\cancel{b}+\cancel{1}$

$= ab-1$, so no df remain for error

Source	df
B	$b-1$
A	$a-1$
AB	$(a-1)(b-1)$

$\text{whole plot}\ \begin{cases} \text{A} \\ \text{AB} \end{cases}$

HW #3 14.2, 14.7

to the original whole plots used for factor A. Figure 14.11 illustrates a situation in which both factors A and B have three levels. Note that the levels of factor A are confounded with the whole plots, and the levels of factor B are confounded with the strips (which can be thought of as a **second** set of whole plots).

A model for the strip-split plot design in Figure 14.11, assuming r replicates, a levels of factor A, and b levels of factor B, is

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + \epsilon_{ijk} \quad \begin{cases} i = 1, 2, \ldots, r \\ j = 1, 2, \ldots, a \\ k = 1, 2, \ldots, b \end{cases}$$

where $(\tau\beta)_{ij}$ and $(\tau\gamma)_{ik}$ are whole-plot errors for factors A and B, respectively, and ϵ_{ijk} is a "subplot" error used to test the AB interaction. Table 14.26 shows an abbreviated analysis of variance assuming A and B are fixed factors and replicates are random. The replicates are sometimes considered as blocks.

14.6 Problems

14.1. A rocket propellant manufacturer is studying the burning rate of propellant from three production processes. Four batches of propellant are randomly selected from the output of each process, and three determinations of burning rate are made on each batch. The results follow. Analyze the data and draw conclusions.

	Process 1				Process 2				Process 3			
Batch	1	2	3	4	1	2	3	4	1	2	3	4
	25	19	15	15	19	23	18	35	14	35	38	25
	30	28	17	16	17	24	21	27	15	21	54	29
	26	20	14	13	14	21	17	25	20	24	50	33

14.2. The surface finish of metal parts made on four machines is being studied. An experiment is conducted in which each machine is run by three different operators and two specimens from each operator are collected and tested. Because of the location of the machines, different operators are used on each machine, and the operators are chosen at random. The data are shown in the following table. Analyze the data and draw conclusions.

	Machine 1			Machine 2			Machine 3			Machine 4		
Operator	1	2	3	1	2	3	1	2	3	1	2	3
	79	94	46	92	85	76	88	53	46	36	40	62
	62	74	57	99	79	68	75	56	57	53	56	47

14.3. A manufacturing engineer is studying the dimensional variability of a particular component that is produced on three machines. Each machine has two spindles, and four components are randomly selected from each spindle. The results follow. Analyze the data, assuming that machines and spindles are fixed factors.

	Machine 1		Machine 2		Machine 3	
Spindle	1	2	1	2	1	2
	12	8	14	12	14	16
	9	9	15	10	10	15
	11	10	13	11	12	15
	12	8	14	13	11	14

14.4. To simplify production scheduling, an industrial engineer is studying the possibility of assigning one time standard to a particular class of jobs, believing that differences between jobs are negligible. To see if this simplification is possible, six jobs are randomly selected. Each job is given to a different group of three operators. Each operator completes the job twice at different times during the week, and the following results are obtained. What are your conclusions about the use of a common time standard for all jobs in this class? What value would you use for the standard?

Job	Operator 1		Operator 2		Operator 3	
1	158.3	159.4	159.2	159.6	158.9	157.8
2	154.6	154.9	157.7	156.8	154.8	156.3
3	162.5	162.6	161.0	158.9	160.5	159.5
4	160.0	158.7	157.5	158.9	161.1	158.5
5	156.3	158.1	158.3	156.9	157.7	156.9
6	163.7	161.0	162.3	160.3	162.6	161.8

14.5. Consider the three-stage nested design shown in Figure 14.5 to investigate alloy hardness. Using the data that follow, analyze the design, assuming that alloy chemistry and heats are fixed factors and ingots are random. Use the restricted form of the mixed model.

	Alloy Chemistry 1					
Heats	1		2		3	
Ingots	1	2	1	2	1	2
	40	27	95	69	65	78
	63	30	67	47	54	45
	Alloy Chemistry 2					
Heats	1		2		3	
Ingots	1	2	1	2	1	2
	22	23	83	75	61	35
	10	39	62	64	77	42

14.6. Reanalyze the experiment in Problem 14.5 using the unrestricted form of the mixed model. Comment on any differences you observe between the restricted and the unrestricted model results. You may use a computer software package.

14.7. Derive the expected mean squares for a balanced three-stage nested design, assuming that A is fixed and that B and C are random. Obtain formulas for estimating the variance components. Assume the restricted form of the mixed model.

14.8. Repeat Problem 14.7 assuming the unrestricted form of the mixed model. You may use a computer software package to do this. Comment on any differences between the restricted and unrestricted model analysis and conclusions.

14.9. Derive the expected mean squares for a balanced three-stage nested design if all three factors are random. Obtain formulas for estimating the variance components.

14.10. Verify the expected mean squares given in Table 14.1.

14.11. ***Unbalanced nested designs.*** Consider an unbalanced two-stage nested design with b_j levels of B under the ith level of A and n_{ij} replicates in the ijth cell.

(**a**) Write down the least squares normal equations for this situation. Solve the normal equations.

(**b**) Construct the analysis of variance table for the unbalanced two-stage nested design.

(**c**) Analyze the following data, using the results in part (b).

Factor A	1		2		
Factor B	1	2	1	2	3
	6	−3	5	2	1
	4	1	7	4	0
	8		9	3	−3
			6		

14.12. ***Variance components in the unbalanced two-stage nested design.*** Consider the model

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{k(ij)} \quad \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b_i \\ k = 1, 2, \ldots, n_{ij} \end{cases}$$

where A and B are random factors. Show that

$$\begin{aligned} E(MS_A) &= \sigma^2 + c_1\sigma_\beta^2 + c_2\sigma_\tau^2 \\ E(MS_{B(A)}) &= \sigma^2 + c_0\sigma_\beta^2 \\ E(MS_E) &= \sigma^2 \end{aligned}$$

where

$$c_0 = \frac{N - \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i} n_{ij}^2/n_{i.}\right)}{b - a}$$

$$c_1 = \frac{\sum_{i=1}^{a}\left(\sum_{j=1}^{b_i} n_{ij}^2/n_{i.}\right) - \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}^2/N}{a - 1}$$

$$c_2 = \frac{N - \dfrac{\sum_{i=1}^{a} n_{i.}^2}{N}}{a - 1}$$

14.13. A process engineer is testing the yield of a product manufactured on three machines. Each machine can be operated at two power settings. Furthermore, a machine has three stations on which the product is formed. An experiment is conducted in which each machine is tested at both power settings, and three observations on yield are taken from each station. The runs are made in random order, and the results are shown in Table P14.1. Analyze this experiment, assuming that all three factors are fixed.

■ **TABLE P14.1**
Yield Experiment in Problem 14.13

	Machine 1			Machine 2		
Station	**1**	**2**	**3**	**1**	**2**	**3**
Power	34.1	33.7	36.2	31.1	33.1	32.8
setting 1	30.3	34.9	36.8	33.5	34.7	35.1
	31.6	35.0	37.1	34.0	33.9	34.3
Power	24.3	28.1	25.7	24.1	24.1	26.0
setting 2	26.3	29.3	26.1	25.0	25.1	27.1
	27.1	28.6	24.9	26.3	27.9	23.9
	Machine 3					
Station	**1**	**2**	**3**			
Power	32.9	33.8	33.6			
setting 1	33.0	33.4	32.8			
	33.1	32.8	31.7			
Power	24.2	23.2	24.7			
setting 2	26.1	27.4	22.0			
	25.3	28.0	24.8			