eNote 7

The Analysis of Split-Plot Experiments

Indhold

7	The	Analysis of Split-Plot Experiments	1
	7.1	Introduction	2
	7.2	The Split-Plot Model	3
	7.3	Example: Tenderness of pork.	6
	7.4	More Complicated Split-Plot Designs	9
		7.4.1 Yield of oats	10
		7.4.2 Example: Rancidness of steaks	14
	7.5	R-TUTORIAL: Analysing Split-plot data	17
		7.5.1 Tenderness of pork	17
		7.5.2 Yield of oats	19
		7.5.3 Rancidness of steaks	23
	7.6	Exercises	25

7.1 Introduction

When planning experiments with several factors it is sometimes necessary or desirable to use experimental units of varying sizes. It might for instance be possible to use fairly small units when dealing with different varieties in a field experiment whereas a given level of fertilizer typically has to be applied to a larger area. A comparison of the effect of the different fertilizer levels is then most logically made with reference to the natural variation between these larger areas, which are called *whole plots*. The effect of the different varieties and the interaction between variety and fertilizer level are naturally compared to the random variation between the smaller units, which are called *subplots*. Therefore it makes good sense to consider a model with random variation on both the whole plot and the subplot levels. In this module a number of exact formulas are given that applies to balanced cases. The ideas, principles and approaches still apply even if everything is NOT well balanced.

7.2 The Split-Plot Model

In the basic split-plot design we have two factors of interest, A with the *k* levels $a_1, ..., a_k$, and B with the *m* levels $b_1, ..., b_m$. We suppose that there are *n* replicates and consider $k \cdot n$ whole plots each consisting of *m* subplots, so that we in total have $k \cdot m \cdot n$ subplots. Ideally the whole plots should be randomized on the levels of A, which is called the *whole plot factor*, and the subplots within each whole plot should be randomized on the levels of B, which is the *subplot factor*. If we have k = 3 fertilizer levels, m = 4 varieties, and n = 2 replicates then one possible split-plot design is given in the following table:

a ₃	b_1	a_1	b_4	a_1	\mathtt{b}_1	a ₂	b ₃	a3	b ₃	a ₂	b ₂
a ₃	b ₃	a_1	b_3	a ₁	b ₃	a ₂	b ₂	a ₃	b_1	a ₂	b_1
a ₃	b ₂	\mathtt{a}_1	\mathtt{b}_1	a_1	b_4	a ₂	\mathtt{b}_1	a3	b ₂	a ₂	b_4
a ₃	b_4	\mathtt{a}_1	b_2	a_1	b_2	a ₂	\mathtt{b}_4	a_3	b_4	a ₂	b ₃

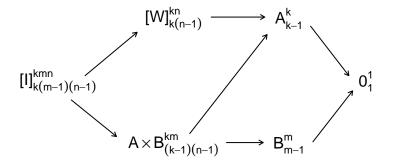
Typically we want to investigate the possibility of an interaction between the two factors A and B so we also consider the product factor $A \times B$. As the whole plots are the experimental units corresponding to the factor A it seems natural to assess the effect of the whole plot factor based on the variation between the whole plots. This implies that we should include the factor W, with levels $1, 2, \dots, k \cdot n$, in the analysis as a random effects factor. Note that W is finer than the whole plot factor A. The factor diagram associated with the split-plot experiment is depicted in Figure 7.1.

We suppose that we have observations Y_1, \ldots, Y_N , where *N*, the total number of observations, equals $k \cdot m \cdot n$. The split-plot model is given by

$$Y_i = \gamma(\mathbf{A}_i, \mathbf{B}_i) + d(\mathbf{W}_i) + \varepsilon_i, \quad i = 1, \dots, N,$$
(7-1)

where

$$\begin{aligned} d(j) &\sim N(0, \sigma_W^2), \quad j = 1, \dots, k \cdot n, \\ \varepsilon_i &\sim N(0, \sigma^2), \quad i = 1, \dots, N, \end{aligned}$$



Figur 7.1: The factor structure diagram for the split-plot experiment.

and all the variables $d(1), \ldots, d(kn)$ and $\varepsilon_1, \ldots, \varepsilon_N$ are mutually independent. The splitplot model is a mixed model. As in the case of the one-way analysis of variance model with a random effect (the "two-layer model") we have that the variance of the observations is the sum of two components,

$$\operatorname{Var} Y_i = \sigma_W^2 + \sigma^2.$$

The variance parameter σ^2 models the variation between subplots within whole plots and σ_W^2 the variation between whole plots.

A test for no interaction between the two factors A and B corresponds to considering the additive split-plot model

$$Y_i = \alpha(\mathbf{A}_i) + \beta(\mathbf{B}_i) + d(\mathbf{W}_i) + \varepsilon_i, \quad i = 1, \dots, N,$$
(7-2)

where the same assumptions hold for the random variables as for the split-plot model given by (7-1). From the factor diagram in Figure 7.1 we see that the interaction has to be tested against the variation between subplots, that is we get exactly the same *F*-test statistic as in the case where W is considered as a fixed effects factor,

$$F=\frac{\mathrm{MS}_{\mathrm{AB}}}{\mathrm{MS}_{\mathrm{e}}},$$

where MS_{AB} is the difference between the residual sums of squares for the two models (7-1) and (7-2) when η is treated as non-random, divided by (k - 1)(m - 1). Similarly MS_e is the residual sum of squares corresponding to the split-plot model (7-1) when η is a fixed effect, divided by k(m - 1)(n - 1). Under the additive split-plot model *F* is F((k - 1)(m - 1), k(m - 1)(n - 1))-distributed.

In the additive split-plot model the difference between two effects corresponding to the whole plot factor, $\alpha(a_j) - \alpha(a_{j'})$, is estimated by

$$\hat{\alpha}(\mathbf{a}_j) - \hat{\alpha}(\mathbf{a}_{j'}) = \bar{Y}_{\mathbf{a}_j} - \bar{Y}_{\mathbf{a}_{j'}}, \quad j, j' = 1, \dots, k,$$

with variance

$$\operatorname{Var}(\hat{\alpha}(\mathbf{a}_j) - \hat{\alpha}(\mathbf{a}_{j'})) = \frac{2(m\sigma_W^2 + \sigma^2)}{mn}.$$
(7-3)

Similarly for the subplot factor B we have that

$$\operatorname{Var}(\hat{\beta}(\mathbf{b}_j) - \hat{\beta}(\mathbf{b}_{j'})) = \frac{2\sigma^2}{kn}, \quad j, j' = 1, \dots, m.$$
(7-4)

From (7-3) and (7-4) we see that the effect of A is relatively less accurately estimated than the effect of B since the expected whole plot variation given by $m\sigma_W^2 + \sigma^2$ is larger than the subplot variation σ^2 . This is especially the case if there is a large variation between whole plots. The reason is that in this experimental design we have randomized the levels of A on the whole plots so that an experimental unit corresponding to A is a whole plot.

Based on the additive split-plot model given by (7-2) a test for an effect of the whole plot factor A, that is the hypothesis H_0 : $\alpha(a_1) = \cdots = \alpha(a_k)$, is equivalent to considering the model

$$Y_i = \beta(\mathsf{B}_i) + d(\mathsf{W}_i) + \varepsilon_i, \quad i = 1, \dots, N,$$
(7-5)

with identical distributional assumptions about the random effects as for the split-plot model given by (7-1). From the factor diagram in Figure 7.1 we see that A has to be tested against the variation between whole plots, that is against the random factor W. The *F*-test statistic is given by

$$F = \frac{\mathrm{MS}_{\mathrm{A}}}{\mathrm{MS}_{\mathrm{W}}},\tag{7-6}$$

where

$$MS_{A} = \frac{mn}{k-1} \sum_{j=1}^{k} (\bar{Y}_{a_{j}} - \bar{Y}_{0})^{2},$$

and MS_W is the difference between the residual sums of squares in the split-plot model (7-1) where W is treated as fixed, and the two-way analysis of variance model with the

factors A and B (and the interaction between them), divided by k(n - 1). Under the model (7-5) *F* is F(k - 1, k(n - 1))-distributed.

In a similar way we can test for an effect of the subplot factor B in the additive split-plot model by considering the hypothesis H_0 : $\beta(b_1) = \cdots = \beta(b_m)$, or equivalently the model

$$Y_i = \alpha(\mathbf{A}_i) + d(\mathbf{W}_i) + \varepsilon_i, \quad i = 1, \dots, N,$$
(7-7)

with the same assumptions about the random effects as previously. Like the interaction between A and B the effect of the factor B has to be tested against the subplot variation, see Figure 7.1. The corresponding *F*-test statistic is given by

$$F = \frac{\mathrm{MS}_{\mathrm{B}}}{\mathrm{MS}_{\mathrm{e}}},\tag{7-8}$$

where

$$MS_{\rm B} = \frac{kn}{k-1} \sum_{j=1}^{m} (\bar{Y}_{{\sf b}_j} - \bar{Y}_{\sf 0})^2$$

and where MS_e now is the residual sum of squares associated with the additive splitplot model (7-2) divided by (kn - 1)(m - 1). If the model (7-7) holds then *F* is F(m - 1, (kn - 1)(m - 1))-distributed.

If we accept to describe data either by model (7-5) or (7-7) we would then proceed to test for an effect of the remaining systematic factor. Since the two factors are tested against different variations (the whole plot factor A against the whole plot variation, and the subplot factor B against the subplot variation) we get exactly the same *F*-test statistics, that is (7-6) in the case of model (7-7) and (7-8) in the case of model (7-5). Suppose that we end up with model (7-7) as the final model then the estimated effects of the whole plot factor are given by

$$\hat{\alpha}(\mathbf{a}_j) = \bar{Y}_{\mathbf{a}_j}, \quad j = 1, \dots, k,$$

and a 95%-confidence interval is given by

$$\alpha(\mathbf{a}_j): \bar{Y}_{\mathbf{a}_j} \pm t_{0.975,k(n-1)} \sqrt{MS_W/(mn)}.$$

7.3 Example: Tenderness of pork.

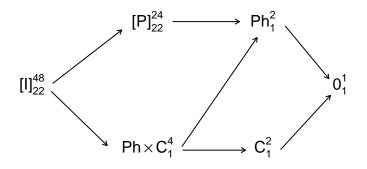
To investigate the effect of the cooling process on the tenderness of meat, 24 porks were slaughtered and cut in halves (right and left side). One side was then cooled by tunnel-cooling (a very quick cooling denoted 'TC' below) and the other side by (conventional) fast cooling (denoted 'FC' below). Furthermore the pigs were divided into two groups:

12 pigs with high pH (ph-group 3) and 12 pigs with low pH (ph-group 2). Two registrations associated with tenderness are recorded below: a sensory (subjective) evaluation of tenderness (denoted tender) and the lenghts of 'sarcomers'. Long sarcomers are suspected to be accociated with less tender meat and could be caused by too fat cooling. The data are given in the following table:

Porker	pH-group	Tunnel	Fast
1	low	7.22	5.56
2	low	3.11	3.33
3	low	7.44	7.00
4	low	4.33	4.89
5	low	6.78	6.56
6	low	5.56	5.67
7	low	7.33	6.33
8	low	4.22	5.67
9	low	3.89	4.00
10	low	5.78	5.56
11	low	6.44	5.67
12	low	8.00	5.33
13	high	8.44	8.44
14	high	7.11	6.00
15	high	6.00	5.78
16	high	7.56	7.67
17	high	5.11	4.56
18	high	8.67	8.00
19	high	5.78	7.67
20	high	6.11	5.67
21	high	7.44	7.56
22	high	7.67	6.11
23	high	8.00	8.22
24	high	8.78	8.44

In this experiment we have the 3 factors, C with levels tunnel and fast, pH with levels low and high, and P with levels 1,..., 24. It is a split-plot design with porker as whole plot and pH as whole plot factor, and side as subplot with C as subplot factor. We have k = m = 2 and n = 12 replicates. The factor diagram is shown in Figure 7.2. We consider the split-plot model given by (7-1) and the additive split-plot model given by (7-2).

The test of fixed effects are given in the following table:



Figur 7.2: The factor structure diagram for the pork tenderness experiment

Source of	Numerator degrees	Denominator degrees	F	P-value
variation	of freedom	of freedom		
phgroup	1	22	8.67	0.0075
cooling	1	22	2.25	0.1479
phgroup*cooling	1	22	0.18	0.6790

The interaction is clearly non-significant. Thus we clearly accept the additive split-plot model. In the additive split-plot model we can then proceed to test whether there is an effect of each of the whole plot factor and the subplot factor:

Source of	Numerator degrees	Denominator degrees	F	P-value
variation	of freedom	of freedom		
phgroup	1	22	8.67	0.0075
cooling	1	22	2.33	0.1403

So that we would not accept to exclude the factor pH from the analysis but find no significant effect of the factor C. Therefore we consider the model

$$Y_i = \alpha(\mathrm{pH}_i) + d(\mathrm{P}_i) + \varepsilon_i, \quad i = 1, \dots, 48,$$

where

$$d(j) \sim N(0, \sigma_W^2), \quad j = 1, \dots, 24,$$

$$\varepsilon_i \sim N(0, \sigma^2), \quad i = 1, \dots, 48,$$

and all the random variables are assumed to be mutually independent. A test for the hypothesis that there is no difference between the two pH-groups, that is

 $H_0: \alpha(low) = \alpha(high),$

results in the same *F*-test statistic of F = 8.6665 and *p*-value of 0.0075 as in the additive split-plot model, and thus we conclude that the systematic part of the model can not be simplified any further.

Estimates of the fixed effects parameters can be seen to be given by

$$\hat{\alpha}(low) = 5.6529,$$

 $\hat{\alpha}(high) = 7.1163,$

corresponding to the averages of the tenderness measurements for each pH-group. Estimates for the variance components are given by

$$\hat{\sigma}_W^2 = 1.2463,$$

 $\hat{\sigma}^2 = 0.4725.$

As

$$MS_{P} = 2.9651$$

we get the following 95%-confidence intervals for the expected tenderness in the two pH-groups,

$$\hat{lpha}(\texttt{low}) \pm t_{0.975,22} \sqrt{2.9651/24} = 5.6529 \pm 0.7289$$

or [4.9240,6.3819]. Similarly we get [6.3873,7.8452] for α (high). All in all we conclude that the tenderness is significantly higher if pH is high compared to if it is low.

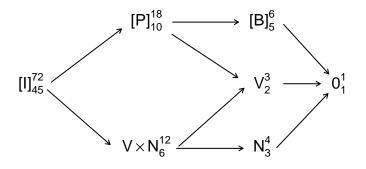
7.4 More Complicated Split-Plot Designs

In the previous section we have considered the simplest split-plot design with only two treatment factors. Naturally this model can be extended in various ways for example by including blocks, covariates, or more complex treatment structures. In this section we shall consider two examples of more complicated split-plot designs. In the first example we look at a split-plot design with blocks.

7.4.1 Yield of oats.

In a field experiment with three oats varieties v_1 , v_2 , and v_3 , and four equidistant levels of nitrogen n_0 , n_1 , n_2 , and n_3 , the yield of oats in quarter pounds was observed. The varieties were sown in 6 randomized blocks with three plots in each block, and with each plot divided into four subplots. The data and design, from Yates (1935), are reproduced in the following table:

710	<i>n</i> ₃	156	<i>n</i> ₂	118]	<i>n</i> ₂	109	n_3	99	710
v_3	n_1	140	n_0	105		n_0	63	n_1	70	v_3
v_1	n_0	111	n_1	130		n_0	80	<i>n</i> ₂	94	710
01	n_3	174	<i>n</i> ₂	157		n_3	126	n_1	82	v_2
v_2	<i>n</i> ₀	117	n_1	114		n_1	90	n_2	100	v_1
υZ	<i>n</i> ₂	161	n_3	141		n_3	116	n_0	62	
	<i>n</i> ₂	104	<i>n</i> ₀	70	1	<i>n</i> ₃	96	n_0	60	
v_3	$\frac{n_2}{n_1}$	89	n_0	117		n_2	89	$\frac{n_0}{n_1}$	102	v_2
			113	117]		07	101	102	
	<i>n</i> ₃	122	n_0	74	1	<i>n</i> ₂	112	n_3	86	
v_1	n_1	89	n_0	81		$\frac{n_2}{n_0}$	68	n_1	64	v_1
			- <u></u>]			. 1		
	<i>n</i> ₁	103	n_0	64]	<i>n</i> ₂	132	n_3	124	
v_2	n_2	132	n_3	133		n_1	129	n_0	89	v_3
			-		1					
		100			1					
v_2	n_1	108	n_2	126	-	<i>n</i> ₂	118	n_0	53	v_1
-	n_3	149	n_0	70]	n_3	113	n_1	74	-
		4.4.4		404	1	[101		0.(
v_3	n_3	144	n_1	124		n_3	104	<i>n</i> ₂	86	v_2
-	n_2	121	n_0	96]	n_0	89	n_1	82	
		(1		100	1	[
v_1	<i>n</i> ₀	61	n_3	100		<i>n</i> ₀	97	<i>n</i> ₁	99	v_3
-	<i>n</i> ₁	91	n_2	97]	<i>n</i> ₂	119	n_3	121	-



Figur 7.3: Factor diagram for the oats experiment.

This is a split-plot design with blocks, meaning that the replicates are in different blocks. Thus we have the factors V (variety), N (nitrogen), P (plot), and B (block) with levels

V
$$v_1, v_2, v_3$$
N n_0, n_1, n_2, n_3 P $1, 2, \dots, 18$ B $1, 2, \dots, 6$

V is the whole plot factor and N is the subplot factor. Furthermore, we have that P is finer than B (knowing which plot the experimental unit is in means that we know which block it is in). In fact P is equivalent to the interaction between B and V. Treating the block factor as random we have the factor diagram given in Figure 7.3.

Letting Y_i denote the yield of the *i*'th subplot we have that the model for the split-plot design with blocks is given by

$$Y_i = \gamma(\mathbf{V}_i, \mathbf{N}_i) + d(\mathbf{B}_i) + f(\mathbf{P}_i) + \varepsilon_i, \quad i = 1, \dots, 72,$$
(7-9)

where

$$\begin{array}{ll} d(j) & \sim & N(0,\sigma_B^2), \\ f(j') & \sim & N(0,\sigma_P^2), \\ \varepsilon_i & \sim & N(0,\sigma^2), \end{array}$$

for j = 1, ..., 6, j' = 1, ..., 18, and i = 1, ..., 72, and with all the random variables assumed to be mutually independent. We are interested in assessing whether there is an interaction between the variety and the nitrogen factors and if not then subsequently whether there is any effect of the two factors at all. The table of tests of fixed effects is given by:

Source of	Numerator degrees	Denominator degrees	F	P-value
variation	of freedom	of freedom		
fertil	3	45	37.69	<.0001
variety	2	10	1.49	0.2724
fertil*variety	6	45	0.30	0.9322

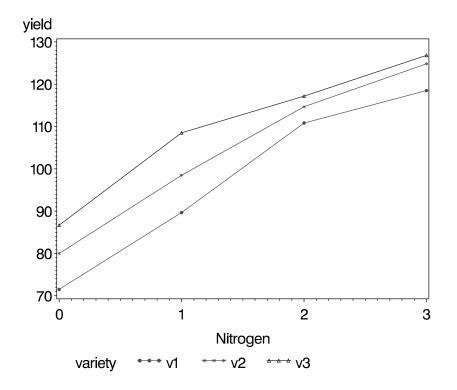
and leaving out the interaction term:

Source of	Numerator degrees	Denominator degrees	F	P-value
variation	of freedom	of freedom		
fertil	3	51	41.05	<.0001
variety	2	10	1.49	0.2724

Thus we accept that there is no difference between the varieties but that there is a clear effect of the nitrogen levels. This last conclusion is not altered when leaving out the whole plot factor V from the analysis as that only influences the mean square corresponding to the plot factor P, and the subplot factor N is tested against the subplot variation. Therefore we have a final model given by

$$Y_i = \beta(N_i) + d(B_i) + f(P_i) + \varepsilon_i, \quad i = 1, ..., 72,$$
 (7-10)

and the same assumptions about the random effects as previously. In Figure 7.4 the average yield for the 12 combinations of variety and nitrogen is given. Since the four nitrogen levels are equidistant, it makes sense to plot the averages versus the numbers 0, 1, 2 and 3.



Figur 7.4: Average yield as a function of nitrogen level for each variety.

Based on Figure 7.4 it might seem surprising that we did not find any significant effect of V in the analysis as the 3 curves are clearly shifted compared to each other. One has to remember, however, that the effect of the varieties is less accurately determined than the effect of the nitrogen level, and that a shift in the curves exactly corresponds to a whole plot variation.

Estimates for the variance components are

$$\hat{\sigma}_B^2 = 210.42,$$

 $\hat{\sigma}_P^2 = 121.85,$
 $\hat{\sigma}^2 = 162.56.$

It is typically not of interest to test whether there is a significant block or plot effect so we will not do that in this example.

Estimates of the expected yield for the four nitrogen levels (LSMEANS) and the 95% confidence bands are given by:

	n_0	n_1	<i>n</i> ₂	n_3
LSMEANS	79.4	98.9	114.2	123.4
Lower	62.3	81.8	97.2	106.3
Upper	96.4	116.0	131.3	140.4

Pair-wise comparisons of nitrogen levels show that n_1 gives a significantly larger yield than n_0 (P-value < 0.0001), n_2 significantly larger than n_1 (P-value 0.0007) and n_3 significantly larger than n_2 (P-value < 0.0358). In other words we see a clear increase in the yield with increasing levels of nitrogen. However, the increase from n_2 to n_3 is not so pronounced though still just significant.

In the second example we consider a split-plot experiment where the subplot factor is the interaction between two factors.

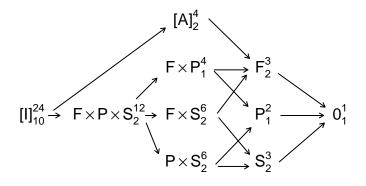
7.4.2 Example: Rancidness of steaks.

In order to investigate the effect of two different types of feed (1 and 2) and 6 different storage conditions on the rancidness of steaks from cattle given the feeds, 4 cows were randomized on the two types of feed so that 2 cows were given each type. After slaugh-tering 6 steaks from each animal were used in a storage experiment. Two types of packaging (A and B) were used along with 3 storage periods (2, 5, and 8 weeks). After the storage period the rancidness of each steak was determined by an LTBA-measurement. The data, which are part of a larger experiment, are given in the following table:

			А			В	
Animal	Feed	2	5	8	2	5	8
1	1	2.0	2.3	2.3	0.6	2.1	2.1
2	2	0.8	1.7	1.4	0.4	0.8	2.1 1.2
3	1	1.2	2.1	2.7	0.9	1.5	1.9
4	2	0.4	0.9	1.3	0.1	1.1	0.9

In this experiment we have 4 factors ANIMAL or A, FEED or F, PACKAGING or P, and STORAGE PERIOD or S with levels

A 1, 2, 3, 4 F 1, 2 P A, B S 2, 5, 8



Figur 7.5: Factor diagram for the rancidness experiment.

A split-plot design is used with animal as whole plot and feed as whole plot factor. The steaks are subplots and the subplot factor is the combination of the packaging type and storage period, that is the interaction between P and S, $P \times S$. All of the three factors (and their interactions) making up the whole plot factor and the subplot factor are of interest in this experiment. The factor diagram is depicted in Figure 7.5.

We let Y_i denote the LTBA-measurement for the *i*'th steak and consider the following split-plot model

$$Y_i = \delta(\mathbf{F}_i, \mathbf{P}_i, \mathbf{S}_i) + d(\mathbf{A}_i) + \varepsilon_i, \quad i = 1, \dots, 24,$$
(7-11)

where

$$\begin{array}{lll} d(j) & \sim & N(0, \sigma_A^2), \\ \varepsilon_i & \sim & N(0, \sigma^2), \end{array}$$

for j = 1, ..., 4, and i = 1, ..., 24, and where we assume that all the random variables are mutually independent. The table of tests of fixed effects is given by:

Source of	Numerator degrees	Denominator degrees	F	P-value
variation	of freedom	of freedom		
feed	1	2	30.37	0.0314
pack	1	10	14.15	0.0037
feed*pack	1	10	1.05	0.3291
storage	2	10	21.90	0.0002
feed*storage	2	10	0.51	0.6154
pack*storage	2	10	0.34	0.7187
feed*pack*storage	2	10	0.29	0.7510

Reducing the full model one step at a time starting with the 3-way interaction followed by each 2-way interaction (we do not give all the details) leads to the model with only main effects:

$$Y_i = \alpha(\mathbf{F}_i) + \beta(\mathbf{P}_i) + \gamma(\mathbf{S}_i) + d(\mathbf{A}_i) + \varepsilon_i, \quad i = 1, \dots, 24,$$
(7-12)

In this model the table of tests of fixed effects is is given by:

Source of	Numerator degrees	Denominator degrees	F	P-value
variation	of freedom	of freedom		
feed	1	2	30.37	0.0314
pack	1	17	18.02	0.0005
storage	2	17	27.90	<.0001

We see that all main effects are significant even though it is fairly close regarding the effect of the type of feed. This might be because of the fact that F is the whole plot factor and as such is tested against the variation between the whole plots, here the animals. It makes sense that it is hard to find a significant effect of the type of feed since we only have 4 experimental units for this factor. Nevertheless there is a significant effect of feed in this experiment. But note how an F-statistic of 30.37 usually extremely significant is just barely so in this case. Often significance tests with very low denominator degrees of freedom are meaningless, since the typical non-significant result really says nothing. This is so because the statistical power of such a test is extremely low, or in different words: even rather large treatment differences will typically NOT be detected.

Estimates of the variance components are

$$\hat{\sigma}^2 = 0.0699,$$

 $\hat{\sigma}^2_A = 0.0145.$

Note that the variation between animals is much smaller than the variation between steaks when it comes to rancidness in this experiment.

Estimates of differences between two effects, with corresponding 95%-confidence intervals, are

$\hat{\alpha}(1) - \hat{\alpha}(2)$	=	0.89,	[0.20,	1.59],
$\hat{\beta}(\mathbf{A}) - \hat{\beta}(\mathbf{B})$	=	0.46,	[0.23,	0.69],
$\hat{\gamma}(2)-\hat{\gamma}(8)$	=	-0.93,	[-1.20, -	-0.65],
$\hat{\gamma}(5)-\hat{\gamma}(8)$	=	-0.16,	[-0.44,	0.12].

So we see an increase in the expected rancidness of steaks from cattle receiving feed of type 1 compared to feed of type 2. Similarly, the expected rancidness of steaks stored in packaging of type A is higher compared to packaging of type B. Finally we see an increase in the expected rancidness with the length of the storage period, though there is not a significant difference between the expected rancidness of steaks stored for 5 weeks and 8 weeks.

7.5 R-TUTORIAL: Analysing Split-plot data

7.5.1 Tenderness of pork

Consider the analysis of the data cooling.txt described above and also in eNote13. First we import the data, and define the factors:

```
cooling <- read.table("cooling.txt", sep=",", header=TRUE)
cooling$pigno <- factor(cooling$pigno)
cooling$phgroup <- factor(cooling$phgroup)</pre>
```

Then we run the full model with the interaction term cooling×phgroup. For comparison with the standard anova table we first run a fully fixed version:

Then we run the proper mixed one:

eNote 7 7.5 R-TUTORIAL: ANALYSING SPLIT-PLOT DATA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
cooling	1	1.04	1.04	2.25	0.1479
phgroup	1	25.70	25.70	55.35	0.0000
pigno	22	65.23	2.96	6.39	0.0000
cooling:phgroup	1	0.08	0.08	0.18	0.6790
Residuals	22	10.21	0.46		

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)
cooling	1.04	1.04	1.00	22.00	2.25	0.1479
phgroup	4.02	4.02	1.00	22.00	8.67	0.0075
cooling:phgroup	0.08	0.08	1.00	22.00	0.18	0.6790

The interaction is not significant and therefore the model can be reduced to the additive model

library(lmerTest)
model2 <- lmer(tender ~ cooling + phgroup + (1 | pigno), data = cooling)
print(xtable(anova(model2)))</pre>

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)
cooling	1.04	1.04	1.00	23.00	2.33	0.1403
phgroup	3.88	3.88	1.00	22.00	8.67	0.0075

The factor cooling is not significant, resulting in the model

model3 <- lmer(tender ~ phgroup-1 + (1 | pigno), data = cooling)</pre>

which cannot be reduced any more. Due to eh simplicity of the model we can get most of the interesting information by using the confint function:

print(xtable(confint(model3)))

	2.5 %	97.5 %
.sig01	0.74	1.52
.sigma	0.53	0.94
phgroup2	4.97	6.34
phgroup3	6.43	7.80

Removal of 1mer objects once the analysis is done using the function rm

```
rm(model1, model2, model3)
```

7.5.2 Yield of oats

Consider the analysis of the data oatyield.txt described above and also in eNote13. First we import the data, and define the necessary factors:

```
oatyield <- read.table("oatyield.txt", sep=",", header=TRUE)
oatyield$block <- factor(oatyield$block)</pre>
```

The oatyield experiment is a split-plot design with blocks. In comparison to the cooling experiment this implies that there are two random factors corresponding to the blocks and the whole plots. The random factors are specified by the factor block and the product factor block×variety.

Again, the full model is fit first by a fully fixed model:

And then the mixed version:

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
fertil	3	20020.50	6673.50	37.69	0.0000
variety	2	1786.36	893.18	5.04	0.0106
block	5	15875.28	3175.06	17.93	0.0000
fertil:variety	6	321.75	53.62	0.30	0.9322
variety:block	10	6013.31	601.33	3.40	0.0023
Residuals	45	7968.75	177.08		

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)
fertil	20020.50	6673.50	3.00	45.00	37.69	0.0000
variety	526.06	263.03	2.00	10.00	1.49	0.2724
fertil:variety	321.75	53.62	6.00	45.00	0.30	0.9322

The interaction is not significant and the reduced model is

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)
fertil	20020.50	6673.50	3.00	51.00	41.05	0.0000
variety	482.91	241.46	2.00	10.00	1.49	0.2724

The final model only contains fertil in the fixed effects part

We could have gotten to this, including a check of the two random effect by the step function of lmerTest: (but some people would prefer to include the two random effects in the model whether they are signnificant or not)

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)
fertil	20020.50	6673.50	3.00	51.00	41.05	0.0000

mystep <- step(model1)
print(xtable(mystep\$rand.table))</pre>

	Chi.sq	Chi.DF	elim.num	p.value
block	5.97	1	kept	0.01
block:variety	9.19	1	kept	0.00

print(xtable(mystep\$anova.table))

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	elim.num	Pr(>F)
fertil:variety	321.75	53.62	6	45.00	0.30	1	0.93
variety	482.91	241.46	2	10.00	1.49	2	0.27
fertil	20020.50	6673.50	3	51.00	41.05	kept	0.00

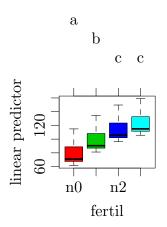
Pairwise comparisons could be obtained by extracting the LSMEANS differences from this:

```
mystep <- step(model1)
print(xtable(mystep$diffs.lsmeans.table))</pre>
```

Remember to use your own 'Bonferroni' correction when looking at the p-values: multiply them all with 6. Or we could do the posthoc analysis with the multcomp-package:

```
library(multcomp)
tuk3 <- glht(model3, linfct = mcp(fertil = "Tukey"))
tuk.cld <- cld(tuk3)
old.par <- par(mai=c(1,1,1.25,1), no.readonly=TRUE)
plot(tuk.cld, col=2:6)</pre>
```

	Estimate	Standard Error	DF	t-value	Lower CI	Upper CI	p-value
fertil n0-n1	-19.50	4.25	51.00	-4.59	-28.03	-10.97	0.00
fertil n0-n2	-34.83	4.25	51.00	-8.20	-43.37	-26.30	0.00
fertil n0-n3	-44.00	4.25	51.00	-10.35	-52.53	-35.47	0.00
fertil n1-n2	-15.33	4.25	51.00	-3.61	-23.87	-6.80	0.00
fertil n1-n3	-24.50	4.25	51.00	-5.76	-33.03	-15.97	0.00
fertil n2-n3	-9.17	4.25	51.00	-2.16	-17.70	-0.63	0.04



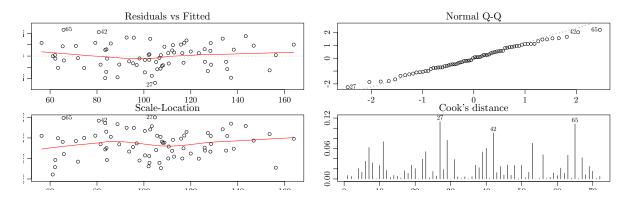
And the confidence limits of the standard deviations in the model:

print(xtable(confint(model3, 1:3)))

	2.5 %	97.5 %
.sig01	6.28	19.25
.sig02	3.21	28.86
.sigma	10.38	15.15

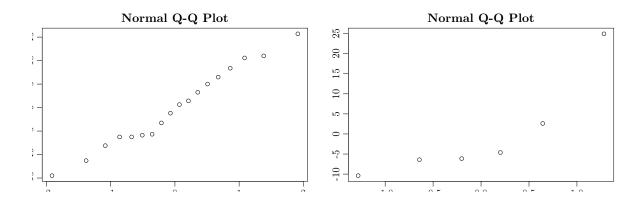
And what about model diagnostics, first from the fixed model:

par(mar=c(1,1,2,3))
par(mfrow=c(2,2))
plot(lm1, which = 1:4)



And then the random effects normality:

par(mar=c(1,1,2,3))
par(mfrow=c(1,2))
qqnorm(ranef(model3)\$'block:variety'[,1])
qqnorm(ranef(model3)\$block[,1])



7.5.3 Rancidness of steaks

Consider the analysis of the data rancid.txt described above and also in eNote13. First we import the data, and define the necessary factors:

```
rancid <- read.table("rancid.txt", sep=",", header=TRUE)
rancid$animal <- factor(rancid$animal)
rancid$feed <- factor(rancid$feed)
rancid$storage <- factor(rancid$storage)</pre>
```

The design in the data set rancid is a split-plot design without blocks. Three factors describe the fixed effects part of the model, and the full model with a three-way interaction is specified

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	elim.num	Pr(>F)
feed:pack:storage	0.05	0.03	2	10.00	0.29	1	0.75
pack:storage	0.06	0.03	2	12.00	0.39	2	0.69
feed:storage	0.09	0.05	2	14.00	0.63	3	0.55
feed:pack	0.09	0.09	1	16.00	1.37	4	0.26
feed	2.12	2.12	1	2.00	30.37	kept	0.03
pack	1.26	1.26	1	17.00	18.02	kept	0.00
storage	3.90	1.95	2	17.00	27.90	kept	0.00

And post hoc:

print(xtable(confint(model1, 1:2)))

	2.5 %	97.5 %
.sig01	0.00	0.27
.sigma	0.16	0.30

print(xtable(mystep\$lsmeans.table))

	feed	pack	storage	Estimate	Standard Error	DF	t-value	Lower CI	Upper C
feed 1	1			1.81	0.11	2.00	15.81	1.32	2.3
feed 2	2			0.92	0.11	2.00	8.01	0.42	1.4
pack A		А		1.59	0.10	4.10	16.37	1.32	1.8
pack B		В		1.13	0.10	4.10	11.65	0.87	1.4
storage 2			2	0.80	0.11	6.50	7.19	0.53	1.0
storage 5			5	1.56	0.11	6.50	14.05	1.30	1.8
storage 8			8	1.73	0.11	6.50	15.51	1.46	1.9

	Estimate	Standard Error	DF	t-value	Lower CI	Upper CI	p-value
feed 1-2	0.89	0.16	2.00	5.51	0.20	1.59	0.03
pack A-B	0.46	0.11	17.00	4.25	0.23	0.69	0.00
storage 2-5	-0.76	0.13	17.00	-5.77	-1.04	-0.48	0.00
storage 2-8	-0.92	0.13	17.00	-7.00	-1.20	-0.65	0.00
storage 5-8	-0.16	0.13	17.00	-1.23	-0.44	0.12	0.24

print(xtable(mystep\$diffs.lsmeans.table))

7.6 Exercises

Exercise 1 Yield of wheat

An experiment with 6 varieties of wheat concerning the treatment of a fungus attack was carried out on an experimental field with 5 blocks each consisting of 12 plots. The experimental design and the yield (hkg/ha) for each plot are given in the following table, where the treatment (T=treated, U=untreated) and the variety (an=anja, ci=citadel, ko=kosack, kr=kraka, sl=sleipner, vu=vuka) are also given.

U	U	U	U	U	U	Т	Т	Т	Т	Т	Т			
kr	sl	ko	an	ci	vu	ko	sl	vu	ci	an	kr			
58.5	58.6	59.6	32.8	57.8	30.6	75.2	74.2	63.1	76.4	69.4	75.6			
U	U	U	U	U	U	Т	Т	Т	Т	Т	Т			
an	vu	sl	kr	ko	ci	ci	ko	vu	sl	an	kr			
30.5	28.5	52.0	54.1	56.1	54.4	79.8	74.4	62.2	76.5	65.4	64.8			
Т	Т	Т	Т	Т	Т	U	U	U	U	U	U			
kr	ko	ci	an	vu	sl	vu	kr	an	ko	ci	sl			
64.2	75.2	80.1	61.3	57.9	78.2	27.1	55.0	31.4	62.7	60.1	59.5			
U	U	U	U	U	U	Т	Т	Т	Т	Т	Т			
an	sl	kr	ci	vu	ko	vu	ko	ci	kr	an	sl			
26.7	54.9	47.9	54.7	24.4	53.8	57.3	77.2	84.4	62.7	58.9	82.2			

Т	Т	Т	Т	Т	Т	U	U	U	U	U	U
an	sl	ko	ci	vu	kr	ci	vu	sl	ko	kr	an
61.9	80.2	73.3	77.5	63.1	62.1	53.2	25.9	60.6	60.4	52.9	31.1

- a) What kind of experimental design is used here? How should the randomization be performed in such an experiment?
- b) Draw the factor diagram and write up a statistical model for the data.
- c) Analyze the data and make conclusions.

The data file can be downloaded as: fungus.txt and is described also in eNote13.

Exercise 2 Yield of oats

Consider the example with the yield of oats from the main text of this module. Assume that the following nitrogen levels were used:

Name	Level
<i>n</i> ₀	0
n_1	1
<i>n</i> ₂	2
<i>n</i> ₃	3

a) Investigate whether the yield dependence of nitrogen level can be expressed by a linear function, and if so use this to summarize the results.

Exercise 3 Rooting of Hibiscus

The data below (kindly supplied by Lise Bertram, KVL) origin from an experiment investigating the effect of conductivity in the water bath on the rooting of Hibiscus.

There were 16 water baths, four for each of the four levels of conductivity: 0.5, 1.0, 1.5 and 2.0 micro-Siemens per square meter (coded as 1, 2, 3, 4 in the data set). There were 10 cuttings of each of three varieties (Casablanca, Red rosa-sinensis, and Holiday) in each water bath, but only data from the latter two varieties are used here. The 30 cuttings were placed at random in the water bath. For each cutting the number roots (rootno) and the weight of roots (rootwt) were registered after 28 days. The four conductivity treatments were randomized on the 16 water baths. The results were (with one line per cutting)

variety	bathno	conduct	rootno	rootwt	
h	1	1	9	.0096	
h	1	1	4	.001	
h	1	1	5	.0007	
h	1	1	6	.0069	
h	1	1	7	.0116	
h	1	1	3	.0102	
h	1	1	5	.0011	
h	1	1	10	.0116	
h	1	1	9	.0167	
h	1	1	10	.0309	
h	5	1	6	.0123	
h	5	1	8	.0025	
		((320 line	es in tot	tal)
r	10	4	5	.0269	
r	10	4	10	.0275	

The data file can be downloaded as: conduct.txt and is described also in eNote13.

- a) Write down all the factors relevant for the analysis, and their levels and mutual structure. Are they crossed or nested, for example?
- b) Specify a statistical model, including the specification of random and fixed effects.

c) Analyse the root weight and draw conclusions.

Exercise 4 Desease in cucumbers

A greenhouse experiment was carried out to investigate how the spread of a disease (agurkesyge) in cucumbers depended on climate and on amount of fertilizer for two varieties. The following data (kindly supplied by Eigil de Neergård, Department of Plant Pathology, KVL) are an extract from the experiment. Two climates were used, (1) change to day temperature 3 hours before sunrise, and (2) normal change to day temperature. Three amounts of fertilizer were applied, normal (2.0 units), high (3.5 units), and extra high (4.0 units). The two varieties were Aminex and Dalibor.

At a certain time the plants were 'standardized' to have equally many leaves, and then (on day 0, say) the plants were contaminated with the disease. On 8 particular subsequent days the amount of infection (in percent) was registred. From the resulting curve of infection two summary measures were calculated (in a way not specified here), namely the rate of spread of the disease, and the level of infection at the end of the period.

There were 3 blocks each consisting of 2 sections, a section being a part of the greenhouse. Each section consisted of 3 plots, which were each divided into 2 subplots, each of which had 6-8 plants. Thus there were a total of 36 subplots. Results were recorded for each subplot.

The experimental factors were randomly allocated to the different units as follows: the 2 climates were allocated to the 2 sections within each block, the 3 amounts of fertilizer were allocated to the 3 plots within each section, and finally the 2 varieties were allocated to the 2 subplots within each plot. Thus, in summary, there were

3 blocks

- 2 sections per block (given 2 different climates)
- 3 plots per section (given 3 different amounts of fertilizer)
- 2 subplots per plot (with 2 different varieties)

The results were

block	section	plot	climate	fert	variety	rate	level
1	1	1	2	2.0	aminex	48.8981	0.06915
1	1	1	2	2.0	dalibor	42.2463	0.06595
1	1	2	2	3.5	aminex	48.2108	0.04679
1	1	2	2	3.5	dalibor	41.6767	0.04881
1	1	3	2	4.0	aminex	55.4369	0.04025
1	1	3	2	4.0	dalibor	40.9562	0.04859
1	2	4	1	2.0	aminex	51.5573	0.09353
1	2	4	1	2.0	dalibor	36.7739	0.10353
1	2	5	1	3.5	aminex	47.9937	0.05327
1	2	5	1	3.5	dalibor	47.8723	0.04397
1	2	6	1	4.0	aminex	57.9171	0.05225
1	2	6	1	4.0	dalibor	37.7185	0.09324
2	3	7	2	2.0	aminex	60.1747	0.04182
2	3	7	2	2.0	dalibor	45.6937	0.06983
2	3	8	2	3.5	aminex	51.0017	0.08863
2	3	8	2	3.5	dalibor	52.2796	0.03622
2	3	9	2	4.0	aminex	51.1251	0.05875
2	3	9	2	4.0	dalibor	48.7217	0.08169
2	4	10	1	2.0	aminex	51.6001	0.07001
2	4	10	1	2.0	dalibor	50.4463	0.09907
2	4	11	1	3.5	aminex	48.3387	0.05788
2	4	11	1	3.5	dalibor	38.6538	0.06834
2	4	12	1	4.0	aminex	51.3147	0.05695
2	4	12	1	4.0	dalibor	38.2488	0.07908
3	5	13	1	2.0	aminex	49.6958	0.07218
3	5	13	1	2.0	dalibor	29.6786	0.11351
3	5	14	1	3.5	aminex	46.6692	
3	5	14	1	3.5	dalibor	36.5892	
3	5	15	1	4.0	aminex	56.0320	
3	5	15	1	4.0	dalibor	36.0955	0.08712
3	6	16	2	2.0	aminex	45.9790	
3	6	16	2	2.0	dalibor	37.2489	
3	6	17	2	3.5	aminex	40.7277	
3	6	17	2	3.5		38.4831	
3	6	18	2	4.0		44.5242	
3	6	18	2	4.0	dalibor	34.3907	0.09651

The data file can be downloaded as: cucumber.txt and is described also in eNote13.

a) Write down all the factors relevant for the analysis, and their levels and mutual

structure. Are they crossed or nested, for example?

- b) Analyse the rate of spread of the disease to investigate the effect of the different factors.
- c) Analyse also the level of infection.