XIII. Incomplete Block Designs

(Box, Hunter and Hunter, 1978, sec.8.3; Cochran and Cox, 1957, sec. 9.1, ch.10, secs 11-1–11.5)

Often, there are reasons as to why the conditions for an RCBD cannot be met; that is, the number of plots per block cannot equal the number of treatments. In these situations an incomplete block design might be appropriate.

For example, you may be running an experiment that involves 6 different sets of ingredients in preparing chicken and only 4 chickens can be cooked at the same time. There are thus 6 treatments but only a possible 4 plots per block. Incomplete block designs (BIBDs) are useful in this situation.

XIII.A Balanced incomplete block designs (BIBDs)

a) Design of BIBDs

Definition XIII.1: A balanced incomplete block design is one in which each of the t treatments is replicated r times and occurs at most once in each of the b blocks that contain k plots and the arrangement of treatments in blocks is such that each pair of treatments occurs together the same number of times (λ) in a block. ■

The first condition means that that the total number of units $= tr = bk$, while the second condition implies that total number of plots with other treatments in the blocks in which a treatment occurs $= \lambda(t-1) = r(k-1)$. A BIBD cannot exist if these two relationships are not met. However, that both of these conditions are satisfied does not imply that a BIBD must exist. For example, a BIBD does not exist for $t = 15$, $k = 5$, $b = 21$, $r = 7$ and $\lambda = 2$, even though both conditions are satisified.

Such designs are **not orthogonal**, however, they are balanced. That is to say they are not orthogonal because treatments are confounded with both blocks and plots within blocks. They are balanced because all comparisons between treatments are confounded with blocks to the same extent, as they are with plot within blocks. It can be shown that for a BIBD the proportion of the information within blocks is $e_2 = (t\lambda)/(kr)$ and between blocks is $e_1 = 1 - e_2$. These proportions are called the

canonical efficiency factors which are always values between zero and one and sum to one for a particular randomized term, in this case Treatments. It is desirable that e_2 is as close to one as possible as this implies that as much of the information as possible is confounded with plots, which are less variable than blocks.

Designs can be obtained from Cochran and Cox (1957) and Box, Hunter and Hunter (1978). The following table is an extract of their table 11.3 and indexes the plans they give.

TABLE 11.3 INDEX TO PLANS

t	*k*	*r*	*b*	$\lambda^{\dagger}$	e_2	Plan	*t*	*k*	*r*	*b*	$\lambda^{\dagger}$	e_2	Plan
4	2	3	6	1	0.67	11.1	11	2	10	55	1	0.55	11.4a
	3	3	4	2	0.89	*		5	5	11	2	0.88	11.19
5	2	4	10	1	0.62	11.2		6	6	11	3	0.92	11.20
	3	6	10	3	0.83	11.1a		10	10	11	9	0.99	*
	4	4	5	3	0.94	*	13	3	6	26	1	0.72	11.21
6	2	5	15	1	0.60	11.3		4	4	13	1	0.81	11.22
	3	5	10	2	0.80	11.4		9	9	13	6	0.96	11.23
	3	10	20	4	0.80	11.5	15	3	7	35	1	0.71	11.24
	4	10	15	6	0.90	11.6		7	7	15	3	0.92	11.25
	5	5	6	4	0.96	*		8	8	15	4	0.94	11.26
7	2	6	21	1	0.58	11.2a	16	6	6	16	2	0.89	11.27
	3	3	7	1	0.78	11.7		6	9	24	3	0.89	11.28
	4	4	7	2	0.88	11.8		10	10	16	6	0.96	11.29
	6	6	7	5	0.97	*	19	3	9	57	1	0.70	11.30
8	2	7	28	1	0.57	11.9		9	9	19	4	0.94	11.31
	4	7	14	3	0.86	11.10		10	10	19	5	0.95	11.32
	7	7	8	6	0.98	*	21	3	10	70	1	0.70	11.33
9	2	8	36	1	0.56	11.3a		5	5	21	1	0.84	11.34
	4	8	18	3	0.84	11.11		7	10	30	3	0.90	11.35
	5	10	18	5	0.90	11.12	25	4	8	50	1	0.78	11.36
	6	8	12	5	0.94	11.13		9	9	25	3	0.93	11.37
	8	8	9	7	0.98	*	28	4	9	63	1	0.78	11.38
10	2	9	45	1	0.56	11.14		7	9	36	2	0.89	11.39
	3	9	30	2	0.74	11.15	31	6	6	31	1	0.86	11.40
	4	6	15	2	0.83	11.16		10	10	31	3	0.93	11.41
	5	9	18	4	0.89	11.17	37	9	9	37	2	0.91	11.42
	6	9	15	5	0.93	11.18	41	5	10	82	1	0.82	11.43
	9	9	10	8	0.99	*	57	8	8	57	1	0.89	11.44
							73	9	9	73	1	0.90	11.45
							91	10	10	91	1	0.91	11.46

†Number of times that two treatments appear together in the same block.

*These plans are constructed by forming all possible combinations of the *t* numbers in groups of size *k*. The number of blocks *b* serves as a check that no group has been missed

Example XIII.1 BIBD for four treatments in blocks of three

Consider an experiment to investigate fabric wear using a Martindale wear tester. Let's suppose it has the feature that only three pieces of cloth can be compared in any one run. The response is the weight loss in tenths of a milligram suffered by a test piece when it is rubbed against a standard grade of emery paper for a 1000 revolutions of the machine. Further suppose that it is desired to test four different cloths. There are thus 4 treatments but only three plots per block are possible.

The following is a BIBD for $t = 4$, $k = 3$, $b = 4$ and $r = 3$ that could be used:

Block				
I	A	B	C	
II	A	B	D	
III	A	C	D	
IV	B	C	D	

Note that each pair of treatments occurs together in $\lambda = 2$ blocks. Consequently, $e_2 = (t\lambda)/(kr) = (4 \times 2)/(3 \times 3) = 8/9 = 0.8888$ and $e_1 = 1 - e_2 = 1 - (8/9) = 1/9 = 0.1111$. That is, 88.9% of the information about treatments is between plots within blocks.

To randomize the design, the treatment combinations are randomized to the blocks and the treatments in a block are randomized to the plots. This can be done in S-Plus using the expressions given in Appendix B. The S-Plus output from running the instructions is as follows:

```
> library(DAE, lib.loc = "f:")
> set.seed(987)
> b <- 4
> k <- 3
> t <- 4
> n <- b * k
> Standard.Order <- factor(1:n)
> Random.Order <- order(rep(runif(b), each = k), runif(n))
> BIBD.Design <- fac.divide(Random.Order, factor.names = list(Blocks = b, Plots
      = k))
> Treat <- factor(c(1, 2, 3, 1, 2, 4, 1, 3, 4, 2, 3, 4), labels = c("A", "B",
     "C", "D"))
> BIBD.Design <- design(Standard.Order, Random.Order, BIBD.Design, Treat)
> BIBD.Design <- sort.col(BIBD.Design, "@ALL", "Random.Order")
> BIBD.Design
   Standard.Order Random.Order Blocks Plots Treat
 8              8            1      1     1     C
 7              7            2      1     2     A
 9              9            3      1     3     D
 3              3            4      2     1     C
 1              1            5      2     2     A
 2              2            6      2     3     B
12             12            7      3     1     D
11             11            8      3     2     C
10             10            9      3     3     B
 5              5           10      4     1     B
 6              6           11      4     2     D
 4              4           12      4     3     A
```

Thus the randomized layout is:

Block				
	I	C	A	D
	II	C	A	B
	III	D	C	B
	IV	B	D	A

Note that the treatments A, B and C, which started off in the first block, ended up in the second block. ■

Some balanced incomplete block designs have the desirable property of being resolvable.

Definition XIII.2: An incomplete block design is **resolvable** if the blocks can be grouped into replicates such that each treatment occurs once and only once in each replicate. ■

Example XIII.2: Resolvable BIBD

The following incomplete block design for $t = 9$, $k = 3$, $b = 12$ and $r = 4$ is resolvable as it consists of 4 complete replicates of the treatments.

		Plots		
		1	2	3
Reps	Blocks			
1	1	1	2	3
	2	4	5	6
	3	7	8	9
2	1	1	4	7
	2	2	5	8
	3	3	6	9
3	1	1	5	9
	2	2	6	7
	3	3	4	8
4	1	1	6	8
	2	2	4	9
	3	3	5	7

Note that each pair of treatments occurs together in $\lambda = 1$ blocks. Consequently, $e_2 = (t\lambda)/(kr) = (9 \times 1)/(3 \times 4) = 3/4 = 0.75$ and $e_1 = 1 - e_2 = 1 - (3/4) = 1/4 = 0.25$. That is, 75% of the information about treatments is between plots within blocks. ■

In randomizing such designs there is the choice of a) ignoring the replicates and randomizing the treatment combinations to the blocks completely at random or b) randomizing the combinations to the blocks within a replicate. Where one can

arrange for the blocks within a replicate to be more homogeneous, then it will be advantageous to keep the replicates together and randomize within a replicate. For example, in an agricultural experiment, it may be possible to divide the experimental site into relatively homogeneous area and to assign the replicates to these; these areas would then be subdivided into blocks that would in turn contain the plots. In an industrial experiment, it may not be possible to carry out the experiment in a single session. Use of resolvable designs enables the experiment to be carried out in stages, with one or more replicates dealt with at each stage. If the experiment has to be discontinued at any time then all the treatments will have occurred equally often.

b) Determining the analysis of variance table

Example XIII.3 Wear testing

In this example, taken from Box, Hunter and Hunter, Appendix 8D, seven types of fabric are to be tested using a wear tester with four positions for testing fabric. The design and data are as follows:

		Position			
		1	2	3	4
	I	563	248	252	627
		F	D	G	B
	II	233	344	226	442
		C	A	G	F
	III	297	211	160	251
		G	D	E	C
Run	IV	195	300	537	337
		E	G	B	A
	V	520	199	278	595
		B	E	C	F
	VI	196	369	185	606
		D	A	E	F
	VII	273	240	602	396
		D	C	B	A

Note that each pair of treatments occurs together on $\lambda = 2$ machines so that

$$e_2 = (t\lambda)/(kr) = (7\times 2)/(4\times 4) = 0.875 \text{ and } e_1 = 1 - 0.875 = 0.125$$

What are the components of the study?

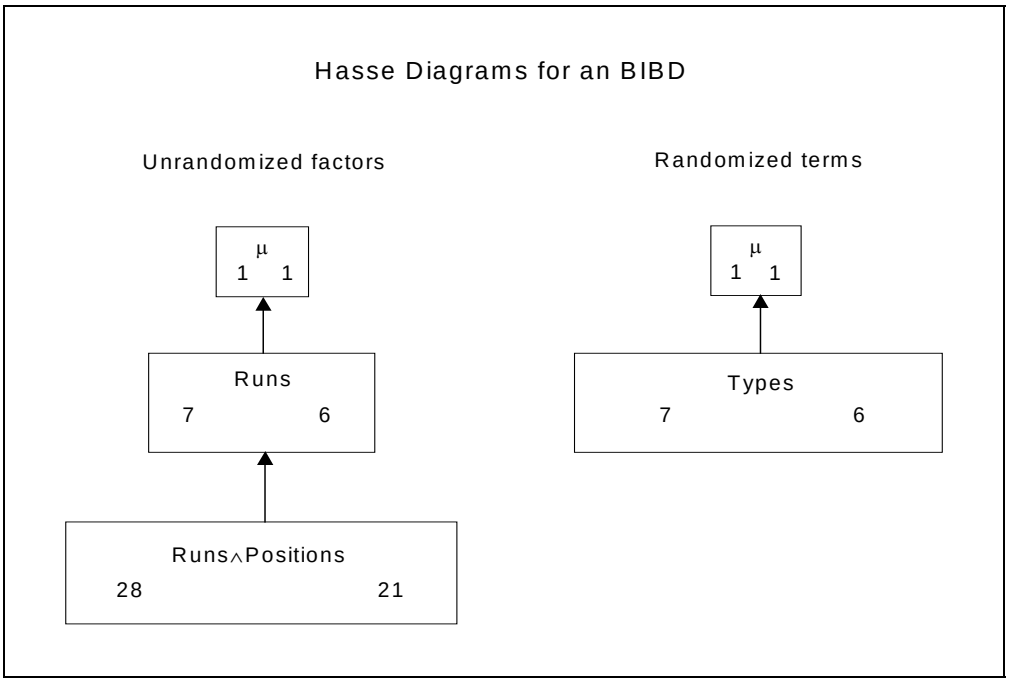
1. Observational unit – a position
2. Response variable – Wear
3. Unrandomized factors – Runs, Positions
4. Randomized factors – Types
5. Type of study – Balanced incomplete block design

The experimental structure is:

Structure	Formula
unrandomized	*7* Runs/*4* Positions
randomized	*7* Types

The formulae expand to give:

Runs/Positions = Runs + Positions[Runs]
and Types = Types



The models for this experiment, based on the unrandomized factors being random factors and the randomized factors being fixed factors are:

E[Y] = Types
and var[Y] = Runs + Runs∧Positions.

The analysis of variance table for this experiment has the following form:

Source	df	E[MSq]
Runs	6	
Types	6	$\sigma^2_{RP} + 4\sigma^2_R + 0.125q_T(\boldsymbol{\psi})$
Positions[Runs]	21	
Types	6	$\sigma^2_{RP} + 0.875q_T(\boldsymbol{\psi})$
Residual	15	σ^2_{RP}
Total	27	

The design chosen has resulted in the differences between treatments confounded with both runs and positions within runs, with most of the information being confounded with the latter. In this case there are no Runs degrees of freedom left over for a Residual. Consequently it is not possible to estimate σ_R^2 separately from $q_T(\boldsymbol{\psi})$. All information about Types will be derived from positions within runs and the 12.5% of the information confounded with runs will have to be forgone. ■

c) Significance test for BIBDs

As for the RCBD, there are two possible maximal models for BIBDs, depending on whether Blocks are deemed to be fixed or random. The maximal model for Blocks fixed is written symbolically as

E[Y] = Blocks + Treatment and Var[Y] = Blocks∧Plot

while that for Blocks random is

E[Y] = Treatment and Var[Y] = Blocks + Blocks∧Plot.

For either model, the analysis of variance involves investigating whether the Blocks and Treatment terms can be omitted from the maximal model. The expected mean squares under the two maximal models are given in the following table.

			E[MSq]		
Source	df	MSq	Blocks fixed	Blocks random	F
Blocks	$b-1$				
Treatments	$t-1$	$s_{T_1}^2$	$\sigma_{BP}^2 + e_1 q_{T_B}(\boldsymbol{\psi})$	$\sigma_{BP}^2 + k\sigma_B^2 + e_1 q_T(\boldsymbol{\psi})$ [&]	$s_{T_1}^2 / s_B^2$
Residual	$b-t$	s_B^2	$\sigma_{BP}^2 + q_B(\boldsymbol{\psi})$ [&]	$\sigma_{BP}^2 + k\sigma_B^2$	s_B^2 / s_R^2
Plots[Blocks]	$b(k-1)$				
Treatments	$t-1$	$s_{T_2}^2$	$\sigma_{BP}^2 + e_2 q_T(\boldsymbol{\psi})$ [&]	$\sigma_{BP}^2 + e_2 q_T(\boldsymbol{\psi})$ [&]	$s_{T_2}^2 / s_R^2$
Residual	$bk-b-t+1$	s_R^2	σ_{BP}^2	σ_{BP}^2	
Total	$bk-1$				

[&] e_1 = efficiency factor for blocks = $1 - e_2$

e_2 = efficiency factor for plots = $(t\lambda)/(kr)$

$q_T(\boldsymbol{\psi})$ is a quadratic form in $\boldsymbol{\psi}$, the observation length vector of parameters for the expectation i.e. $\boldsymbol{\psi} = \mathbf{X}\boldsymbol{\theta}$, and is a function of the τs.

$q_{T_B}(\boldsymbol{\psi})$ is a quadratic form in $\boldsymbol{\psi}$ that depends on both the βs and the τs

Note that there are two Treatment lines in the analysis, the first being referred to as the "interblock" Treatment line and the second as the "intrablock" Treatment line. Generally, one tries to have e_2 as close to one as possible and to base conclusions on the intrablock Treatment effects.

Because, when Blocks are fixed, $q_{T_B}(\boldsymbol{\psi})$ involves both βs and τs, it is not possible to separately test for treatment difference between blocks in this case — the intrablock test for treatments will be the only test for treatments that can be performed here. Thus it is preferable to designate Blocks as random, if this is appropriate.

The significance test for the case of Blocks random is a follows:

Step 1: Set up hypotheses

a) H_0: $\tau_1 = \tau_2 = ... = \tau_t$
H_1: at least one pair of population treatment means is different

b) H_0: $\sigma_B^2 = 0$
H_1: $\sigma_B^2 \neq 0$

Step 2: Calculate test statistics

The analysis of variance table for a BIBD is:

Source	df	MSq	E[MSq]	F
Blocks	$b-1$			
Treatments	$t-1$	$s_{T_1}^2$	$\sigma_{BP}^2 + k\sigma_B^2 + e_1 q_T(\boldsymbol{\psi})$ &	$s_{T_1}^2 / s_B^2$
Residual	$b-t$	s_B^2	$\sigma_{BP}^2 + k\sigma_B^2$	s_B^2 / s_R^2
Plots[Blocks]	$b(k-1)$			
Treatments	$t-1$	$s_{T_2}^2$	$\sigma_{BP}^2 + e_2 q_T(\boldsymbol{\psi})$ &	$s_{T_2}^2 / s_R^2$
Residual	$bk-b-t+1$	s_R^2	σ_{BP}^2	
Total	$bk-1$			

& e_1 = efficiency factor for blocks = 1 − e_2
e_2 = efficiency factor for plots = $(t\lambda)/(kr)$
$q_T(\boldsymbol{\psi})$ is a quadratic form in $\boldsymbol{\psi}$, the observation length vector of parameters for the expectation i.e. $\boldsymbol{\psi} = \mathbf{X}\boldsymbol{\theta}$, and is a function of the τs.

Step 3: Decide between hypotheses

If *p* for F is less than α, reject H_0.

d) Analysis for the example

Example XIII.3 Wear testing (continued)

The S-Plus expressions for obtaining the analysis for this example are given below.

```
attach(BIBDWear)
BIBDWear
library(DAE, lib.loc="f:")
boxplot(split(Wear, Runs), style.bxp="old", medchar=T, medpch=8)
boxplot(split(Wear, Types), style.bxp="old", medchar=T, medpch=8)
BIBDWear.aov <- aov(Wear ~ Runs + Types, BIBDWear)
BIBDWear.aov <- aov(Wear ~ Types + Error(Runs/Positions), BIBDWear)
summary(BIBDWear.aov)
#
# Diagnostic checking
#
res <- resid.errors(BIBDWear.aov)
fit <- fitted.errors(BIBDWear.aov)
data.frame(Runs,Positions,Types,Wear,res,fit)
plot(fit, res)
qqnorm(res)
qqline(res)
tukey.1df(BIBDWear.aov, BIBDWear, error.term="Positions %in% Runs")
#
# multiple comparisons
#
BIBDWear.tab <- model.tables(BIBDWear.aov, type="means")
BIBDWear.tab
sed <- as.vector(se.contrast(BIBDWear.aov, list(Types == "A", Types == "B"),
   data=BIBDWear))
q <- qtukey(0.95, 7, 15)
hsd <- q/sqrt(2)*sed
data.frame(sed,q,hsd)
#
# plotting
#
title("Fitted values for Wear")
Types.Mean <- BIBDWear.tab$tables$Types
BIBDWear.Means <- data.sheet(Types.Mean)
#then use 2D-Graphs to produce bar charts
```

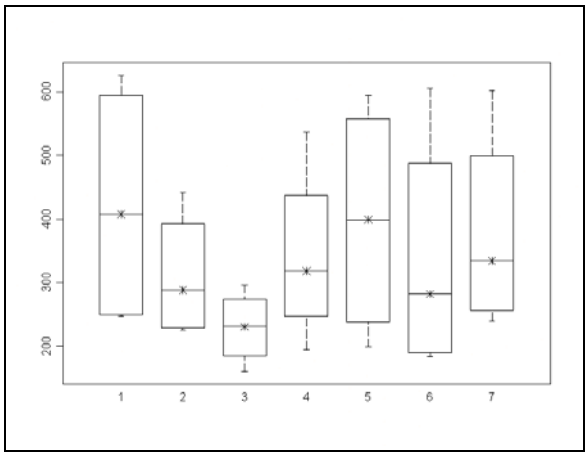
In analyzing a BIBD experiment an `Error` function must be used to obtain the full interblock and intrablock analyses. Consequently, the `Error` term needs to be specified in `tukey.1df` and `resid.errors` and `fitted.errors` are used to extract the residuals and fitted values.

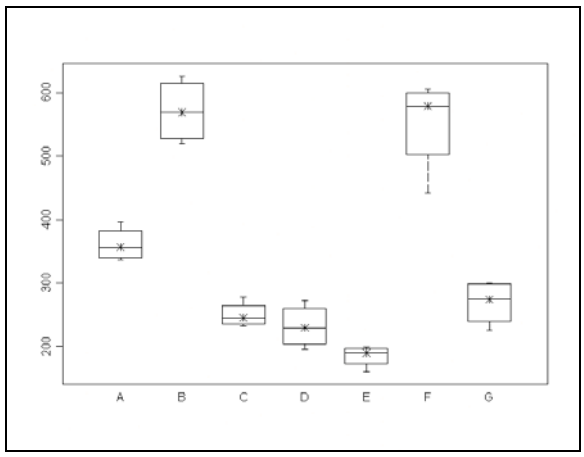
The S-Plus output produced by these expressions is:

```
> attach(BIBDWear)
> BIBDWear
   Runs Positions Types Wear
 1    1         1     F  563
 2    1         2     D  248
 3    1         3     G  252
 4    1         4     B  627
 5    2         1     C  233
 6    2         2     A  344
 7    2         3     G  226
 8    2         4     F  442
```

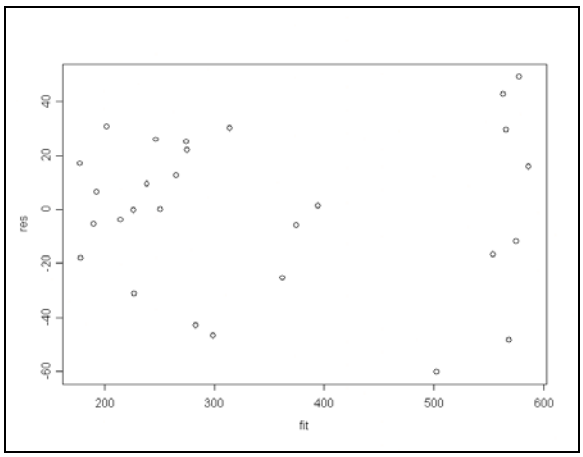
```
 9    3         1     G  297
10    3         2     D  211
11    3         3     E  160
12    3         4     C  251
13    4         1     E  195
14    4         2     G  300
15    4         3     B  537
16    4         4     A  337
17    5         1     B  520
18    5         2     E  199
19    5         3     C  278
20    5         4     F  595
21    6         1     D  196
22    6         2     A  369
23    6         3     E  185
24    6         4     F  606
25    7         1     D  273
26    7         2     C  240
27    7         3     B  602
28    7         4     A  396
> library(DAE, lib.loc = "f:")
> boxplot(split(Wear, Runs), style.bxp = "old", medchar = T, medpch = 8)
> boxplot(split(Wear, Types), style.bxp = "old", medchar = T, medpch = 8)
> BIBDWear.aov <- aov(Wear ~ Runs + Types, BIBDWear)
> summary(BIBDWear.aov)
          Df Sum of Sq  Mean Sq  F Value         Pr(F)
     Runs  6   97394.7 16232.45 11.03176 0.00008936571
    Types  6  506798.6 84466.43 57.40437 0.00000000169
Residuals 15   22071.4  1471.43
> BIBDWear.aov <- aov(Wear ~ Types + Error(Runs/Positions), BIBDWear)
> summary(BIBDWear.aov)
Error: Runs
      Df Sum of Sq  Mean Sq
Types  6  97394.71 16232.45

Error: Positions %in% Runs
          Df Sum of Sq  Mean Sq  F Value         Pr(F)
    Types  6  506798.6 84466.43 57.40437 1.687115e-009
Residuals 15   22071.4  1471.43

> #
# Diagnostic checking
#
> res <- resid.errors(BIBDWear.aov)
Refitting model to allow projection
> fit <- fitted.errors(BIBDWear.aov)
Refitting model to allow projection
> data.frame(Runs, Positions, Types, Wear, res, fit)
   Runs Positions Types Wear            res      fit
 1    1         1     F  563 -1.178571e+001 574.7857
 2    1         2     D  248  9.285714e+000 238.7143
 3    1         3     G  252 -4.678571e+001 298.7857
 4    1         4     B  627  4.928571e+001 577.7143
 5    2         1     C  233  3.064286e+001 202.3571
 6    2         2     A  344  3.007143e+001 313.9286
 7    2         3     G  226 -3.571429e-001 226.3571
 8    2         4     F  442 -6.035714e+001 502.3571
 9    3         1     G  297  2.200000e+001 275.0000
10    3         2     D  211 -3.928571e+000 214.9286
11    3         3     E  160 -1.807143e+001 178.0714
12    3         4     C  251 -4.507505e-014 251.0000
13    4         1     E  195  1.707143e+001 177.9286
14    4         2     G  300  2.514286e+001 274.8571
15    4         3     B  537 -1.678571e+001 553.7857
16    4         4     A  337 -2.542857e+001 362.4286
17    5         1     B  520 -4.842857e+001 568.4286
18    5         2     E  199  6.428571e+000 192.5714
19    5         3     C  278  1.250000e+001 265.5000
20    5         4     F  595  2.950000e+001 565.5000
21    6         1     D  196 -3.128571e+001 227.2857
22    6         2     A  369 -5.928571e+000 374.9286
23    6         3     E  185 -5.428571e+000 190.4286
24    6         4     F  606  4.264286e+001 563.3571
```

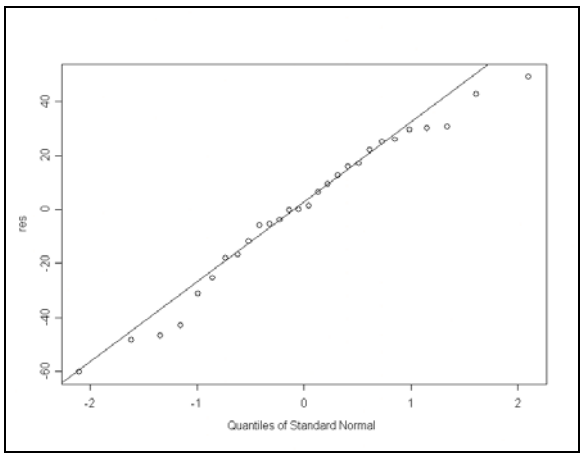
```
25    7         1     D  273  2.592857e+001 247.0714
26    7         2     C  240 -4.314286e+001 283.1429
27    7         3     B  602  1.592857e+001 586.0714
28    7         4     A  396  1.285714e+000 394.7143
> plot(fit, res)
> qqnorm(res)
> qqline(res)
> tukey.1df(BIBDWear.aov, BIBDWear, error.term = "Positions %in% Runs")
Refitting model to allow projection
Refitting model to allow projection
$SS.1df:
[1] 6340.96

$SS.res.1:
[1] 15730.47

$F.1df:
[1] 5.643407

$p.1df:
[1] 0.03234489
```





The boxplot indicate that the responses of the blocks are different and that there appears to be very strong differences between the types. However, as there is confounding between the blocks and types, these plots can only be taken as a rough guide as they do not separate the effects of the two factors.





The residual-versus-fitted-values plot is showing a mild tendency for the variance to increase with the magnitude of the fitted values. The normal probability plot indicates that the data are approximately normally distributed.

The significance test for this example is now given.

Step 1: Set up hypotheses

a) H_0: $\tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = \tau_6 = \tau_7$
H_1: at least one pair of population type means is different

b) H_0: $\sigma^2_R = 0$
H_1: $\sigma^2_R \neq 0$

Step 2: Calculate test statistics

The analysis of variance table for the example BIBD is:

Source	df	SSq	MSq	E[MSq]	F	Prob
Runs	6	97395	16232			
Types	6	506799	84466	$\sigma^2_{RP} + 4\sigma^2_R + 0.125 q_T(\boldsymbol{\psi})$ [&]		
Position[Runs]	21	528870				
Types	6	506799	84466	$\sigma^2_{RP} + 0.875 q_T(\boldsymbol{\psi})$ [&]	57.40	<0.001
Residual	15	22071	1471	σ^2_{RP}		
Nonadditivity	1	6341	6341		5.64	0.032
Deviation	14	15730	1124			
Total	27	626265				

[&] $e_2 = (t\lambda)/(kr) = (7\times 2)/(4\times 4) = 0.875$,
$e_1 = 1 - 0.875 = 0.125$
$q_T(\boldsymbol{\psi}) = 4\sum(\tau_k - \bar{\tau}_.)^2/6$

Step 3: Decide between hypotheses

Clearly, there are differences between the Types as evidenced by the line for Types under Positions within Runs ("intrablock"). A test for Runs and Types is not possible in this instance as $b = t$ and there is no Residual for Blocks. There is evidence of transformable non-additivity. It would appear that variance is increasing as the fitted values increase. This would need to be rectified before proceeding further. ■

e) Treatment differences for BIBDs

The same general considerations for the examination of treatment differences, as apply for the other designs, apply in the case of the BIBD; that is, one would use multiple comparison procedures or fit submodels as appropriate.

In the case of the multiple comparison procedures, the computation has to be amended to take into account the efficiency factor, e_2. The formulae for Tukey's HSD procedure becomes:

$$w(5\%) = \frac{q_{t,\nu,0.05}}{\sqrt{2}} s\sqrt{\frac{2}{e_2 r}}.$$

Note that $s\sqrt{2/e_2 r}$ is the **s**tandard **e**rror of **d**ifference or s.e.d.

Further, the quantities being compared are no longer simple treatment means; they are called *adjusted treatment means*, they having been adjusted to remove the influence of blocks. They are obtained from the intrablock information only.

Example XIII.3 Wear testing (continued)

```
> #
# multiple comparisons
#
> BIBDWear.tab <- model.tables(BIBDWear.aov, type = "means")
Refitting model to allow projection
> BIBDWear.tab

Tables of means
Grand mean

 345.79

 Types
      A      B      C      D      E      F      G
 367.43 558.79 255.86 219.79 182.93 555.86 279.86
> sed <- as.vector(se.contrast(BIBDWear.aov, list(Types == "A", Types == "B"),
     data = BIBDWear))
Refitting model to allow projection
> q <- qtukey(0.95, 7, 15)
> hsd <- q/sqrt(2) * sed
> data.frame(sed, q, hsd)
       sed     q      hsd
1 28.99683 4.782 98.04945
> #
# Plotting
#
> title("Fitted values for Wear")
> Types.Mean <- BIBDWear.tab$tables$Types
> BIBDWear.Means <- data.sheet(Types.Mean)
> #then use 2D-Graphs to produce bar charts
```

The adjusted treatment means, in increasing order, are:

	Type							
	E	D	C	G	A	F	B	w(5%)
Adjusted Means	182.9	219.8	255.9	279.9	367.4	555.9	558.8	98.05

$$w(5\%) = \frac{q_{t,\nu,0.05}}{\sqrt{2}} s\sqrt{\frac{2}{e_2 r}} = \frac{4.782}{\sqrt{2}}\sqrt{1471\frac{2}{4\times 0.875}}$$
$$= \frac{4.782}{\sqrt{2}} 28.99683$$
$$= 98.05$$

It would appear that F and B are greater than the rest and that, of the rest, A is greater than E, D and C with G intermediate between these two groups. However, given the significance of the test for nonadditivity, these conclusions should be ignored and an analysis for which the assumptions are met should be sought. ■

XIII.B Partially balanced incomplete block designs (PBIBDs)

A problem with BIBDs is that in order to attain balance *r* has to be large. In order to obtain designs with smaller values of *r*, designs that lack the complete symmetry of the balanced designs must be used. Instead of all pairs of treatments occurring together in a block the same number of times, the number of times pairs of treatments occur together in the blocks varies.

Definition XIII.3: An **associate class** of treatments is the set of treatments for which the number of times any pair of them occurs together in the same block is equal. Let this number be λ_i. ■

Definition XIII.4: A block design with $m > 1$ associate classes is called a partially balanced incomplete block design with *m* associate classes — PBIB(*m*). ■

The simplest partially balanced designs are those with two associate classes or PBIB(2)s.

Example XIII.4 PBIBD for a soya bean experiment

The following gives a design, in standard order, and the results for an experiment to investigate the effects of 25 varieties on the yield of soyabeans.

					Rep A					
Plot	1		2		3		4		5	
Block	Var	Yield	Var	Yield	Var	Yield	Var	Yield	Var	Yield
I	**1**	6	**2**	7	**3**	5	**4**	8	**5**	6
II	**6**	16	**7**	12	**8**	12	**9**	13	**10**	8
III	**11**	17	**12**	7	**13**	7	**14**	9	**15**	14
IV	**16**	18	**17**	16	**18**	13	**19**	13	**20**	14
V	**21**	14	**22**	15	**23**	11	**24**	14	**25**	14
					Rep B					
I	**1**	24	**6**	13	**11**	24	**16**	11	**21**	8
II	**2**	21	**7**	11	**12**	14	**17**	11	**22**	23
III	**3**	16	**8**	4	**13**	12	**18**	12	**23**	12
IV	**4**	17	**9**	10	**14**	30	**19**	9	**24**	23
V	**5**	15	**10**	15	**15**	22	**20**	16	**25**	19

It is a two-associate class design with $\lambda_1 = 0$ and $\lambda_2 = 1$. For example varieties 1 and 2 occur in the same block once, whereas varieties 1 and 7 never occur together in the same block. It is also a resolvable design with two complete replicates of the treatments. To obtain a BIBD design for this combination of treatments (25) and plots per block (5) would require 6 replicates. ■

The ANOVA for such experiments, now that we have computers, is not much more difficult to obtain than for a BIBD. However, multiple comparisons procedures are more complicated because there are different standard errors of difference for pairs of means, one corresponding to each associate class.

a) Design of PBIBDs

Cochran and Cox (1957) gives some plans for PBIBDs and there are tables that have been produced for others. However, these days the simplest way to obtain incomplete block designs is using computer software, of which CycDesN is an excellent example.

The randomization for a PBIBDs is similar to the BIBDs — the sets of treatments are randomized across the blocks and the treatments within sets are randomized to the plots within blocks.

b) Determining the analysis of variance table

Example XIII.4 PBIBD for a soya bean experiment (continued)

What are the components of the study?

1. Observational unit – a plot
2. Response variable – Yield
3. Unrandomized factors – Reps, Blocks, Plots
4. Randomized factors – Varieties
5. Type of study – Partially balanced incomplete block design (with two associate classes)

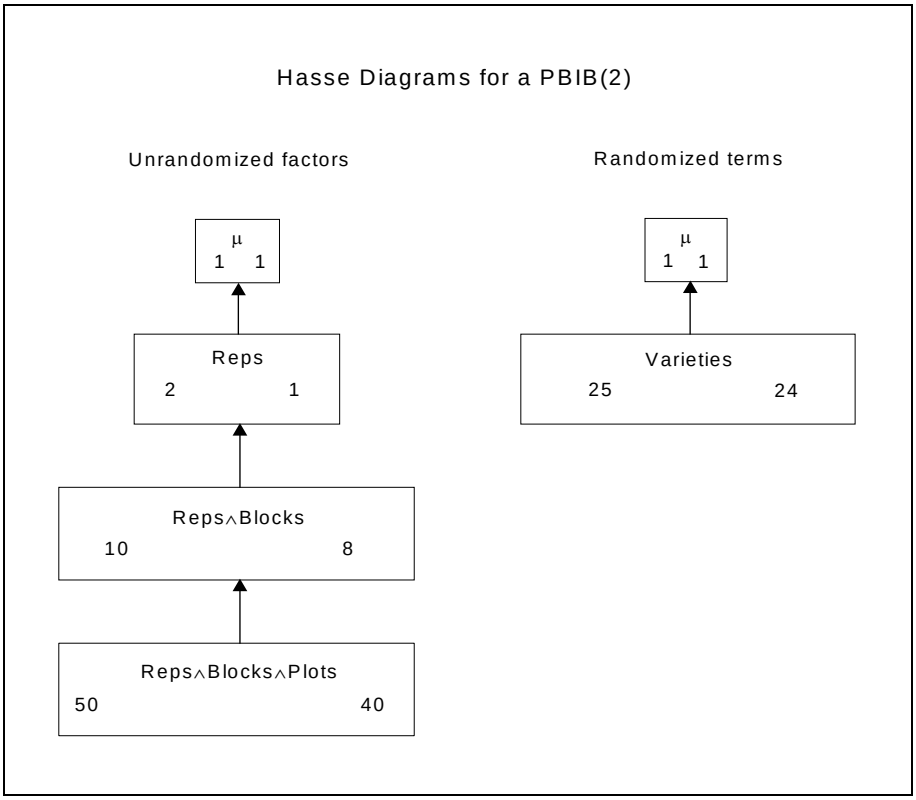
The experimental structure is:

Structure	Formula
unrandomized	*2* Reps/*5* Blocks/*5* Plots
randomized	*7* Varieties

Note that whereas nonresolvable IBDs have Blocks/Plots, resolvable IBDs have Reps/Blocks/Plots, provided the randomization respects the resolvability by randomizing treatment combinations to the blocks within a replicate. If treatment combinations are completely randomized to *all* blocks then the structure would be Blocks/Plots.

The formulae expand to give:

Reps/Blocks/Plots = Reps + Blocks[Reps] + Plots[Reps∧Blocks]
and Varieties = Varieties



The models for this experiment, based on the unrandomized factors being random factors and the randomized factors being fixed factors are:

E[Y] = Varieties

and var[Y] = Reps + Blocks[Reps] + Plots[Reps∧Blocks].

The analysis of variance table for this experiment has the following form:

Source	df	E[MSq]
Reps	1	$\sigma^2_{RBP} + 5\sigma^2_{RB} + 25\sigma^2_{R}$
Blocks[Reps]	8	
Varieties	8	$\sigma^2_{RBP} + 5\sigma^2_{RB} + q_V(\boldsymbol{\psi})$
Plots[Runs∧Blocks]	40	
Varieties	24	$\sigma^2_{RBP} + q_V(\boldsymbol{\psi})$
Residual	16	σ^2_{RBP}
Total	49	

The design chosen has resulted in only some of the differences between varieties, eight degrees of freedom worth, confounded with both blocks and plots within blocks; the other 16 degrees of freedom is confounded only with plots within blocks. This time we do not have a formula to compute the efficiency factors. However it can be shown that the efficiency factors for the eight degrees of freedom confounded with both blocks and plots is 0.5. For the 16 degrees of freedom confounded only with plots, the efficiency factor is clearly 1. Now as there are no Blocks[Reps] degrees of freedom left over for a Residual, it is not possible to estimate σ^2_{RB} separately from $q_V(\boldsymbol{\psi})$. All information about Varieties will be derived from plots within blocks and the 50% of the information on the eight degrees of freedom for Varieties confounded with Blocks will have to be forgone. ■

c) Analysis for the example

Example XIII.4 PBIBD for a soya bean experiment (continued)

The S-Plus expressions for obtaining the analysis for this example are given below.

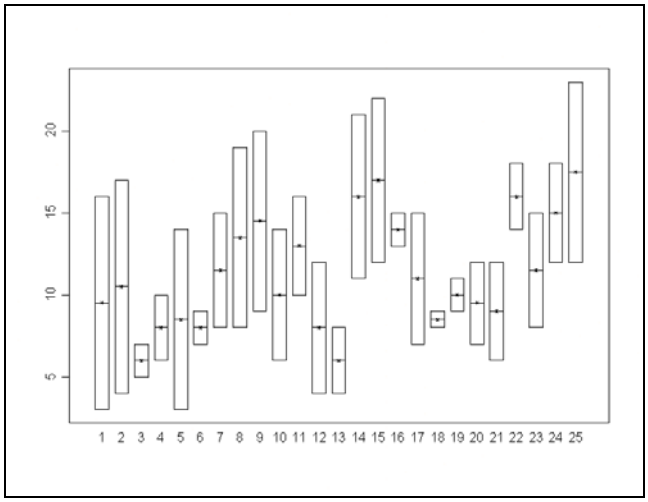
```
attach(PBIBDSoy)
PBIBDSoy
library(DAE, lib.loc="f:")
boxplot(split(Yield, Varieties), style.bxp="old", medchar=T, medpch=8)
PBIBDSoy.aov <- aov(Yield ~ Varieties + Error(Reps/Blocks/Plots), PBIBDSoy)
summary(PBIBDSoy.aov)
#
# Diagnostic checking
#
res <- resid.errors(PBIBDSoy.aov)
fit <- fitted.errors(PBIBDSoy.aov)
data.frame(Blocks,Plots,Varieties,Yield,res,fit)
plot(fit, res)
qqnorm(res)
qqline(res)
```

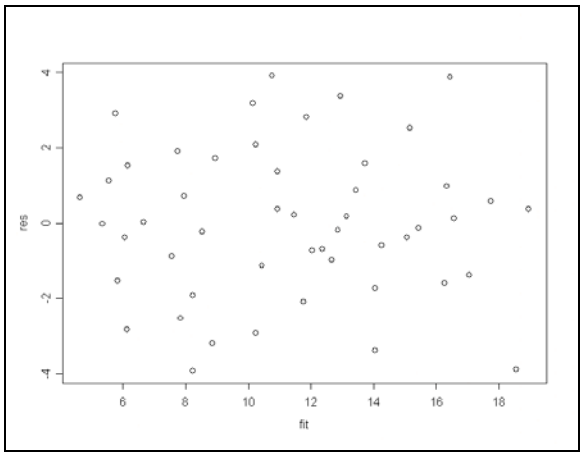
```
tukey.1df(PBIBDSoy.aov, PBIBDSoy, error.term="Plots %in% (Reps/Blocks)")
#
# multiple comparisons
#
PBIBDSoy.tab <- model.tables(PBIBDSoy.aov, type="means")
PBIBDSoy.tab
sed.1 <- as.vector(se.contrast(PBIBDSoy.aov, list(Varieties == "1", Varieties ==
   "2"), data=PBIBDSoy))
sed.2 <- as.vector(se.contrast(PBIBDSoy.aov, list(Varieties == "1", Varieties ==
   "7"), data=PBIBDSoy))
q <- qtukey(0.95, 7, 15)
hsd.1 <- q/sqrt(2)*sed.1
hsd.2 <- q/sqrt(2)*sed.2
data.frame(q,sed.1,hsd.1,sed.2,hsd.2)
```

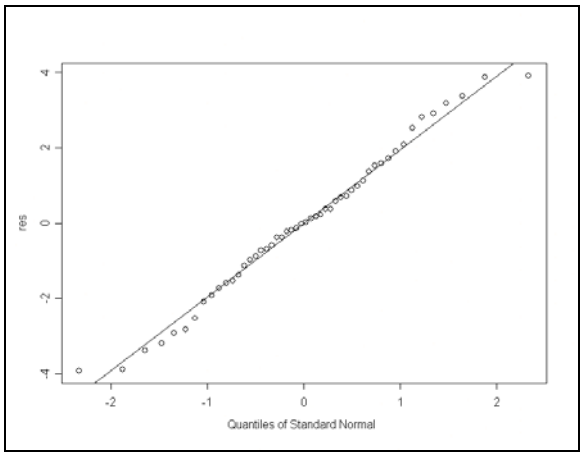
The S-Plus output produced by these commands is:

```
> attach(PBIBDSoy)
> PBIBDSoy
   Reps Blocks Plots Varieties Yield
 1    1      1     1         1     3
 2    1      1     2         2     4
 3    1      1     3         3     5
 4    1      1     4         4     6
 5    1      1     5         5     3
 6    1      2     1         6     7
 7    1      2     2         7     8
 8    1      2     3         8     8
 9    1      2     4         9     9
10    1      2     5        10     6
11    1      3     1        11    10
12    1      3     2        12     4
13    1      3     3        13     4
14    1      3     4        14    11
15    1      3     5        15    12
16    1      4     1        16    13
17    1      4     2        17     7
18    1      4     3        18     9
19    1      4     4        19     9
20    1      4     5        20    12
21    1      5     1        21    12
22    1      5     2        22    14
23    1      5     3        23    15
24    1      5     4        24    12
25    1      5     5        25    12
26    2      1     1         1    16
27    2      1     2         6     9
28    2      1     3        11    16
29    2      1     4        16    15
30    2      1     5        21     6
31    2      2     1         2    17
32    2      2     2         7    15
33    2      2     3        12    12
34    2      2     4        17    15
35    2      2     5        22    18
36    2      3     1         3     7
37    2      3     2         8    19
38    2      3     3        13     8
39    2      3     4        18     8
40    2      3     5        23     8
41    2      4     1         4    10
42    2      4     2         9    20
43    2      4     3        14    21
44    2      4     4        19    11
45    2      4     5        24    18
46    2      5     1         5    14
47    2      5     2        10    14
48    2      5     3        15    22
49    2      5     4        20     7
50    2      5     5        25    23
> library(DAE, lib.loc = "f:")
```

```
> boxplot(split(Yield, Varieties), style.bxp = "old", medchar = T, medpch = 8)
> PBIBDSoy.aov <- aov(Yield ~ Varieties + Error(Reps/Blocks/Plots), PBIBDSoy)
> summary(PBIBDSoy.aov)
Error: Reps
          Df Sum of Sq Mean Sq F Value Pr(F)
Residuals  1    359.12  359.12

Error: Blocks %in% Reps
          Df Sum of Sq Mean Sq
Varieties  8    351.76   43.97

Error: Plots %in% (Reps/Blocks)
          Df Sum of Sq Mean Sq  F Value     Pr(F)
Varieties 24    398.88  16.620 1.368464 0.2611893
Residuals 16    194.32  12.145

> #
# Diagnostic checking
#
> res <- resid.errors(PBIBDSoy.aov)
Refitting model to allow projection
> fit <- fitted.errors(PBIBDSoy.aov)
Refitting model to allow projection
> plot(fit, res)
> qqnorm(res)
> qqline(res)
> tukey.1df(PBIBDSoy.aov, PBIBDSoy, error.term = "Plots %in% (Reps/Blocks)")
Refitting model to allow projection
Refitting model to allow projection
$SS.1df:
[1] 2.96113

$SS.res.1:
[1] 191.3589

$F.1df:
[1] 0.2321134

$p.1df:
[1] 0.6369149

> #
# multiple comparisons
#
PBIBDSoy.tab <- model.tables(PBIBDSoy.aov, type = "means")
Refitting model to allow projection
> PBIBDSoy.tab

Tables of means
Grand mean

 11.28

 Varieties
      1      2      3     4      5      6     7      8      9     10     11
 12.619 12.061 10.053 8.937 9.4949 9.0486 11.28 16.524 14.292 9.2717 14.292
     12     13     14     15     16     17     18     19    20     21     22
 7.0403 7.8213 15.631 16.747 14.404 9.3833 9.6064 7.9329 7.375 7.1519 13.288
    23     24     25
 11.28 11.838 14.627
> sed.1 <- as.vector(se.contrast(PBIBDSoy.aov, list(Varieties == "1", Varieties ==
      "2"), data = PBIBDSoy))
Refitting model to allow projection
> sed.2 <- as.vector(se.contrast(PBIBDSoy.aov, list(Varieties == "1", Varieties ==
      "7"), data = PBIBDSoy))
Refitting model to allow projection
> q <- qtukey(0.95, 7, 15)
> hsd.1 <- q/sqrt(2) * sed.1
> hsd.2 <- q/sqrt(2) * sed.2
> data.frame(q, sed.1, hsd.1, sed.2, hsd.2)
      q    sed.1   hsd.1    sed.2    hsd.2
1 4.782 3.688668 12.4728 3.477709 11.75947
```







The output indicates that there are no problems with the assumptions and that there are no significant differences between the treatments. ■

XIII.C Exercises

XIII.1 The following is an unrandomized plan for a partially balanced incomplete block design.

Block Plot	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10
2	4	5	6	7	8	9	10	1	2	3
3	3	4	5	6	7	8	9	10	1	2

Use S-Plus with a seed of 1012 to obtain a randomized layout for the experiment.

XIII.2 An experiment was designed to compare five brands of perfume. People were to rate the perfume on a ten-point scale but it was decided that each person should only evaluate three perfumes at a time. This resulted in the use of a

BIBD with the judge acting as a block. The design is given in the following table and is resolvable, although this was ignored in assigning the treatment combinations to judges.

	Judge (Person)									
	I	II	III	IV	V	VI	VII	VIII	IX	X
Test										
1	1	1	1	1	1	1	2	2	2	3
2	2	2	2	3	3	4	3	3	4	4
3	3	4	5	4	5	5	4	5	5	5

What are the components of the study?

1. Observational unit _____
2. Response variable _____
3. Unrandomized factors _____
4. Randomized factors _____
5. Type of study _____

What is the experimental structure for this experiment?

Structure	Formula
unrandomized	
randomized	

What are the terms derived from the experimental structure? Write out the Hasse diagram for each structure formula.

What are the expected mean squares for the lines in the analysis of variance table based on all unrandomized factors being random and all randomized factors being fixed?

Write down the analysis of variance table, including the expected mean squares for the lines in it.

Source	df	E[MSq]
Total		

Comment on the advisability of using just the intrablock information about Perfume differences given the design employed.

The results of the experiment are given in the following table, in the same order as the treatments are given in the first table. The data are available in *BIBDPerfume.sdd* in the *AdvDesign* directory of the data share \\cwpool0\briencj and from the web site.

	Judge (Person)									
	I	II	III	IV	V	VI	VII	VIII	IX	X
Test										
1	4	4	3	6	5	8	6	7	5	6
2	5	4	4	9	8	10	9	10	8	7
3	8	6	5	8	6	7	9	8	6	4

Analyze the data using S-Plus, including diagnostic checking and the examination of treatment differences.