Chapter 7

Block Designs

Definition 7.1. A **block** is a group of k similar or homogenous units. In a **block design**, each unit in a block is randomly assigned to one of k treatments. The meaning of "similar" is that the units are likely to have similar values of the response when given identical treatments.

In agriculture, adjacent plots of land are often used as blocks since adjacent plots tend to give similar yields. Litter mates, siblings, twins, time periods (eg different days) and batches of material are often used as blocks.

Following Cobb (1998, p. 247), there are 3 ways to get blocks. i) Sort units into groups (blocks) of k similar units. ii) Divide large chunks of material (blocks) into smaller pieces (units). iii) Reuse material or subjects (blocks) several times. Then the time slots are the units.

Example 7.1. For i), to study the effects of k different medicines, sort people into groups of size k according to similar age and weight. For ii) suppose there are b plots of land. Divide each plot into k subplots. Then each plot is a block and the subplots are units. For iii), give the k different treatments to each person over k months. Then each person has a block of time slots and the ith month = time slot is the unit.

Suppose there are b blocks and $n = kb$. The one way Anova design randomly assigns b of the units to each of the k treatments. Blocking places a constraint on the randomization, since within each block of units, exactly one unit is randomly assigned to each of the k treatments.

Hence a one way Anova design would use the R command `sample(n)` and the first b units would be assigned to treatment 1, the second b units to

treatment 2, ... and the last b units would be assigned to treatment k.

For the completely randomized block designs described in the following section, the command `sample(k)` is done b times: once for each block. The ith command is for the units of the ith block. If $k = 5$ and the `sample(5)` command yields 2 5 3 1 4, then the 2nd unit in the ith block is assigned to treatment 1, the 5th unit to treatment 2, the 3rd unit to treatment 3, the 1st unit to treatment 4 and the 4th unit to treatment 5.

Remark 7.1. Blocking and randomization often makes the iid error assumption hold to a useful approximation.

For example, if grain is planted in n plots of land, yields tend to be similar (correlated) in adjacent identically treated plots, but the yields from all of the plots vary greatly, and the errors are not iid. If there are 4 treatments and blocks of 4 adjacent plots, then randomized blocking makes the errors approximately iid.

7.1   One Way Block Designs

Definition 7.2. For the **one way block design** or **completely randomized block design (CRBD)**, there is a factor A with k levels and there are b blocks. The CRBD model is

$$Y_{ij} = \mu_{ij} + e_{ij} = \mu + \tau_i + \beta_j + e_{ij}$$

where τ_i is the ith treatment effect and $\sum_{i=1}^{k} \tau_i = 0$, β_j is the jth block effect and $\sum_{j=1}^{b} \beta_j = 0$. The indices $i = 1, ..., k$ and $j = 1, ..., b$. Then

$$\mu_i \equiv \frac{\mu_{io}}{b} = \frac{1}{b} \sum_{j=1}^{b} (\mu + \tau_i + \beta_j) = \mu + \tau_i.$$

So the μ_i are all equal if the τ_i are all equal. The errors e_{ij} are iid with 0 mean and constant variance σ^2.

Notice that the CRBD model is additive: there is no block treatment interaction. The ANOVA table for the CRBD is like the ANOVA table for a two way Anova main effects model. Shown below is a CRBD ANOVA table in symbols. Sometimes "Treatment" is replaced by "Factor" or "Model." Sometimes "Blocks" is replaced by the name of the blocking variable. Sometimes "Error" is replaced by "Residual."

Source	df	SS	MS	F	p-value
Blocks	b-1	SSB	MSB	"F_{block}"	"p_{block}"
Treatment	k-1	SSTR	MSTR	F_0=MSTR/MSE	pval for Ho
Error	$(k-1)(b-1)$	SSE	MSE		

Be able to perform the 4 step completely randomized block design ANOVA F test of hypotheses. This test is similar to the fixed effects one way Anova F test.

i) Ho: $\mu_1 = \mu_2 = \cdots = \mu_k$ and H_A: not Ho.

ii) Fo = MSTR/MSE is usually given by output.

iii) The p-value = $P(F_{k-1,(k-1)(b-1)} > F_o)$ is usually given by output.

iv) If the p–value $< \delta$, reject Ho and conclude that the mean response depends on the level of the factor. Otherwise fail to reject Ho and conclude that the mean response does not depend on the level of the factor. Give a nontechnical sentence.

Rule of thumb 7.1. If $p_{block} \geq 0.1$, then blocking was not useful. If $0.05 \leq p_{block} < 0.1$, then the usefulness was borderline. If $p_{block} < 0.05$, then blocking was useful.

Remark 7.2. The response, residual and transformation plots are almost used in the same way as for the one and two way Anova model, but all of the dot plots have sample size $m = 1$. Look for the plotted points falling in roughly evenly populated bands about the identity line and $r = 0$ line.

Definition 7.3. The **block response scatterplot** plots blocks versus the response. The plot will have b dot plots of size k with a symbol corresponding to the treatment. Dot plots with clearly different means suggest that blocking was useful. A symbol pattern within the blocks suggests that the response depends on the factor.

Definition 7.4. Graphical Anova for the CRBD model uses the residuals as a reference set instead of a F distribution. The scaled treatment deviations $\sqrt{b-1}(\overline{Y}_{i0} - \overline{Y}_{00})$ have about the same variability as the residuals if Ho is true. The scaled block deviations $\sqrt{k-1}(\overline{Y}_{0j} - \overline{Y}_{00})$ also have about the same variability as the residuals if blocking is ineffective. A dot plot of the scaled block deviations is placed above the dot plot of the scaled treatment deviations which is placed above the dot plot of the residuals. For small $n \leq 40$, suppose the distance between two scaled deviations (A and B,

250

say) is greater than the range of the residuals = $\max(r_{ij}) - \min(r_{ij})$. Then declare μ_A and μ_B to be significantly different. If the distance is less than the range, do not declare μ_A and μ_B to be significantly different. Scaled deviations that lie outside the range of the residuals are significant: the corresponding treatment means are significantly different from the overall mean.

For $n \geq 100$, let $r_{(1)} \leq r_{(2)} \leq \cdots \leq r_{(n)}$ be the order statistics of the residuals. Then instead of the range, use $r_{(\lceil 0.975n \rceil)} - r_{(\lceil 0.025n \rceil)}$ as the distance where $\lceil x \rceil$ is the smallest integer $\geq x$, eg $\lceil 7.7 \rceil = 8$. So effects outside of the interval $(r_{(\lceil 0.025n \rceil)}, r_{(\lceil 0.975n \rceil)})$ are significant. See Box, Hunter and Hunter (2005, p. 150-151).

```
Output for Example 7.2.
          Df   Sum Sq  Mean Sq F value    Pr(>F)
block      3 79308210 26436070  54.310 4.348e-06
treatment  3  1917416   639139   1.313    0.3292
Residuals  9  4380871   486763

> ganova2(x,block,y)  scaled block deviations
                   -3790.377  4720.488  2881.483 -3811.594
block                 1          2         3          4

                      scaled treatment deviations
                   -266.086  -833.766   733.307  366.545
Treatments          "A"        "B"        "C"       "D"
```

Example 7.2. Ledolter and Swersey (2007, p. 60) give completely randomized block design data. The block variable = market had 4 levels (1 Binghamton, 2 Rockford, 3 Albuquerque, 4 Chattanooga) while the treatment factor had 4 levels (A no advertising, B \$6 million, C \$12 million, D \$18 million advertising dollars in 1973). The response variable was average cheese sales (in pounds per store) sold in a 3 month period.

a) From the graphical Anova in Figure 7.1, were the blocks useful?

b) Perform an appropriate 4 step test for whether advertising helped cheese sales.

Solution: a) In Figure 7.1, the top dot plot is for the scaled block deviations. The leftmost dot corresponds to blocks 4 and 1, the middle dot to block 3 and the rightmost dot to block 1 (see output from the regpack
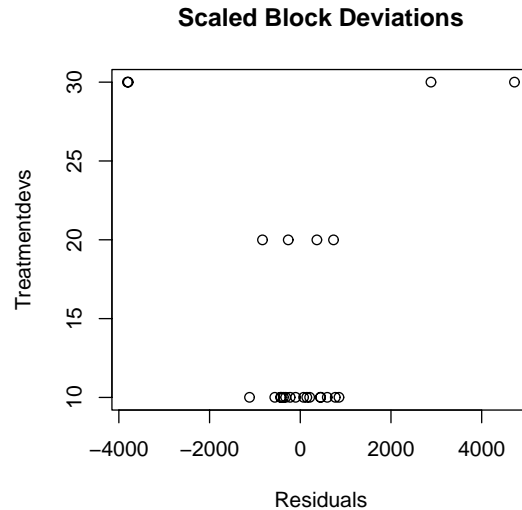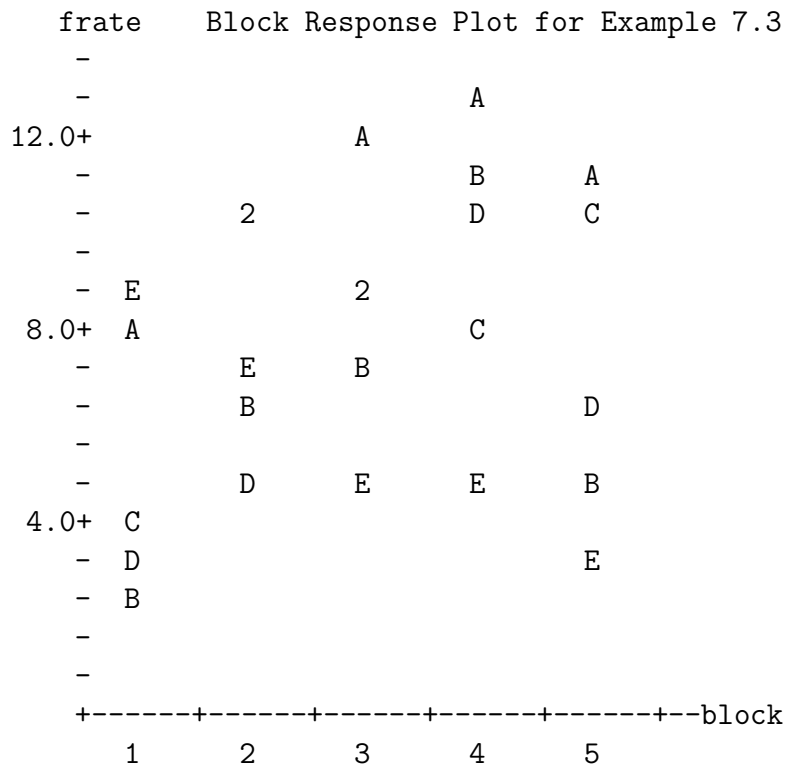
Scaled Block Deviations

Figure 7.1: Graphical Anova for a One Way Block Design

function `ganova2`). Yes, the blocks were useful since some (actually all) of the dots corresponding to the scaled block deviations fall outside the range of the residuals. This result also agrees with $p_{block} = 4.348\text{e--}06 < 0.05$.

b) i) Ho: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_A: not Ho
ii) Fo = 1.313
iii) pval = 0.3292
iv) Fail to reject Ho, the mean sales does not depend on advertising level.

In Figure 7.1, the middle dot plot is for the scaled treatment deviations. From left to right, these correspond to B, A, D and C since the output shows that the deviation corresponding to C is the largest with value 733.3. Since the four scaled treatment deviations all lie within the range of the residuals, the four treatments again do not appear to be significant.

Example 7.3. Snedecor and Cochran (1967, p. 300) give a data set with 5 types of soybean seed. The response frate = number of seeds out of 100 that failed to germinate. Five blocks were used. On the following page is a block response plot where A, B, C, D and E refer to seed type. The 2 in the second block indicates that A and C both had values 10. Which type of seed

252

```
   frate     Block Response Plot for Example 7.3
     -
     -                              A
12.0+                     A
     -                              B        A
     -           2                  D        C
     -
     -  E                 2
 8.0+  A                           C
     -           E        B
     -           B                           D
     -
     -           D        E        E        B
 4.0+  C
     -  D                                    E
     -  B
     -
     -
     +------+------+------+------+------+--block
         1      2      3      4      5
```

has the highest germination failure rate?
a) A   b) B   c) C   d) D   e) E

Solution: a) A since A is on the top for blocks 2–5 and second for block
1. (The Bs and Es suggest that there may be a block treatment interaction.)

7.2   Blocking with the K Way Anova Design

Blocking is used to reduce the MSE so that inference such as tests and
confidence intervals are more precise. On the following page is a partial
Anova table for a k way Anova design with one block where the degrees of
freedom are left blank. For A, use $H_0 : \mu_{10\cdots0} = \mu_{20\cdots0}$. The other main effects
have similar null hypotheses. For interaction, use H_0 : no interaction.

These models get complex rapidly as k and the number of levels l_i in-
crease. As k increases, there are a large number of models to consider. For
experiments, usually the 3 way and higher order interactions are not signif-
icant. Hence a full model that includes the blocks, all k main effects and

all $\binom{k}{2}$ two way interactions is a useful starting point for response, residual and transformation plots. The higher order interactions can be treated as potential terms and checked for significance. As a rule of thumb, significant interactions tend to involve significant main effects.

Source	df	SS MS	F	p-value
block		SSblock MSblock	"F_{block}"	"p_{block}"
k main effects		eg SSA = MSA	F_A	p_A
$\binom{k}{2}$ 2 way interactions		eg SSAB = MSAB	F_{AB}	p_{AB}
$\binom{k}{3}$ 3 way interactions		eg SSABC = MSABC	F_{ABC}	p_{ABC}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\binom{k}{k-1}$ $k-1$ way interactions				
the k way interaction		SSA$\cdots$L = MSA$\cdots$L	$F_{A\cdots L}$	$p_{A\cdots L}$
Error		SSE MSE		

The following example has one block and 3 factors. Hence there are 3 two way interactions and 1 three way interaction.

```
Source  df  SS      MS       F      pvalue
block    1  0.1334 0.1334  4.85    0.0379
L        3  0.0427 0.0142  0.5164 0.6751
M        2  0.0526 0.0263  0.9564 0.3990
P        1  0.5355 0.5355  19.47   0.0002
LM       6  0.2543 0.0424  1.54    0.2099
LP       3  0.2399 0.0800  2.91    0.0562
MP       2  0.0821 0.0410  1.49    0.2463
LMP      6  0.0685 0.0114  0.4145 0.8617
error   23  0.6319 0.0275
```

Example 7.4. Snedecor and Cochran (1967, p. 361-364) describe a block design (2 levels) with three factors: food supplements Lysine (4 levels), Methionine (3 levels) and Protein (2 levels). Male pigs were fed the supplements in a $4 \times 3 \times 2$ factorial arrangement and the response was average daily weight gain. The ANOVA table is shown above. The model could be described as $Y_{ijkl} = \mu_{ijkl} + e_{ijkl}$ for $i = 1, 2, 3, 4$; $j = 1, 2, 3$; $k = 1, 2$ and $l = 1, 2$ where i, j, k are for L,M,P and l is for block. Note that μ_{i000} is the mean corresponding to the ith level of L.

a) There were 24 pigs in each block. How were they assigned to the $24 = 4 \times 3 \times 2$ runs (a run is a L,M,P combination forming a pig diet)?

b) Was blocking useful?

c) Perform a 4 step test for the significant main effect.

d) Which, if any, of the interactions were significant?

Solution: a) Randomly.

b) Yes, $0.0379 < 0.05$.

c) H_0 $\mu_{0010} = \mu_{0020}$ H_A not H_0

$F_P = 19.47$

pval $= 0.0002$

Reject H_0, the mean weight gain depends on the protein.

d) None.

Remark 7.3. There are 3 basic principles of DOE. Randomization, factorial crossing and blocking can be used to create many DOE models.

i) Use **randomization** to assign units to treatments.

ii) Use **factorial crossing** to compare the effects of 2 or more factors in the same experiment: if $A_1, A_2, ..., A_k$ are the k factors where the ith factor A_i has l_i levels, then there are $(l_1)(l_2) \cdots (l_k)$ treatments where a treatment has one level from each factor.

iii) Use **blocking** to increase precision. Divide units into blocks of similar homogeneous units where "similar" implies that the units are likely to have similar values of the response if given the same treatment. Within each block, randomly assign units to treatments.

7.3 Latin Square Designs

Latin square designs have a lot of structure. The design contains a row block factor, a column block factor and a treatment factor, each with a levels. The two blocking factors and the treatment factor are crossed, but it is assumed that there is no interaction. A capital letter is used for each of the a treatment levels. So $a = 3$ uses A, B, C while $a = 4$ uses A, B, C, D.

Definition 7.5. In an $a \times a$ *Latin square*, each letter appears exactly once in each row and in each column. A *standard Latin square* has letters written in alphabetical order in the first row and in the first column.

Five Latin squares are shown below. The first, third and fifth are standard. If $a = 5$, there are 56 standard Latin squares.

```
A B C     A B C     A B C D     A B C D E     A B C D E
B C A     C A B     B A D C     E A B C D     B A E C D
C A B     B C A     C D A B     D E A B C     C D A E B
                    D C B A     C D E A B     D E B A C
                                B C D E A     E C D B A
```

Definition 7.6. The model for the **Latin square design** is

$$Y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + e_{ijk}$$

where τ_i is the ith treatment effect, β_j is the jth row block effect, γ_k is the kth column block effect with i, j and $k = 1, ..., a$. The errors e_{ijk} are iid with 0 mean and constant variance σ^2. The ith treatment mean $\mu_i = \mu + \tau_i$.

Shown below is an ANOVA table for the Latin square model given in symbols. Sometimes "Error" is replaced by "Residual," or "Within Groups." Sometimes rblocks and cblocks are replaced by the names of the blocking factors. Sometimes "p-value" is replaced by "P", "$Pr(> F)$" or "PR > F."

Source	df	SS	MS	F	p-value
rblocks	$a - 1$	"SSRB"	"MSRB"	"F_{row}"	"p_{row}"
cblocks	$a - 1$	"SSCB"	"MSCB"	"F_{col}"	"p_{col}"
treatments	$a - 1$	SSTR	MSTR	F_o=MSTR/MSE	pval
Error	$(a - 1)(a - 2)$	SSE	MSE		

Rule of thumb 7.2. Let p_{block} be p_{row} or p_{col}. If $p_{block} \geq 0.1$, then blocking was not useful. If $0.05 \leq p_{block} < 0.1$, then the usefulness was borderline. If $p_{block} < 0.05$, then blocking was useful.

Be able to perform the 4 step Anova F test for the Latin square design. This test is similar to the fixed effects one way Anova F test.
i) Ho: $\mu_1 = \mu_2 = \cdots = \mu_a$ and H_A: not Ho.
ii) Fo = MSTR/MSE is usually given by output.
iii) The p-value = $P(F_{a-1,(a-1)(a-2)} > F_o)$ is usually given by output.
iv) If the p–value $< \delta$, reject Ho and conclude that the mean response depends on the level of the factor. Otherwise fail to reject Ho and conclude that the

mean response does not depend on the level of the factor. Give a nontechnical sentence. Use $\delta = 0.05$ if δ is not given.

Remark 7.4. The response, residual and transformation plots are almost used in the same way as for the one and two way Anova models, but all of the dot plots have sample size $m = 1$. Look for the plotted points falling in roughly evenly populated bands about the identity line and $r = 0$ line.

Source	df	SS	MS	F	P
rblocks	3	774.335	258.1117	2.53	0.1533
cblocks	3	133.425	44.4750	0.44	0.7349
fertilizer	3	1489.400	496.4667	4.87	0.0476
error	6	611.100	101.8500		

Example 7.5. Dunn and Clark (1974, p. 129) examine a study of four fertilizers on yields of wheat. The row blocks were 4 types of wheat. The column blocks were 4 plots of land. Each plot was divided into 4 subplots and a Latin square design was used. (Ignore the fact that the data had an outlier.)

a) Were the row blocks useful? Explain briefly.

b) Were the column blocks useful? Explain briefly.

c) Do an appropriate 4 step test.

Solution:
a) No, $p_{row} = 0.1533 > 0.1$.
b) No, $p_{col} = 0.7349 > 0.1$.
c) i) H_0 $\mu_1 = \mu_2 = \mu_3 = \mu_4$ H_A not H_0
ii) $F_0 = 4.87$
iii) pval $= 0.0476$
iv) Reject H_0. The mean yield depends on the fertilizer.

Remark 7.5. The Latin square model is additive, but the model is often incorrectly used to study "nuisance factors" that can interact. Factorial or fractional factorial designs should be used when interaction is possible.

Remark 7.6. The randomization is done in 3 steps. Draw 3 random permutations of 1, ..., a. Use the 1st permutation to randomly assign row block levels to the numbers 1, ..., a. Use the 2nd permutation to randomly assign column block levels to the numbers 1, ..., a. Use the 3rd permutation

to randomly assign treatment levels to the 1st a letters (A, B, C and D if $a = 4$).

Example 7.6. In the social sciences, often a blocking factor is *time*: the levels are a time slots. Following Cobb (1998, p. 254), a Latin square design was used to study the response $Y = $ *blood sugar level*, where the row blocks were 4 *rabbits*, the column blocks were 4 *time slots*, and the treatments were 4 levels of *insulin*. Label the rabbits as I, II, III and IV; the dates as 1, 2, 3, 4; and the 4 insulin levels $i_1 < i_2 < i_3 < i_4$ as 1, 2, 3, 4. Suppose the random permutation for the rabbits was 3, 1, 4, 2; the permutation for the dates 1, 4, 3, 2; and the permutation for the insulin levels was 2, 3, 4, 1. Then i_2 is treatment A, i_3 is treatment B, i_4 is treatment C and i_1 is treatment D. Then the data are as shown below on the left. The data is rearranged for presentation on the right.

```
            raw data                     presentation data
              date                             date
rabbit  4/23 4/27 4/26 4/25     rabbit 4/23 4/25 4/26 4/27
III      57A  45B  60C  26D      I       24B  46C  34D  48A
I        24B  48A  34D  46C      II      33D  58A  57B  60C
IV       46C  47D  61A  34B      III     57A  26D  60C  45B
II       33D  60C  57B  58A      IV      46C  34B  61A  47D
```

Example 7.7. Following Cobb (1998, p. 255), suppose there is a rectangular plot divided into 5 rows and 5 columns to form 25 subplots. There are 5 treatments which are 5 varieties of a plant, labelled 1, 2, 3, 4, 5; and the response Y is yield. Adjacent subplots tend to give similar yields under identical treatments, so the 5 rows form the row blocks and the 5 columns form the column blocks. To perform randomization, three random permutations are drawn. Shown on the following page are 3 Latin squares. The one on the left is an unrandomized Latin square.

Suppose 2, 4, 3, 5, 1 is the permutation drawn for rows. The middle Latin square with randomized rows has 1st row which is the 2nd row from the original unrandomized Latin square. The middle square has 2nd row that is the 4th row from the original, the 3rd row is the 3rd row from the original, the 4th row is the 5th row from the original, and the 5th row is the 1st row from the original.

258

```
unrandomized    randomized rows randomized Latin square
rows columns    rows columns     rows columns
  1 2 3 4 5       1 2 3 4 5         1 4 2 5 3
1 A B C D E     2 B C D E A       2 B E C A D
2 B C D E A     4 D E A B C       4 D B E C A
3 C D E A B     3 C D E A B       3 C A D B E
4 D E A B C     5 E A B C D       5 E C A D B
5 E A B C D     1 A B C D E       1 A D B E C
```

Suppose 1, 4, 2, 5, 3 is the permutation drawn for columns. Then the randomized Latin square on the right has 1st column which is the 1st column from the middle square, the 2nd column is the 4th column from the middle square, the 3rd column is the 2nd column from the middle square, the 4th column is the 5th column from the middle square, and the 5th column is the 3rd column from the middle square.

Suppose 3, 2, 5, 4, 1 is the permutation drawn for variety. Then variety 3 is treatment A, 2 is B, 5 is C, 4 is D and variety 1 is E. Now sow each subplot with the variety given by the randomized Latin square on the right. Hence the northwest corner gets $B = $ variety 2, the northeast corner gets $D = $ variety 4, the southwest corner gets $A = $ variety 3, the southeast corner gets $C = $ variety 5, et cetera.

7.4   Summary

1) A block is a group of similar (homogeneous) units in that the units in a block are expected to give similar values of the response if given the same treatment.

2) In agriculture, adjacent plots of land are often used as blocks since adjacent plots tend to give similar yields. Litter mates, siblings, twins, time periods (eg different days) and batches of material are often used as blocks.

3) The *completely randomized block design* with k treatments and b blocks of k units uses randomization within each block to assign exactly one of the block's k units to each of the k treatments. This design is a generalization of the matched pairs procedure.

4) The *Anova F test for the completely randomized block design* with k treatments and b blocks is nearly the same as the fixed effects one way Anova F test.

i) Ho: $\mu_1 = \mu_2 = \cdots = \mu_k$ and H_A: not Ho.

ii) Fo = MSTR/MSE is usually given by output.

iii) The p-value = $P(F_{k-1,(k-1)(b-1)} > F_o)$ is usually given by output.

iv) If the p–value $< \delta$, reject Ho and conclude that the mean response depends on the level of the factor. Otherwise fail to reject Ho and conclude that the mean response does not depend on the level of the factor. Give a nontechnical sentence.

5) Shown below is an ANOVA table for the completely randomized block design.

Source	df	SS	MS	F	p-value
Blocks	b-1	SSB	MSB	"F_{block}"	"p_{block}"
Treatment	k-1	SSTR	MSTR	F_0=MSTR/MSE	pval for Ho
Error	$(k-1)(b-1)$	SSE	MSE		

6) Rule of thumb: If $p_{block} \geq 0.1$, then blocking was not useful. If $0.05 \leq p_{block} < 0.1$, then the usefulness was borderline. If $p_{block} < 0.05$, then blocking was useful.

7) The response, residual and transformation plots are used almost in the same way as for the one and two way Anova model, but all of the dot plots have sample size $m = 1$. Look for the plotted points falling in roughly evenly populated bands about the identity line and $r = 0$ line.

8) The **block response scatterplot** plots blocks versus the response. The plot will have b dot plots of size k with a symbol corresponding to the treatment. Dot plots with clearly different means suggest that blocking was useful. A symbol pattern within the blocks suggests that the response depends on the factor.

9) Shown is an ANOVA table for the Latin square model given in symbols. Sometimes "Error" is replaced by "Residual," or "Within Groups." Sometimes rblocks and cblocks are replaced by the blocking factor name. Sometimes "p-value" is replaced by "P", "$Pr(> F)$" or "PR > F."

Source	df	SS	MS	F	p-value
rblocks	$a-1$	"SSRB"	"MSRB"	"F_{row}"	"p_{row}"
cblocks	$a-1$	"SSCB"	"MSCB"	"F_{col}"	"p_{col}"
treatments	$a-1$	SSTR	MSTR	F_o=MSTR/MSE	pval
Error	$(a-1)(a-2)$	SSE	MSE		

10) Let p_{block} be p_{row} or p_{col}. Rule of thumb: If $p_{block} \geq 0.1$, then blocking was not useful. If $0.05 \leq p_{block} < 0.1$, then the usefulness was borderline. If $p_{block} < 0.05$, then blocking was useful.

11) The *Anova F test for the Latin square design* with a treatments is nearly the same as the fixed effects one way Anova F test.
i) Ho: $\mu_1 = \mu_2 = \cdots = \mu_a$ and H_A: not Ho.
ii) Fo = MSTR/MSE is usually given by output.
iii) The p-value = $P(F_{a-1,(a-1)(a-2)} > F_o)$ is usually given by output.
iv) If the p–value $< \delta$, reject Ho and conclude that the mean response depends on the level of the factor. Otherwise fail to reject Ho and conclude that the mean response does not depend on the level of the factor. Give a nontechnical sentence.

12) The response, residual and transformation plots are almost used in the same way as for the one and two way Anova models, but all of the dot plots have sample size $m = 1$. Look for the plotted points falling in roughly evenly populated bands about the identity line and $r = 0$ line.

13) The randomization is done in 3 steps. Draw 3 random permutations of 1, ..., a. Use the 1st permutation to randomly assign row block levels to the numbers 1, ..., a. Use the 2nd permutation to randomly assign column block levels to the numbers 1, ..., k. Use the 3rd permutation to randomly assign treatment levels to the 1st a letters (A, B, C and D if $a = 4$).

14) Graphical Anova for the **completely randomized block** design makes a dotplot of the scaled block deviations $\tilde{\beta}_j = \sqrt{k-1}\hat{\beta}_j = \sqrt{k-1}(\overline{y}_{0j0} - \overline{y}_{000})$ on top, a dotplot of scaled treatment deviations (effects) $\tilde{\alpha}_i = \sqrt{b-1}\hat{\alpha}_i = \sqrt{b-1}(\overline{y}_{i00} - \overline{y}_{000})$ in the middle and a dotplot of the residuals on the bottom. Here k is the number of treatments and b is the number of blocks.

15) Graphical Anova uses the residuals as a reference distribution. Suppose the dotplot of the residuals looks good. Rules of thumb: i) An effect is marginally significant if its scaled deviation is as big as the biggest residual or as negative as the most negative residual. ii) An effect is significant if it is well beyond the minimum or maximum residual. iii) Blocking was effective if at least one scaled block deviation is beyond the range of the residuals. iv) The treatments are different if at least one scaled treatment effect is beyond the range of the residuals. (These rules depend on the number of residuals n. If n is very small, say 8, then the scaled effect should be well beyond the range of the residuals to be significant. If the n is 40, the value of the minimum residual and the value of the maximum residual correspond to a

$1/40 + 1/40 = 1/20 = 0.05$ critical value for significance.)

7.5 Complements

Box, Hunter and Hunter (2005, p. 150-156) explain Graphical Anova for the CRBD and why randomization combined with blocking often makes the iid error assumption hold to a reasonable approximation.

The R package `granova` may be useful for graphical Anova. It is available from (http://streaming.stat.iastate.edu/CRAN/) and authored by R.M. Pruzek and J.E. Helmreich. Also see Hoaglin, Mosteller, and Tukey (1991).

Matched pairs tests are a special case of CRBD with $k = 2$.

A **randomization test** has H_0: *the different treatments have no effect.* This null hypothesis is also true if within each block, all k pdfs are from the same location family. Let $j = 1, ..., b$ index the b blocks. There are b pdfs, one for each block, that come from the same location family but possibly different location parameters: $f_Z(y - \mu_{0j})$. Let A be the treatment factor with k levels a_i. Then $Y_{ij}|(A = a_i) \sim f_Z(y - \mu_{0j})$ where j is fixed and $i = 1, ..., k$. Thus the levels a_i have no effect on the response, and the Y_{ij} are iid within each block if H_0 holds. Note that there are $k!$ ways to assign $Y_{1j}, ... Y_{kj}$ to the k treatments within each block. An impractical randomization test uses all $M = [k!]^b$ ways of assigning responses to treatments. Let F_0 be the usual CRBD F statistic. The F statistic is computed for each of the M permutations and H_0 is rejected if the proportion of the M F statistics that are larger than F_0 is less than δ. The distribution of the M F statistics is approximately $F_{k-1, (k-1)(b-1)}$ for large n under H_0. The randomization test and the usual CBRD F test also have the same power, asymptotically. See Hoeffding (1952) and Robinson (1973). These results suggest that the usual CRBD F test is semiparametric: the pvalue is approximately correct if n is large and if all k pdfs $Y_{ij}|(A = a_i) \sim f_Z(y - \mu_{0j})$ are the same for each block where j is fixed and $i = 1, ..., k$. If H_0 does not hold, then there are kb pdfs $Y_{ij}|(A = a_i) \sim f_Z(y - \mu_{ij})$ from the same location family. Hence the location parameter depends on both the block and treatment.

Olive (2009b) shows that practical randomization tests that use a random sample of $\max(1000, [n \log(n)])$ randomizations have level and power similar to the tests that use all M possible randomizations. Here each "randomization" uses b randomly drawn permutations of $1, ..., k$.

Hunter (1989) discusses some problems with the Latin square design.

7.6 Problems

Problems with an asterisk * are especially important.

```
Output for 7.1.
source      Df Sum Sq Mean Sq F value  Pr(>F)
block        4  49.84   12.46  2.3031 0.10320
seed         4  83.84   20.96  3.8743 0.02189
Residuals   16  86.56    5.41
```

7.1. Snedecor and Cochran (1967, p. 300) give a data set with 5 types of soybean seed. The response frate = number of seeds out of 100 that failed to germinate. Five blocks were used. Assume the appropriate model can be used (although this assumption may not be valid due to a possible interaction between the block and the treatment).

  a) Did blocking help? Explain briefly.

  b) Perform the appropriate 4 step test using the output above.

```
Output for 7.2.
Source      df      SS        MS      F      P
blocks       3    197.004   65.668   9.12   0.001
treatment    5    201.316   40.263   5.59   0.004
error       15    108.008    7.201
```

7.2. Current nitrogen fertilization recommendations for wheat include applications of specified amounts at specified stages of plant growth. The treatment consisted of six different nitrogen application and rate schedules. The wheat was planted in an irrigated field that had a water gradient in one direction as a result of the irrigation. The field plots were grouped into four blocks, each consisting of six plots, such that each block occurred in the same part of the water gradient. The response was the observed nitrogen content from a sample of wheat stems from each plot. The experimental units were the 24 plots. Data is from Kuehl (1994, p. 263).

  a) Did blocking help? Explain briefly.

  b) Perform the appropriate 4 step test using the output above.

7.3. An experimenter wants to test 4 types of an altimeter. There are eight helicopter pilots available for hire with from 500 to 3000 flight hours of experience. The response variable is the altimeter reading error. Perform the appropriate 4 step test using the output below. Data is from Kirk (1982, p. 244).

```
Output for Problem 7.3
 Source       df    SS      MS       F      P
treatment    3    194.50   64.833   47.78  0.000
blocks       7     12.50    1.786    1.32
error        21    28.50    1.357
```

One way randomized block designs in SAS, Minitab and R

7.4. This problem is for a one way block design and uses data from Box, Hunter and Hunter (2005, p. 146).

a) Copy and paste the SAS program for this problem from (www.math.siu.edu/olive/reghw.txt). Print out the output but only turn in the Anova table, residual plot and response plot.

b) Do the plots look ok?

c) Copy the SAS data into Minitab much as done for Problem 6.4. Right below C1 type "block", below C2 type "treat" and below C3 type "yield".

d) Select Stat>ANOVA>Two-way, select "C3 yield" as the response and "C1 block" as the row factor and "C2 treat" as the column factor. Click on "Fit additive model," click on "Store residuals" and click on "Store fits." Then click on "OK."

e) **block response scatterplot**: Use file commands "Edit>Command Line Editor" and write the following lines in the window.
GSTD
LPLOT 'yield' vs 'block' codes for 'treat'

f) Click on the submit commands box and print the plot. Click on the output and then click on the *printer* icon.

g) Copy (www.math.siu.edu/olive/regdata.txt) into *R*.

Type the following commands to get the following Anova table.

```
z<-aov(yield~block+treat,pen)
summary(z)
```

```
            Df  Sum Sq Mean Sq F value  Pr(>F)
block        4 264.000  66.000  3.5044 0.04075 *
treat        3  70.000  23.333  1.2389 0.33866
Residuals   12 226.000  18.833
```

h) Did blocking appear to help?

i) Perform a 4 step F test for whether yield depends on treatment.

Latin Square Designs in SAS and R

(Latin square designs can be fit by Minitab, but not with Students' version of Minitab.)

7.5. This problem is for a Latin square design and uses data from Box, Hunter and Hunter (2005, p. 157-160).

Copy and paste the SAS program for this problem from (www.math.siu.edu/olive/reghw.txt).

a) Click on the output and use the menu commands "Edit>Select All" and "Edit>Copy. In *Word* use the menu commands "Edit>Paste" then use the left mouse button to highlight the first page of output. Then use the menu command "Edit>Cut." Then there should be one page of output including the Anova table. Print out this page.

b) Copy the data for this problem from (www.math.siu.edu/olive/regdata.txt) into R. Use the following commands to create a residual plot. Copy and paste the plot into *Word*. (Click on the plot and simultaneously hit the *Ctrl* and *c* buttons. Then go to *Word* and use the menu commands "Edit>Paste.")

```
z<-aov(emissions~rblocks+cblocks+additives,auto)
summary(z)
plot(fitted(z),resid(z))
title("Residual Plot")
abline(0,0)
```

c) Use the following commands to create a response plot. Copy and paste the plot into *Word*. (Click on the plot and simultaneously hit the *Ctrl* and *c* buttons. Then go to *Word* and use the menu commands "Edit>Paste.")

```
attach(auto)
FIT <- auto$emissions - z$resid
```

```
plot(FIT,auto$emissions)
title("Response Plot")
abline(0,1)
detach(auto)
```

d) Do the plots look ok?

e) Were the column blocks useful? Explain briefly.

f) Were the row blocks useful? Explain briefly.

g) Do an appropriate 4 step test.

7.6. Obtain the Box, Hunter and Hunter (2005, p. 146) penicillin data from (www.math.siu.edu/olive/regdata.txt) and the R program `ganova2` from (www.math.siu.edu/olive/regpack.txt). The program does graphical Anova for completely randomized block designs.

a) Enter the following commands and include the plot in *Word* by simultaneously pressing the *Ctrl* and *c* keys, then using the menu commands "Copy>Paste" in *Word*.

```
attach(pen)
ganova2(pen$treat,pen$block,pen$yield)
detach(pen)
```

b) Blocking seems useful because some of the scaled block deviations are outside of the spread of the residuals. The scaled treatment deviations are in the middle of the plot. Do the treatments appear to be significantly different?