

STATS660.2 Case studies in linear mixed models

Anonymous

November 18, 2010

Copyright This work is licensed under the Creative Commons Attribution-ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/1.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Contents

1	Irrigation	3
1.1	The Data	3
1.2	The Model and Hypothesis	3
1.3	Matrix formulation	4
1.4	Anova tables, tests and estimates of variance components.	6
1.5	Estimating equations (for μ and β) and solutions.	7
1.6	REML estimates (of variances).	8
1.7	Estimate with all effects fixed.	9
1.8	The SS-formulas in the ANOVA tables.	10
1.9	SS computed by hand.	10
1.10	ML estimates.	11
1.11	Compare estimates of fixed effects (sl 30 vs. sl 60)	11
1.12	Compare variance and SS estimates (sl 25, sl 35, sl 55, sl 60)	12
1.13	Compare results from assuming fixed vs. random effects.	12
1.14	Discussion	13
2	Barley seed example	14
2.1	Anova tables, tests and estimates of variance components	14
2.2	The data	15
2.3	The model and hypothesis	15
2.4	Matrix formulation	16
2.5	Matrix formulation	17
2.6	Estimating equations for α and β	17
2.7	Restricted maximum likelihood estimates of variances	17
2.8	Likelihood ratio test of hypothesis	18
2.9	Estimate with all effects fixed	18
2.10	The ss-formulas in the ANOVA tables	19
2.11	SS computed by 'hand'	19
2.12	Maximum likelihood estimates	19
2.13	Compare estimates of fixed effects	19
2.14	Compare variance and SS estimates	19
2.15	Compare results from assuming fixed vs random effects	20

2.16	Discussion	20
2.17	The model and hypothesis	20
3	Handball	20
3.1	The data	20
3.2	The model and hypothesis	21
3.3	Matrix formulation	23
3.4	Anova tables, tests and estimates of variance components	26
3.5	REML estimates	26
3.6	Estimating equations and solutions	27
3.7	ML estimates	28
3.8	LRT of hypothesis	28
3.9	Estimates with all effects fixed	28
3.10	Compare estimates of fixed effects REML vs ML	30
3.11	Compare variance and SS estimates	30
3.12	Compare results from assuming fixed vs random effects	31
3.13	T-test and comparison	31
3.14	Equality of F-tests in the full model and the Δ -model	32
3.15	Discussion	33
4	Machines	33
4.1	Discussion	33
4.2	The data	33
4.3	The model and hypothesis	33
4.4	Matrix formulation	34
4.5	Anova tables, tests and estimates of variance components	35
4.6	Estimating equations and solutions	36
4.7	REML estimates	37
4.8	LRT of hypothesis	37
4.9	Estimates with all effects fixed	38
4.10	The SS-formulas in the ANOVA tables	38
4.11	SS computed by 'hand'	39
4.12	Maximum likelihood estimates	40
4.13	Compare variance and SS estimates	40

4.14	Compare variance and SS estimates	41
5	Rail	42
5.1	The Data	42
5.2	The Model and Hypothesis	43
5.3	Matrix formulation	43
5.4	Anova tables, tests and estimates of variance components.	44
5.5	Estimating equations (for mu and beta) and solutions.	44
5.6	REML estimates (of variances).	45
5.7	LRT of hypothesis	45
5.8	Estimates with all effects fixed	46
5.9	The SS - formulas in the ANOVA table	46
5.10	SS computed "by hand"	46
5.11	ML estimates	47
5.12	Compare estimates of fixed effects [sl30 vs sl60]	48
5.13	Compare variance and SS estimates [sl25, sl 35, sl55, sl60]	48
5.14	Compare results from assuming fixed vs random effects	49
5.15	Discussion	49

1 Irrigation

1.1 The Data

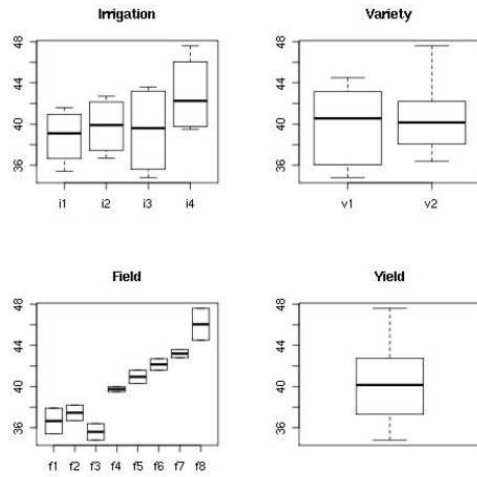


Figure 1: Boxplots of the yield group according to each of the three effects: irrigation, variety and field. The last boxplot shows the yield total.

The objective is to determine the effects of four different irrigation methods and two crop varieties in an agricultural field trial. Eight fields are available, but we can only apply one type of irrigation to each field. We can however split each field in two and plant different varieties in each half. The irrigation method is the whole plot factor, which should be assigned randomly to each field. Within each field the variety is assigned randomly. Irrigation i1 is used on fields f1 and f5, irrigation i2 is used on fields f2 and f6, irrigation i3 is used on fields f3 and f7 and irrigation i4 is used on fields f4 and f8. In each field variety v1 is planted in half of the field and variety v2 is planted in the other half of the field.

Boxplots of the yield grouped according to irrigation, variety and field can be seen in the figure. There seems to be difference between some fields and also between i4 and the other irrigation methods.

Table I shows the data.

1.2 The Model and Hypothesis

$$y_{fiv} = \mu + \mathbf{u}_i + \mathbf{v}_v + (\mathbf{tv})_{iv} + \phi_f + \epsilon_{fiv}$$

The irrigation and variety are obviously fixed effects since we are testing exactly those types. The fields on the other hand are random, since the soil and weather for example can not be chosen exactly. There is also interaction between irrigation and variety and also between field and variety. We can not assume that there is interaction between field and irrigation since each field only has one type of irrigation. The irrigation/variety interaction is fixed but the field/variety interaction is random. We therefore propose the model

$$y_{fiv} = \mu + \mathbf{u}_i + \mathbf{v}_v + (\mathbf{tv})_{iv} + \phi_f + (\phi\mathbf{v})_{fv} + \epsilon_{fiv}$$

where μ is the overall mean, $\mathbf{u}_i, \mathbf{v}_v, (\mathbf{tv})_{iv}$ are the fixed effects which represent the deviation from the mean due to irrigation, variety and their interaction and $\phi_f, (\phi\mathbf{v})_{fv}$ are random variables which represent the

Table 1: The Data

Field	Irrigation	Variety	Yield
f1	i1	v1	35.4
f1	i1	v2	37.9
f2	i2	v1	36.7
f2	i2	v2	38.2
f3	i3	v1	34.8
f3	i3	v2	36.4
f4	i4	v1	39.5
f4	i4	v2	40.0
f5	i1	v1	41.6
f5	i1	v2	40.3
f6	i2	v1	42.7
f6	i2	v2	41.6
f7	i3	v1	43.6
f7	i3	v2	42.8
f8	i4	v1	44.5
f8	i4	v2	47.6

deviation from the mean due to field and the interaction of field and variety. The random variable ε_{fiv} represents the deviation of the yield of variety v , irrigated with irrigation i on field f . Here $\phi_f \sim n(0, \sigma_\phi^2)$, $(\phi v)_{fv} \sim n(0, \sigma_{(\phi v)}^2)$, $\varepsilon_{fiv} \sim n(0, \sigma^2)$.

However, since we only have one observation per variety within each field, we must use a simpler model which does not account for the variety/field interaction. This model becomes

$$y_{fiv} = \mu + \mathbf{v}_i + \mathbf{v}_v + (\mathbf{tv})_{iv} + \phi_f + \varepsilon_{fiv}$$

where we have $\varepsilon_{fiv} \sim n(0, \sigma^2)$, $\phi_f \sim n(0, \sigma_\phi^2)$. We also have the restrictions

$$\sum_i \mathbf{v}_i = \sum_v \mathbf{v}_v = \sum_i (\mathbf{tv})_{iv} = \sum_v (\mathbf{tv})_{iv} = 0.$$

In the second part of the analysis we will assume that ϕ_f is also a fixed effect.

Null hypothesis are

$$H_{0(iv)} : (\mathbf{tv})_{iv} = 0, \quad H_{0i} : \mathbf{v}_i = 0, \quad H_{0v} : \mathbf{v}_v = 0, \quad H_{0\phi} : \phi_f = 0.$$

and we will find estimates for the variances σ^2, σ_ϕ^2 and also for the parameters in our model.

1.3 Matrix formulation

We want to find matrices X, Z such that

$$\mathbf{y} = X\boldsymbol{\beta} + Z\mathbf{u} + \boldsymbol{\varepsilon}$$

where β contains the fixed effects/parameters and \mathbf{u} contains the random effects. What we get is the following

$$\begin{pmatrix} y_{111} \\ y_{112} \\ y_{221} \\ y_{222} \\ y_{331} \\ y_{332} \\ y_{441} \\ y_{442} \\ y_{511} \\ y_{512} \\ y_{621} \\ y_{622} \\ y_{731} \\ y_{732} \\ y_{841} \\ y_{842} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \iota_1 \\ \iota_2 \\ \iota_3 \\ \iota_4 \\ \nu_1 \\ \nu_2 \\ (\iota\nu)_{11} \\ (\iota\nu)_{12} \\ (\iota\nu)_{21} \\ (\iota\nu)_{22} \\ (\iota\nu)_{31} \\ (\iota\nu)_{32} \\ (\iota\nu)_{41} \\ (\iota\nu)_{42} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & \ddots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \\ \phi_6 \\ \phi_7 \\ \phi_8 \end{pmatrix} + \varepsilon.$$

Because of the restrictions

$$\sum_i \iota_i = \sum_v \nu_v = \sum_i (\iota\nu)_{iv} = \sum_v (\iota\nu)_{iv} = 0$$

we can reduce the X matrix and write

$$\begin{pmatrix} y_{111} \\ y_{112} \\ y_{221} \\ y_{222} \\ y_{331} \\ y_{332} \\ y_{441} \\ y_{442} \\ y_{511} \\ y_{512} \\ y_{621} \\ y_{622} \\ y_{731} \\ y_{732} \\ y_{841} \\ y_{842} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \iota_2 \\ \iota_3 \\ \iota_4 \\ \nu_2 \\ (\iota\nu)_{22} \\ (\iota\nu)_{32} \\ (\iota\nu)_{42} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & \ddots & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \\ \phi_6 \\ \phi_7 \\ \phi_8 \end{pmatrix} + \varepsilon.$$

In further calculations X, Z, β, \mathbf{u} will be in the reduced form.

1.4 Anova tables, tests and estimates of variance components.

We set up the ANOVA table for the model. It is divided into two parts. The upper part takes care of the whole-plot, that is the irrigation within each field. The lower part takes care of the split-plot, that is, the variety within each irrigation. We have a total of 16 observations which makes 15 degrees of freedom. The sums of squares are calculated using the formulas

$$\begin{aligned}
 SSI &= 2V \sum_i (\bar{y}_{\cdot i} - \bar{y}_{\dots})^2 \\
 SSWPE &= 2 \sum_i \sum_f (\bar{y}_{f\cdot} - \bar{y}_{\cdot i})^2 \\
 SSWPT &= 2 \sum_f (\bar{y}_{f\cdot} - \bar{y}_{\dots})^2 \\
 SSV &= 2I \sum_v (\bar{y}_{\cdot v} - \bar{y}_{\dots})^2 \\
 SSIV &= 2 \sum_i \sum_v (\bar{y}_{vi} - \bar{y}_{\cdot v} - \bar{y}_{\cdot i} + \bar{y}_{\dots})^2 \\
 SSE &= \sum_f \sum_i \sum_v (\bar{y}_{vi} - \bar{y}_{f\cdot} - \bar{y}_{\cdot v} + \bar{y}_{fvi})^2 \\
 SST &= \sum_f \sum_v \sum_i (\bar{y}_{fvi} - \bar{y}_{\dots})^2
 \end{aligned}$$

where $i = 1, 2, 3, 4, v = 1, 2$ and $f = 1, \dots, 8$. We also calculate the expected values of the mean sums of squares and the results can be seen in table 2.

Table 2: ANOVA-table in the mixed model.

Source	DF	SS	MSS	E[MSS]
Irrigation	$I - 1$	SSI	$MSI = \frac{SSI}{(I-1)}$	$\frac{2V}{I-1} \sum_i v_i^2 + \frac{2V}{I-1} \sum_i (\bar{t}\bar{v})_i^2 + \sigma^2$
W.P.E.	$F - I$	SSWPE	$MSWPE = \frac{SSWPE}{F-I}$	$V\sigma_\phi^2 + \sigma^2$
W.P.T.	$F - 1$	SSWPT		
Variety	$V - 1$	SSV	$MSV = \frac{SSV}{(V-1)}$	$\frac{2I}{V-1} \sum_v v_v^2 + \frac{2I}{V-1} \sum_v (\bar{t}\bar{v})_v^2 + \sigma^2$
IV	$(V - 1)(I - 1)$	SSIV	$MSIV = \frac{SSIV}{(V-1)(I-1)}$	$\frac{2}{(I-1)(V-1)} \sum_i \sum_v (\bar{t}\bar{v})_{iv}^2 + \sigma^2$
S.P.E.	$I(V - 1)$	SSE	$MSE = \frac{SSE}{I(V-1)}$	σ^2
Total	$2VI - 1$	SST		

We can now test our null-hypothesis. We start by testing if the interaction between irrigation and variety is significant, the hypothesis $H_{0(tv)}$. We do that by calculating the ratio

$$MSIV/MSE$$

and comparing it to the appropriate F-value. We can now test the hypothesis H_{0t} and H_{0v} which correspond to effect of irrigation and variety. It is done by computing the ratios

$$MSV/MSE, \quad MSI/MSE$$

respectively. Now we have found out if each of the fixed effects are significant and we move on to the random effect. To see if the random effect, field (null-hypothesis $H_{0\phi}$), is significant we compute

$$MSWPE/MSE$$

Table 3: ANOVA-table of mixed model with numerical values.

Source of variation	DF	Sum of Squares	Mean Square	Ratio	Contrast
Irrigation	3	40.190	13.397	0.388	$< F_{3,4} = 6.59$
W.P.E.	4	138.030	34.507	16.37	$> F_{4,4} = 6.39$
W.P.T.	7	178.22			
Variety	1	2.250	2.250	1.068	$< F_{1,4} = 7.71$
IV	3	1.550	0.517	0.2452	$< F_{3,4} = 6.59$
Split Plot Error	4	8.43	2.1075		
Total	15	190.45			

and compare to the right F-value. If a computed ratio is lower than the appropriate F-value the effect is not significant and we can not reject the corresponding null-hypothesis. If on the other hand, the ratio is greater than the F-value we reject the null hypotheses that all levels of a single effect are the same and say that there is infact difference between the levels.

We can use the ANOVA-table to calculate estimates for σ^2 and σ_ϕ^2 after we check which effects are significant. We get the estimate for σ^2 directly from the table, that is

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

After we have found the estimate for σ^2 , and if the effect of fields is significant, we can find $\hat{\sigma}_\phi^2$ using the equation

$$MSF = V\sigma_\phi^2 + \sigma^2.$$

We are now ready to look at the ANOVA-table with all the numerical values, see table 3.

As we can see from the table the effects of irrigation, variety and their interaction are not significant. However, the effect of field is significant. In null-hypothesis this means that we reject $H_{0\phi}$ but can not reject $H_{0I}, H_{0V}, H_{0(IV)}$.

Finally we get the variance estimates

$$\hat{\sigma}^2 = 2.1075, \quad \hat{\sigma}_\phi^2 = 16.2.$$

In the above estimates we used 95% significance level.

1.5 Estimating equations (for \mathbf{u} and β) and solutions.

We have the equation

$$\mathbf{y} = X\beta + Z\mathbf{u} + \varepsilon$$

from slide 20 and we want to estimate β, \mathbf{u} using BLUP.

We find matrices G, R so that

$$\text{Var}[\mathbf{y}] = \sigma^2 \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} = \begin{bmatrix} \sigma_\phi^2/\sigma^2 \mathbb{I}_8 & 0 \\ 0 & \mathbb{I}_{16} \end{bmatrix}$$

and we thus have $G = \sigma_\phi^2/\sigma^2 \mathbb{I}_8$ and $R = \mathbb{I}_{16}$. Now we can estimate β, \mathbf{u} with the equations

$$\hat{\beta} = (X'(R + ZGZ')^{-1}X)^{-1}X'(R + ZGZ')^{-1}\mathbf{y}$$

$$\hat{u} = (Z'R^{-1}Z + G^{-1})^{-1}(Z'R^{-1} - Z'R^{-1}X(X'(R + ZGZ')^{-1}X)^{-1}X'(R + ZGZ')^{-1})y$$

These equations give us the BLUP-estimates

$$\hat{\beta} = \begin{pmatrix} \mu \\ i_2 \\ i_3 \\ i_4 \\ v_2 \\ (iv)_{22} \\ (iv)_{32} \\ (iv)_{42} \end{pmatrix} = \begin{pmatrix} 38.5 \\ 1.2 \\ 0.7 \\ 3.5 \\ 0.6 \\ -0.4 \\ -0.2 \\ 1.2 \end{pmatrix}, \quad \hat{u} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \end{pmatrix} = \begin{pmatrix} -2.018692 \\ -2.206477 \\ -3.567920 \\ -2.957618 \\ 2.018692 \\ 2.206477 \\ 3.567920 \\ 2.957618 \end{pmatrix}.$$

1.6 REML estimates (of variances).

We want to use REML to estimate the variances. We start by finding a matrix K such that

$$KX \equiv 0.$$

We can use QR -decomposition and write

$$X = QR = [q_1, \dots, q_p, \dots, q_n]R$$

where R is $n \times p$ matrix like X and Q is $n \times n$ matrix and the vectors q_1, \dots, q_n form an orthonormal basis. Then it holds that

$$[q_{p+1}, \dots, q_n]'X = 0.$$

so we let $K = [q_{p+1}, \dots, q_n]$.

We look at

$$\tilde{y} = Ky \sim n(K\Sigma_y K')$$

where $\Sigma_y = \sigma^2(ZGZ' + R)$, we denote $A = ZGZ' + R$. The likelihood function of \tilde{y} is

$$L(\theta) = \frac{1}{(2\pi)^{n/2} |K\Sigma_y K'|^{1/2}} e^{-\frac{1}{2}\tilde{y}'(K\Sigma_y K')^{-1}\tilde{y}}$$

and we find $\hat{\theta} = (\hat{\sigma}^2, \hat{\sigma}_\phi^2)$ that solves

$$\max_{\theta} L(\theta) \quad \text{or} \quad \max_{\theta} \ln(L(\theta)).$$

The function we need to maximize with respect to σ^2 and σ_ϕ^2 is

$$\ln(L(\theta)) = -\frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |KAK'| - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \tilde{y}'(KAK')^{-1}\tilde{y}$$

and some calculations give us the maximum value of $\ln(L)$

$$\max_{\theta} (\ln(L(\theta))) = -27.66671$$

and the REML-estimates of σ^2, σ_ϕ^2

$$\hat{\theta} = \begin{pmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_\phi^2 \end{pmatrix} = \begin{pmatrix} 0.7024995 \\ 16.9025017 \end{pmatrix}.$$

1.7 Estimate with all effects fixed.

If we assume that all effects are fixed the only random variable is ε_{fiv} and we only need to estimate σ^2 , not σ_ϕ^2 . The model is still

$$y_{fiv} = \mu + i_i + v_v + (iv)_{iv} + \phi_f + \varepsilon_{fiv}$$

but since we only have 16 observations which is not a full model, we can not estimate all the parameters. We have a total of $16 - 1 = 15$ degrees of freedom to work with, which means that we must either exclude the effect of irrigation or the effect of 3 fields. We will first exclude the effects of three fields and then the effects of the irrigations. We can write the model on matrix form as

$$\mathbf{y} = \tilde{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ i_2 \\ i_3 \\ i_4 \\ v_2 \\ (iv)_{22} \\ (iv)_{32} \\ (iv)_{42} \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \\ \phi_6 \\ \phi_7 \\ \phi_8 \end{pmatrix} + \boldsymbol{\varepsilon}$$

and use regular methods to estimate $\boldsymbol{\beta}$. We get

$$\hat{\boldsymbol{\beta}} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\mathbf{y} = \begin{pmatrix} \mu \\ i_2 \\ i_3 \\ i_4 \\ v_2 \\ (iv)_{22} \\ (iv)_{32} \\ (iv)_{42} \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \end{pmatrix} = \begin{pmatrix} 36.35 \\ 5.70 \\ 6.65 \\ 8.80 \\ 0.60 \\ -0.40 \\ -0.20 \\ 1.20 \\ -4.70 \\ -7.60 \\ -6.30 \\ 4.30 \end{pmatrix}.$$

The second estimation where we leave out the effect of irrigation is

$$\hat{\boldsymbol{\beta}} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\mathbf{y} = \begin{pmatrix} \mu \\ v_2 \\ (iv)_{22} \\ (iv)_{32} \\ (iv)_{42} \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \\ \phi_6 \\ \phi_7 \\ \phi_8 \end{pmatrix} = \begin{pmatrix} 36.35 \\ 0.60 \\ -0.40 \\ -0.20 \\ 1.20 \\ 1.00 \\ -0.95 \\ 2.50 \\ 4.30 \\ 5.70 \\ 6.65 \\ 8.80 \end{pmatrix}.$$

We will use the first model, the one without three fields, in the subsequent slides.

1.8 The SS-formulas in the ANOVA tables.

We can not estimate all levels in the factors and must leave three out. We chose to leave out three levels of the fields and therefore we have 5 levels of the fields and the degrees of freedom are 4. We can use MSE to calculate whether each of the factors is significant or not and therefore we can simplify the ANOVA table from slide 25 and it looks like table 4. The sums of squares are

Table 4: ANOVA-table for the fixed model.

Source of variation	DF	Sum of Squares	Mean Square	Ratio F-test
Field	$F - 1$	SSF	$MSF = \frac{SSF}{F-1}$	MSF/MSE
Irrigation	$I - 1$	SSI	$MSI = \frac{SSI}{I-1}$	MSI/MSE
Variety	$V - 1$	SSV	$MSV = \frac{SSV}{V-1}$	MSV/MSE
IV	$(V - 1)(I - 1)$	SSIV	$MSIV = \frac{SSIV}{(V-1)(I-1)}$	$MSIV/MSE$
Error	$VI + I - F$	SSE	$MSE = \frac{SSE}{VI+I-F}$	
Total	$2VI - 1$	SST		

$$SSI = 2V \sum_i (\bar{y}_{\cdot i} - \bar{y}_{\dots})^2$$

$$SSF = 2 \sum_i \sum_f (\bar{y}_{f \cdot} - \bar{y}_{\cdot i})^2$$

$$SSV = 2I \sum_v (\bar{y}_{\cdot v} - \bar{y}_{\dots})^2$$

$$SSIV = 2 \sum_i \sum_v (\bar{y}_{vi} - \bar{y}_{\cdot v} - \bar{y}_{\cdot i} + \bar{y}_{\dots})^2$$

$$SSE = \sum_f \sum_i \sum_v (\bar{y}_{\cdot i} - \bar{y}_{f \cdot} - \bar{y}_{\cdot v} + \bar{y}_{fvi})^2$$

$$SST = \sum_f \sum_v \sum_i (\bar{y}_{fvi} - \bar{y}_{\dots})^2$$

1.9 SS computed by hand.

We get, with some help from ANOVA tables in R:

$$SSF = 138.03, \quad SSI = 40.190, \quad SSV = 2.25, \quad SSIV = 1.55$$

$$SSE = 8.43, \quad SST = 190.45$$

Table 5 shows the ANOVA table.

Notice that the effect of fields only has four degrees of freedom instead of seven. This is because we only have 16 observations. Therefore we can not take the effect of all fields into account and we chose to exclude the last three fields, f_6, f_7, f_8 .

The table shows that only the effect of fields is significant, that is, we reject that all the fields are the same. This does not come as a surprise if we look at the boxplot of the fields and this is also the same result as in the mixed model.

As before MSE gives us an estimate of σ^2 , that is

$$\hat{\sigma}^2 = MSE = 2.1075.$$

An ANOVA-table that leaves out the irrigations instead of three fields, gives the same result.

Table 5: The ANOVA table for the fixed model with numerical values.

Source of variation	DF	Sum of Squares	Mean Square	Ratio F-test	Contrast F-value
Field	4	138.03	34.507	16.37	>6.39
Irrigation	3	40.19	13.397	6.36	<6.59
Variety	1	2.25	2.25	1.07	<7.71
IV	3	1.55	0.517	0.25	<6.59
Error	4	8.43	2.1075		
Total	15	190.45			

1.10 ML estimates.

The ML estimate of β is the same as on slide 45, that is

$$\hat{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\mathbf{y} = \begin{pmatrix} \mu \\ i_2 \\ i_3 \\ i_4 \\ v_2 \\ (iv)_{22} \\ (iv)_{32} \\ (iv)_{42} \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \end{pmatrix} = \begin{pmatrix} 36.35 \\ 5.70 \\ 6.65 \\ 8.80 \\ 0.60 \\ -0.40 \\ -0.20 \\ 1.20 \\ -4.70 \\ -7.60 \\ -6.30 \\ 4.30 \end{pmatrix}.$$

The estimate we get for σ^2 from the table is not the ML estimate, but it is easily computed since

$$\begin{aligned} \hat{\sigma}_{MLE}^2 &= \frac{1}{N} \sum_i \sum_v \sum_f (y_{fiv} - \hat{y}_{fiv})^2 \\ &= \frac{1}{16} 8.43 = 0.527 \end{aligned}$$

where \hat{y}_{fiv} are our estimated values calculated according to the model and $\hat{\beta}$.

This estimate is much lower than the one we got from the ANOVA-table.

1.11 Compare estimates of fixed effects (sl 30 vs. sl 60)

In both models we got that the effects of irrigation, variety and their interaction were not significant. That means that there is no difference between the levels of those effects, and we assume that their estimates are equal to zero. We still calculated their estimates.

In the mixed model we got the following estimates for the fixed effects $\mu, \iota, v, (\iota v)$

$$\hat{\beta} = \begin{pmatrix} \mu \\ \iota_2 \\ \iota_3 \\ \iota_4 \\ v_2 \\ (\iota v)_{22} \\ (\iota v)_{32} \\ (\iota v)_{42} \end{pmatrix} = \begin{pmatrix} 38.5 \\ 1.2 \\ 0.7 \\ 3.5 \\ 0.6 \\ -0.4 \\ -0.2 \\ 1.2 \end{pmatrix}$$

and in the fixed effect model we got the estimates

$$\hat{\beta} = \begin{pmatrix} \mu \\ \iota_2 \\ \iota_3 \\ \iota_4 \\ \nu_2 \\ (\iota\nu)_{22} \\ (\iota\nu)_{32} \\ (\iota\nu)_{42} \\ \phi_2 \\ \phi_3 \\ \phi_4 \\ \phi_5 \end{pmatrix} = \begin{pmatrix} 36.35 \\ 5.70 \\ 6.65 \\ 8.80 \\ 0.60 \\ -0.40 \\ -0.20 \\ 1.20 \\ -4.70 \\ -7.60 \\ -6.30 \\ 4.30 \end{pmatrix}.$$

As can be seen the total mean estimates μ are close to each other, they are both positive and in the same range. Although the estimates of ι_i are all positive in both cases, they are different, the estimates in the mixed model being much lower. The estimates of ν_2 are the same in both cases and the same can be said about the $(\iota\nu)_{ij}$ interaction estimates. In the fixed effects model ϕ_2, \dots, ϕ_5 are also estimated, which is not the case in the mixed model.

We see that the estimates of the effects that we have concluded to be zero usually turn out to be close to zero, and if we would calculate their confidence intervals zero would be an element in each of them.

On the other hand, the effect of field is significant, so those estimates are different from zero.

1.12 Compare variance and SS estimates (sl 25, sl 35, sl 55, sl 60)

We start by comparing the variances. On slide 25 we got estimates of σ^2 and σ_ϕ^2 from the ANOVA-table and they were

$$\hat{\sigma}^2 = 2.1075, \quad \hat{\sigma}_\phi^2 = 16.2$$

and on slide 35 we used the REML method to find MLE for the two sigmas, and the estimates were

$$\hat{\sigma}^2 = 0.702, \quad \hat{\sigma}_\phi^2 = 16.90.$$

As can be seen, the estimates for σ^2 are quite different but the estimates for σ_ϕ^2 are almost the same.

On slide 55 we got an estimate for σ^2 from the ANOVA-table and it was

$$\hat{\sigma}^2 = 2.1075$$

which is the same as in the mixed model. The MLE of σ^2 in the fixed model was

$$\hat{\sigma}^2 = 0.527$$

as can be seen on slide 60. This is much lower than the estimate of σ^2 on slide 55.

We will now compare the SS-estimates for the models. In the mixed model, we don't have a sum for the fields, they are included in the residuals of the whole plot. In the fixed effects model we do not need those residuals, and therefore we can use that sum for the fields. The sums of squares for irrigation, variety and their interaction are the same in both models and also the error term.

1.13 Compare results from assuming fixed vs. random effects.

In the mixed effect model we assumed that the effect of fields was random and that the other effects were fixed. Therefore we had two variances σ^2 and σ_ϕ^2 to think about. We could estimate all the levels of the

Level	Estimate in mixed model	Difference $f_i - f_1$	Estimate in fixed model
f_1	-2.018692	0	0
f_2	-2.206477	0.18	1.00
f_3	-3.567920	1.55	-0.95
f_4	-2.957618	0.94	2.50
f_5	2.018692	4.04	4.30
f_6	2.206477	4.23	5.70
f_7	3.567920	5.59	6.65
f_8	2.957618	4.98	8.80

fields and also all the levels of the fixed effects. We estimated the sigmas with two different methods, and they gave the same estimate for σ_ϕ^2 but different estimates for σ^2 .

In the fixed effects model we assumed that all effects were fixed. Due to lack of observations we could not estimate all the levels of all effects, and we chose to leave out effects of three fields. All other levels were estimated. As before we estimated σ^2 in two different ways, and the estimates gave different results.

In both models, only the effect of fields was significant. We can therefore assume that all the levels of the other factors equal zero. We got the same ANOVA-estimate for σ^2 in both models but the ML estimates were quite different.

1.14 Discussion

We have now gone through two approaches to analyse our problem and build a model for it. First we built a model with both random and fixed effects and then we built a model with only fixed effects.

We proposed the equation

$$y_{fiv} = \mu + \mathbf{v}_i + \mathbf{v}_v + (\mathbf{1v})_{iv} + \phi_f + \varepsilon_{fiv}.$$

In both approaches the effects of irrigation, variety and their interaction were non-significant so we can reduce the model and write it as

$$y_{fiv} = \mu + \phi_f + \varepsilon_{fiv}$$

where ϕ_f is either a random or fixed effect. We estimated the levels of ϕ_f and the results from each model can be seen in the table. First come the estimates in the mixed model which we obtained using BLUP. The next column shows the difference $f_i - f_1$ for $i = 1, \dots, 8$ in the first column. The last column shows the estimates from the fixed model using ML, and there $f_1 = 0$ and the other f_i represent the difference between f_1 and f_i . As can be seen in the table the difference between the estimates in the two models is not much.

In the fixed model we were able to estimate all levels of all effects but in the mixed model we had to leave out three levels because we had so few observations. This problem with the number of observations and not being able to estimate all levels makes the fixed effects model worse than the mixed model. More things could be derived from the mixed model, which makes it more accurate. It is also natural to think of the fields as a random effect since the soil and weather can not be controlled. Some of the fields may have gotten more rain and others could have gotten warmer weather during the experiment, for example.

We therefore propose the mixed model (which is actually a random model)

$$y_{fiv} = \mu + \phi_f + \varepsilon_{fiv}$$

where $\phi_f \sim n(0, \sigma_\phi^2)$ and $\varepsilon_{fiv} \sim n(0, \sigma^2)$.

We estimated the variances using two methods, the estimates from the ANOVA-tables are

$$\hat{\sigma}^2 = 2.1075, \quad \hat{\sigma}_\phi^2 = 16.2$$

and the ML estimates are

$$\hat{\sigma}^2 = 0.702, \quad \hat{\sigma}_\beta^2 = 16.9.$$

The estimates of $\hat{\sigma}_\beta^2$ are almost the same, and it should not matter much which one we use. On the other hand, there is a big difference between the $\hat{\sigma}^2$ estimates. It is not obvious which one to use.

2 Barley seed example

2.1 Anova tables, tests and estimates of variance components

The relevant sums of squares for the model described in 14 are:

$$\begin{aligned} SSA &= JK \sum_{i=1}^I (\bar{y}_{i..} - \bar{y}_{...})^2 \\ SSB &= K \sum_{i=1}^I \sum_{j=1}^J (\bar{y}_{ij.} - \bar{y}_{i..})^2 \\ SSE &= \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (y_{ijk} - \bar{y}_{ij.})^2 \end{aligned} \quad (1)$$

The expected sums of squares are:

$$\begin{aligned} E[SSA] &= JK \left(\sum_{i=1}^I (\alpha_i - \bar{\alpha}.)^2 + E[(\bar{\beta}_{.(i)} - \bar{\beta}_{.(.)})^2] + E[(\bar{\epsilon}_{i..} - \bar{\epsilon}_{...})^2] \right) \\ &= JK \sum_{i=1}^I (\alpha_i - \bar{\alpha}.)^2 + K(I-1)\sigma_\beta^2 + (I-1)\sigma^2 \\ E[SSB] &= K(E[(\bar{\beta}_{j(i)} - \bar{\beta}_{.(i)})^2] + E[(\bar{\epsilon}_{ij.} - \bar{\epsilon}_{i..})^2]) \\ &= KI(J-1)\sigma_\beta^2 + I(J-1)\sigma^2 \\ E[SSE] &= IJ(K-1)\sigma^2 \end{aligned} \quad (2)$$

so the ANOVA table for this problem is:

Source	SS	DF	MSS	E[MSS]
Treatment	SSA	$I-1$	$\frac{SSA}{I-1}$	$\frac{JK}{I-1} \sum_{i=1}^I (\alpha_i - \bar{\alpha}.)^2 + K\sigma_\beta^2 + \sigma^2$
Seedbag	SSB	$I(J-1)$	$\frac{SSB}{I(J-1)}$	$K\sigma_\beta^2 + \sigma^2$
Residuals	SSE	$IJ(K-1)$	$\frac{SSE}{IJ(K-1)}$	σ^2

To test our hypothesis from 16 we note that under the null hypothesis the ratio

$$F = \frac{MSA}{MSB}$$

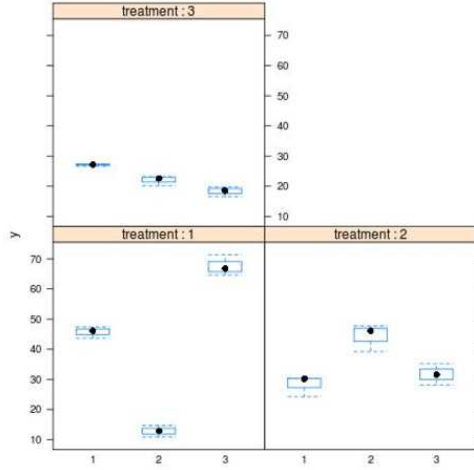
has an centred F-distribution. From table 4.5 we can calculate our estimate of F is 1.035 with a p-value

	Df	Sum Sq	Mean Sq
treatment	2.00	1758.72	879.36
treatment:seedgrp	6.00	5099.26	849.88
Residuals	18.00	137.87	7.66

Table 6: ANOVA table

of 0.589, with 2 and 6 degrees of freedom. MSE serves as an estimate for σ^2 and $\frac{MSB-MSE}{K}$ for σ_β^2 . Here $\hat{\sigma}^2 = 7.660$ and $\hat{\sigma}_\beta^2 = 280.739$.

2.2 The data



A barley seed test example. This originally came in as a request from an MSc student (Aðalheiður Einarsdóttir, advisor Ólafur Andrésson) in biology and resulted in a bit of discussion. The dataset describes the results from an experiment on barley and its response to different fertilizers. The design of the experiment is a nested design where the seedgroup (- bag) varies within the treatment levels. There were treatment levels, one group was given water, the second urea and the third water. For each treatment level three seedgroups were planted with three repetitions. The experimental results can be seen in figure 2 and table 7.

Figure 2: Boxplot of the response as a function of treatment and seedgroups.

2.3 The model and hypothesis

The treatment is obviously not random while the seedgroup within treatments could be considered to be a random effect. The model for the measurement is therefore

$$\begin{aligned}
 y_{ijk} &= \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk} \\
 i &= 1, \dots, I \\
 j &= 1, \dots, J \\
 k &= 1, \dots, K
 \end{aligned} \tag{3}$$

where μ is the overall mean, α_i is the effect of the i -th treatment and $\beta_{j(i)} \sim n(0, \sigma_\beta^2)$ is the effect of the seedgroup j within treatment i . The formulation of μ as the overall mean of y introduces the following restraints:

$$\sum_i \alpha_i = 0 \tag{4}$$

The hypothesis of interest is :

$$H_0 : \alpha_i = 0 \forall i \quad \text{vs.} \quad H_a : \exists i \text{ s.t. } \alpha_i \neq 0 \tag{5}$$

	y	treatment	seedgrp	st	s	t
1	46.10	1	1	11	1	1
2	47.40	1	1	11	1	1
3	43.60	1	1	11	1	1
4	14.70	1	2	21	2	1
5	12.90	1	2	21	2	1
6	10.90	1	2	21	2	1
7	71.30	1	3	31	3	1
8	66.80	1	3	31	3	1
9	64.60	1	3	31	3	1
10	30.40	2	1	12	1	2
11	30.20	2	1	12	1	2
12	24.30	2	1	12	1	2
13	46.10	2	2	22	2	2
14	47.70	2	2	22	2	2
15	39.20	2	2	22	2	2
16	35.20	2	3	32	3	2
17	31.60	2	3	32	3	2
18	28.20	2	3	32	3	2
19	26.70	3	1	13	1	3
20	27.20	3	1	13	1	3
21	27.40	3	1	13	1	3
22	23.20	3	2	23	2	3
23	22.60	3	2	23	2	3
24	20.20	3	2	23	2	3
25	19.80	3	3	33	3	3
26	18.60	3	3	33	3	3
27	16.60	3	3	33	3	3

Table 7: The Barley-dataset

2.4 Matrix formulation

In matrix form the model described in equation 14 becomes

$$y = X\alpha + Z\beta + \varepsilon \quad (6)$$

where $\varepsilon \sim n(0, \sigma^2 I_{27})$, $\beta \sim n(0, \sigma_\beta^2 I_9)$ and $\alpha = (\mu, \alpha_1, \alpha_2, \alpha_3)$ fixed. If we disregard the restriction in equation 15 we have that the matrices in 8 are

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 1 \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 1 \end{bmatrix} \quad (7)$$

To include the restraints in 15 we will here follow the approach in R and introduce $\mu' = \mu + \alpha_1$ and delete the column representing α_1 , i.e. the second column in X .

2.5 Matrix formulation

In matrix form the model described in equation 14 becomes

$$y = X\alpha + Z\beta + \varepsilon \quad (8)$$

where $\varepsilon \sim n(0, \sigma^2 I_{27})$, $\beta \sim n(0, \sigma_\beta^2 I_9)$ and $\alpha = (\mu, \alpha_1, \alpha_2, \alpha_3)$ fixed. If we disregard the restriction in equation 15 we have that the matrices in 8 are

$$X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 1 \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 1 \\ 0 & \dots & 0 & 1 \end{bmatrix} \quad (9)$$

To include the restraints in 15 we will here follow the approach in R and introduce $\mu' = \mu + \alpha_1$ and delete the column representing α_1 , i.e. the second column in X .

2.6 Estimating equations for α and β

From the matrix formulation above we can derive the BLUP estimates for α and β from the following formulae:

$$\begin{aligned} \hat{\alpha} &= (X^T(R + ZGZ^T)X)^{-1}X^T(R + ZGZ^T)^{-1}y \\ \hat{\beta} &= (Z^TR^{-1}Z + G^{-1})^{-1}(Z^TR^{-1} - Z^TR^{-1}X(X^T(R + ZGZ^T)^{-1}X)^{-1}X^T(R + ZGZ^T)^{-1})y \end{aligned} \quad (10)$$

where R and G solve

$$\text{Var}(y) = \sigma^2 \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} = \begin{bmatrix} \sigma_\beta^2/\sigma^2 I_{IJ} & 0 \\ 0 & I_{JK} \end{bmatrix} \quad (11)$$

and the BLUP estimates are

	alpha	beta_1	beta_2	beta_3
1	42.03	0.28	-2.21	1.93
2	-7.27	-0.49	0.72	-0.23
3	-19.56	0.35	-0.04	-0.31

2.7 Restricted maximum likelihood estimates of variances

To find the REML estimate for the variances we need to find an appropriate linear transformation to eliminate the fixed effects in 8. That find K such that

$$KX = 0$$

To find a possible K we can use a QR -decomposition for X , that is

$$X = QR$$

were R is an $n \times p$ matrix like X and Q is an $n \times n$ orthonormal matrix, that is the columns q_1, \dots, q_n of Q form an orthonormal basis. Furthermore, $K = [q_{p+1}, \dots, q_n]$, the matrix containing the last $n - p$ columns, has the property

$$KX = 0$$

Using this as our selected K we use it to transform our measurements and we see that

$$\tilde{\mathbf{y}} = K\mathbf{y} \sim n(0, K\Sigma_y K')$$

where $\Sigma_y = \sigma^2(ZGZ' + R)$. The problem now translates into finding the pair $(\sigma^2, \sigma_\beta^2)$ that maximises the log-likelihood function of $n(0, K\Sigma_y K')$. Using the nonlinear optimiser in R (nlm) with starting values close to the estimate which we found with the anova table the REML estimate is $\hat{\sigma}_{\beta, REML}^2 = 280.739$ and $\hat{\sigma}_{REML}^2 = 7.659$ which is exactly the same estimate as estimated with the ANOVA up to the third digit.

2.8 Likelihood ratio test of hypothesis

We will here test the hypothesis presented in 16 with a likelihood ratio test. The model in 14 will be compared with

$$y_{ijk} = \mu + \beta_{j(i)} + \varepsilon_{ijk}$$

Here the test statistic is

$$D = -2\log(L_{full}) - \log(L_{red})$$

where $D \sim \chi_d^2$ and d is the difference in degrees of freedom between the models. We have $D = 2.192$ with $p = 0.334$.

2.9 Estimate with all effects fixed

Assuming that the seedgroup effects are constant we can estimate α and β with ordinary linear regression. The augmented X - matrix will include columns from the original X from earlier and Z excluding the second column from X and first, fourth and seventh from Z to be comparable with R . Equation 8 reduces to

$$y = X\alpha + \varepsilon \tag{12}$$

where

$$\alpha = \begin{bmatrix} \mu \\ \alpha_2 \\ \alpha_3 \\ \beta_{1(2)} \\ \beta_{1(3)} \\ \beta_{2(2)} \\ \beta_{2(3)} \\ \beta_{3(2)} \\ \beta_{3(3)} \end{bmatrix} \tag{13}$$

	alpha	beta_2	beta_3
1	45.70	-32.87	21.87
2	-17.40	16.03	3.37
3	-18.60	-5.10	-8.77

Table 8: Estimate with all effects fixed

2.10 The ss-formulas in the ANOVA tables

Same as from the mixed effects except for $\beta_{j(i)}$ is considered fixed and $\sum_{j=1}^J \beta_{j(i)} = 0$ for all i . Therefore the test statistic for the hypothesis in equation 16 is

$$F = \frac{MSA}{MSE}$$

The results can be seen in table 9.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
treatment	2	1758.72	879.36	114.80	0.0000
treatment:seedgrp	6	5099.26	849.88	110.96	0.0000
Residuals	18	137.87	7.66		

Table 9: ANOVA table for the model in efecq:model assuming all effects are fixed

2.11 SS computed by 'hand'

	y	sum_y_ij.	bar_y_ij.	sum_y_i..	bar_y_i..	diff_ij	diff_i
1:1	46.1 47.4 43.6	137.1	45.7	378.3	42.03	3.67	8.94
1:2	14.7 12.9 10.9	38.5	12.83	378.3	42.03	-29.2	8.94
1:3	71.3 66.8 64.6	202.7	67.57	378.3	42.03	25.53	8.94
2:1	30.4 30.2 24.3	84.9	28.3	312.9	34.77	-6.47	1.67
2:2	46.1 47.7 39.2	133	44.33	312.9	34.77	9.57	1.67
2:3	35.2 31.6 28.2	95	31.67	312.9	34.77	-3.1	1.67
3:1	26.7 27.2 27.4	81.3	27.1	202.3	22.48	4.62	-10.61
3:2	23.2 22.6 20.2	66	22	202.3	22.48	-0.48	-10.61
3:3	19.8 18.6 16.6	55	18.33	202.3	22.48	-4.14	-10.61

2.12 Maximum likelihood estimates

The maximum likelihood estimators for all effects of the model can be seen in table 8. The maximum likelihood estimate for σ is computed from the residuals from the fitted model

$$\hat{\sigma}_{MLE}^2 = \frac{1}{IJK} \sum_{i,j,k=1}^{I,J,K} (y_{ijk} - \hat{y}_{ijk})^2$$

where \hat{y}_{ijk} is the fitted value for y_{ijk} . The estimate is $\hat{\sigma}_{MLE} = 5.303$.

2.13 Compare estimates of fixed effects

The estimates of the fixed effects can be seen in table 10. We can see that the estimates differ somewhat, which can be attributed to how the seedgroup effect is handled in the different estimation processes.

2.14 Compare variance and SS estimates

We see that the estimates of σ^2 and σ_{β}^2 are identical when estimated with REML and ANOVA. The maximum likelihood estimator for σ^2 when assuming that the seed group effect is fixed is lower than when using REML. That however is made at the expense 4 more degrees of freedom being used.

	1	2	3
1	mu	45.7	42.03
2	alpha_2	-17.4	-7.27
3	alpha_3	-18.6	-19.56

Table 10: Comparison of estimates of α_i

2.15 Compare results from assuming fixed vs random effects

When assuming that the seed group effect is random the effects of the treatment is dwarfed by the seed group noise. When they are considered fixed the effects of the seed group influences greatly all comparisons.

2.16 Discussion

The results from this analysis is inconclusive as the effect of the seedgroup, as noted earlier, is such that the effects of the treatment cannot be measured with the current setup with a mixed effects models. A fixed effects model seems implausible as one would like to draw inferences on the treatment effects for all barley seedgroups. Therefore the number of seedbag should be increased to effectively determine the effects of the treatment.

2.17 The model and hypothesis

The treatment is obviously not random while the seedgroup within treatments could be considered to be a random effect. The model for the measurement is therefore

$$\begin{aligned}
 y_{ijk} &= \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk} \\
 i &= 1, \dots, I \\
 j &= 1, \dots, J \\
 k &= 1, \dots, K
 \end{aligned} \tag{14}$$

where μ is the overall mean, α_i is the effect of the i -th treatment and $\beta_{j(i)} \sim n(0, \sigma_\beta^2)$ is the effect of the seedgroup j within treatment i . The formulation of μ as the overall mean of y introduces the following restraints:

$$\sum_i \alpha_i = 0 \tag{15}$$

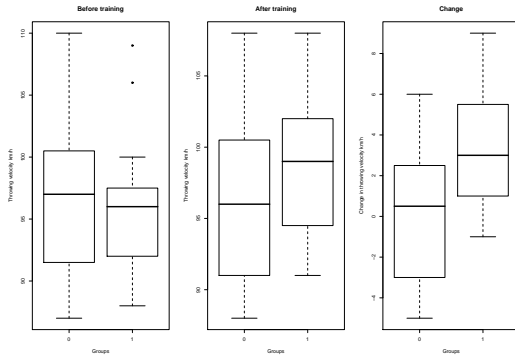
The hypothesis of interest is :

$$H_0 : \alpha_i = 0 \forall i \quad \text{vs.} \quad H_a : \exists i \text{ s.t. } \alpha_i \neq 0 \tag{16}$$

3 Handball

3.1 The data

This data comes from a BSc project in physical therapy. The projects name is: "Effects of training with a weighted implement on throwing velocity of male team handball players" and it's authors are Hildur Sólveig Sigurðardóttir and Tinna Jökulsdóttir who worked under the instructions of Dr. Árni Árnason. An abstract of the project can be seen at: <http://www2.hi.is/Apps/WebObjects/HI.woa/1/wa/dp?id=1024878wosid=kR8n2Adn>
40 handball players were randomly assign to two groups. Players in the research group trained three



times a week for 8 weeks with a heavy handball. Players in the control group did no extra training. Injured players were excluded from the research. 15 players in the research group and 12 players in the control group finished the research. Players throwing velocity was measured before and after the 8 week training period. We are interested in seeing if the extra training has an effect on the players throwing velocity.

3.2 The model and hypothesis

We have throwing velocity (y) as an response variable. We have 3 explanatory variables: The time (τ), the group factor (γ) and the players effect (ρ). We can there for suggest the model:

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \rho_{k(j)} + \varepsilon_{ijk}$$

- μ is the overall mean
- τ_i is the time effect and is fixed. $i = 1, 2$ stands for before and after training.
- γ_j is the group effect and is fixed. $j = 0, 1$ stands for no extra training and training with a heavy handball.
- $(\tau\gamma)_{ij}$ stands for the interaction between the time and group effect. It is also fixed.
- $\rho_{k(j)}$ stands for the players effect which is nested within the groups and is assumed to be random, independent and identically normally distributed with mean 0 and variance σ_ρ^2 . In short $\rho_{k(j)} \sim N(0, \sigma_\rho^2)$ iid.
- ε_{ijk} is the mesurment error and is also assumed to be random and iid. $\varepsilon_{ijk} \sim N(0, \sigma^2)$.

We also assume that $\rho_{k(j)}$ and ε_{ijk} are independent. We will refer to this model as the full model.

We can't estimate all these parameters and there for have to have some restrictions on them. The most common restrictions are.

$$\begin{aligned} \sum_i \tau_i &= 0 \\ \sum_j \gamma_j &= 0 \\ \sum_i (\tau\gamma)_{ij} &= \sum_j (\tau\gamma)_{ij} = 0 \end{aligned}$$

Because much of our analysis will take place in the R statistical package we will have different restrictions. Our restrictions will be:

$$\begin{aligned}\tau_1 &= 0 \\ \gamma_0 &= 0 \\ (\tau\gamma)_{10} &= (\tau\gamma)_{20} = (\tau\gamma)_{11} = 0\end{aligned}$$

Instead of having these restrictions on the parameters we could have stated the model as follows:

$$y_{ijk} = \mu_{ij} + \rho_{k(j)} + \varepsilon_{ijk}$$

where μ_{ij} represents the four first parameters in the model.

If we look at the change in performance for individual players we get a much simpler model.

$$\begin{aligned}\Delta y_{jk} &= y_{2jk} - y_{1jk} \\ &= \mu + \tau_2 + \gamma_j + (\tau\gamma)_{2j} + \rho_{k(j)} + \varepsilon_{2jk} - (\mu + \tau_1 + \gamma_j + (\tau\gamma)_{1j} + \rho_{k(j)} + \varepsilon_{1jk}) \\ &= \tau_2 - \tau_1 + (\tau\gamma)_{2j} - (\tau\gamma)_{1j} + \varepsilon_{2jk} - \varepsilon_{1jk} \\ &= \Delta\tau + \Delta(\tau\gamma)_j + \Delta\varepsilon_{jk}\end{aligned}$$

an one-way anova with fixed effects. Note that $\Delta\varepsilon_{jk} \sim N(0, 2\sigma^2)$. Our restrictions in this model are $\Delta(\tau\gamma)_0 = 0$. Like the full model we can state this model in a simpler form:

$$\Delta y_{jk} = \Delta\mu_j + \Delta\varepsilon_{jk}$$

We will refer to this model as the Δ -model.

The hypothesis we want to test are:

1. If the players are have any effect. That is can we drop the random effect from the model. Our null-hypothesis is therefor

$$H_1 : \sigma_p^2 = 0$$

2. Is the interaction significant?

$$H_2 : \text{All } (\tau\gamma)_{ij} = 0$$

We will show that testing this is the same as testing the group effect in the Δ -model. And it's known that testing the group effect in a one-way anova with only two groups is the same as doing a regular t-test for means assuming equal variance.

In the hypothesis above we are really asking if we can simplify the full model. We also like to test, if we find the interaction significant, which groups means are different.

3. Were the groups just different before the training period?

$$H_3 : \mu_{10} = \mu_{11}$$

With our restrictions this simplifies to

$$H_3 : \gamma_1 = 0$$

4. Are the groups different after the training period?

$$H_4 : \mu_{20} = \mu_{21}$$

simplifies to

$$H_4 : \gamma_1 + (\tau\gamma)_{21} = 0$$

5. Did throwing velocity change for players in the control group?

$$H_5 : \mu_{10} = \mu_{20}$$

in the Δ -model this would be $\Delta\mu_0 = 0$. This simplifies to

$$H_5 : \tau_2 = 0$$

6. Did throwing velocity change for players in the treatment group?

$$H_6 : \mu_{21} = \mu_{11}$$

We can also test this hypothesis with the Δ -model. Taking the restrictions into account this simplifies to

$$H_6 : \tau_2 + (\tau\gamma)_{21} = 0$$

3.3 Matrix formulation

If we state the full model in matrix form:

$$\mathbf{y} = X\boldsymbol{\beta} + Z\mathbf{u} + \boldsymbol{\varepsilon}$$

where β contains the fixed effects and \mathbf{u} the random effects. We have the following vectors and matrixes.

$$\begin{bmatrix} y_{1,0,1} \\ y_{1,0,2} \\ \vdots \\ y_{1,0,12} \\ y_{1,1,13} \\ y_{1,1,14} \\ \vdots \\ y_{1,1,27} \\ y_{2,0,1} \\ y_{2,0,2} \\ \vdots \\ y_{2,0,12} \\ y_{2,1,13} \\ y_{2,1,14} \\ \vdots \\ y_{2,1,27} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdot & 1 & 0 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & 1 & 0 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & 1 & 0 & 1 & \cdot & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdot & 1 & 0 & 0 & 1 & \cdot & 1 & 0 & \cdot \\ \cdot & 0 & 1 & 1 & 0 & \cdot & 0 & 1 & \cdot \\ \cdot & 0 & 1 & 1 & 0 & \cdot & 0 & 1 & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdot & 1 & 0 & \cdot & \cdot & 1 & 0 \\ \cdot & \cdot & \cdot & 0 & 1 & \cdot & \cdot & 0 & 1 \\ \cdot & \cdot & \cdot & 0 & 1 & \cdot & \cdot & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \gamma_0 \\ \gamma_1 \\ (\tau\gamma)_{10} \\ (\tau\gamma)_{11} \\ (\tau\gamma)_{20} \\ (\tau\gamma)_{21} \end{bmatrix} \\
 + \begin{bmatrix} 1 & 0 & \dots & \cdot & \cdot & \cdot & \dots & 0 \\ 0 & 1 & \ddots & \cdot & \cdot & \cdot & \ddots & \cdot \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \cdot & \cdot & \ddots & 1 & \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot & 1 & \cdot & \ddots & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot & 1 & \ddots & \cdot \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdot & \dots & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ 1 & 0 & \ddots & \cdot & \cdot & \cdot & \ddots & 0 & 0 \\ 0 & 1 & \ddots & \cdot & \cdot & \cdot & \ddots & \cdot & \cdot \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \cdot & \cdot & \ddots & 1 & \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & \ddots & \cdot & 1 & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot & 1 & \ddots & \cdot & \cdot \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & \cdot \\ 0 & \cdot & \dots & \cdot & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \begin{bmatrix} \rho_{1(0)} \\ \rho_{2(0)} \\ \vdots \\ \rho_{12(0)} \\ \rho_{13(1)} \\ \rho_{14(1)} \\ \vdots \\ \rho_{27(1)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,0,1} \\ \varepsilon_{1,0,2} \\ \vdots \\ \varepsilon_{1,0,12} \\ \varepsilon_{1,1,13} \\ \varepsilon_{1,1,14} \\ \vdots \\ \varepsilon_{1,1,27} \\ \varepsilon_{2,0,1} \\ \varepsilon_{2,0,2} \\ \vdots \\ \varepsilon_{2,0,12} \\ \varepsilon_{2,1,13} \\ \varepsilon_{2,1,14} \\ \vdots \\ \varepsilon_{2,1,27} \end{bmatrix}$$

Because of our restrictions the fixed effect part of the model is reduced to:

$$X\beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdot & 0 & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot \\ \vdots & \vdots & \vdots & \vdots \\ \cdot & 0 & 1 & \cdot \\ \cdot & 1 & 0 & \cdot \\ \cdot & 1 & 0 & \cdot \\ \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_2 \\ \gamma_1 \\ (\tau\gamma)_{21} \end{bmatrix}$$

We don't have any random effects for the Δ -model. We can therefore state it as:

$$\mathbf{y} = X_{\Delta}\beta_{\Delta} + \varepsilon_{\Delta}$$

$$\begin{bmatrix} \Delta y_{0,1} \\ \Delta y_{0,2} \\ \vdots \\ \Delta y_{0,12} \\ \Delta y_{1,13} \\ \Delta y_{1,14} \\ \vdots \\ \Delta y_{1,1,27} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ \cdot & 1 & 0 \\ \cdot & 0 & 1 \\ \cdot & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta\tau \\ \Delta(\tau\gamma)_0 \\ \Delta(\tau\gamma)_1 \end{bmatrix} + \begin{bmatrix} \Delta\varepsilon_{0,1} \\ \Delta\varepsilon_{0,2} \\ \vdots \\ \Delta\varepsilon_{0,12} \\ \Delta\varepsilon_{1,13} \\ \Delta\varepsilon_{1,14} \\ \vdots \\ \Delta\varepsilon_{1,1,27} \end{bmatrix}$$

Because of the restrictions the reduced form of the X_{Δ} -matrix and the β_{Δ} -vector are:

$$X_{\Delta}\beta_{\Delta} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ \vdots & \vdots \\ \cdot & 0 \\ \cdot & 1 \\ \cdot & 1 \\ \vdots & \vdots \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta\tau \\ \Delta(\tau\gamma)_1 \end{bmatrix}$$

3.4 Anova tables, tests and estimates of variance components

Let $N = 27$ represent the total number of players in the study. Let $n_0 = 12$ and $n_1 = 15$ represent the number of players in the control and the training group. We have the following sums of squares.

$$\begin{aligned}
 SST &= N \sum_i (\bar{y}_{i..} - \bar{y}...)^2 \\
 SSG &= 2 \sum_j n_j (\bar{y}_{.j.} - \bar{y}...)^2 \\
 SSTG &= \sum_i \sum_j n_j (\bar{y}_{ij.} + \bar{y}... - \bar{y}_{i..} - \bar{y}_{.j.})^2 \\
 SSP(G) &= 2 \sum_j \sum_k (\bar{y}_{.jk} - \bar{y}_{.j.})^2 \\
 SSE &= \sum_i \sum_j \sum_k (y_{ijk} + \bar{y}_{.j.} - \bar{y}_{ij.} - \bar{y}_{.jk})^2 \\
 SSTOT &= \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}...)^2
 \end{aligned}$$

We see from expected MS in table 11 that we can test H_1 with a F-test where $F = MSP(G)/MSE$. We

Table 11: Anova table: formulas

Source	df	SS	MS	$E[MS]$
Time	$I - 1$	SST	SST/df_T	$\sigma^2 + N \sum_i \tau_i^2$
Group	$J - 1$	SSG	SSG/df_G	$\sigma^2 + 2\sigma_p^2 + 2 \sum_j n_j \gamma_j^2$
Time \times Group	$(I - 1)(J - 1)$	SSTG	$SSTG/df_{TG}$	$\sigma^2 + \sum_i \sum_j n_j (\tau\gamma)_{ij}$
Players	$\sum_j (n_j - 1)$	SSP(G)	$SSP(G)/df_{P(G)}$	$\sigma^2 + 2\sigma_p^2$
Error	$\sum_j (n_j - 1)$	SSE	SSE/df_E	σ^2

can also test H_2 with a F-test, $F = MSTG/MSE$. The result from this tests are in table 12. We reject both of these hypothesis and therefor have to stick to the full model.

Table 12: Anova table: Estimates

Source	df	SS	MS	F	p
Time	1	44.46	44.46		
Group	1	9.07	9.07		
Time \times Group	1	32.38	32.38	5.71	0.0247
Players	25	1825.93	73.04	12.89	$6.25 \cdot 10^9$
Error	25	141.66	5.67		
Total	53	2053.5	38.75		

We also see that MSE is an unbiased estimator of σ^2 and that we can get an unbiased estimator of σ_p^2 . Our estimates are:

$$\begin{aligned}
 \hat{\sigma}^2 &= MSE = 5.66633 \\
 \hat{\sigma}_p^2 &= \frac{MSP(G) - MSE}{2} = \frac{73.04 - 5.67}{2} = 33.68533
 \end{aligned}$$

3.5 REML estimates

We fitted the full model with the lme4 package in R. The background of REML can be view in other lectures here on the web. The estimates we got from REML are:

$$\begin{aligned}
 \hat{\sigma}^2 &= 5.666334 \\
 \hat{\sigma}_p^2 &= 33.68533
 \end{aligned}$$

Which are the same as our previous estimates.

3.6 Estimating equations and solutions

If we think about the full model in matrix form

$$\mathbf{y} = X\boldsymbol{\beta} + Z\mathbf{u} + \boldsymbol{\varepsilon}$$

we see that

$$\text{Var}[\mathbf{u}] = \sigma_p^2 \cdot \mathbf{I}_{27} = \sigma^2 \cdot G$$

$$\text{Var}[\boldsymbol{\varepsilon}] = \sigma^2 \cdot \mathbf{I}_{54} = \sigma^2 \cdot R$$

Estimating equations for $\boldsymbol{\beta}$ and \mathbf{u} are

$$\hat{\boldsymbol{\beta}} = \{X'(R + ZGZ')^{-1}X\}^{-1} X'(R + ZGZ')^{-1}\mathbf{y}$$

$$\hat{\mathbf{u}} = (Z'R^{-1}Z + G^{-1})^{-1} [Z'R^{-1} - Z'R^{-1}X \cdot \{X'(R + ZGZ')^{-1}X\}^{-1} \cdot X'(R + ZGZ')^{-1}] \mathbf{y}$$

We have the following estimates for $\boldsymbol{\beta}$ and \mathbf{u} :

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_2 \\ \hat{\gamma}_1 \\ (\hat{\tau\gamma})_{21} \end{bmatrix} = \begin{bmatrix} 96.667 \\ 0.083333 \\ -0.73333 \\ 3.1167 \end{bmatrix}$$

$$\hat{\mathbf{u}} = \begin{bmatrix} \hat{\rho}_{1(0)} \\ \hat{\rho}_{2(0)} \\ \vdots \\ \hat{\rho}_{12(0)} \\ \hat{\rho}_{13(1)} \\ \hat{\rho}_{14(1)} \\ \vdots \\ \hat{\rho}_{27(1)} \end{bmatrix} = \begin{bmatrix} 3.959 \\ -4.804 \\ -7.572 \\ 3.036 \\ 11.339 \\ -7.572 \\ -0.653 \\ -1.576 \\ 0.269 \\ -5.727 \\ -0.192 \\ 9.494 \\ -7.410 \\ 2.275 \\ 0.892 \\ -4.182 \\ -3.259 \\ -3.259 \\ 1.353 \\ -1.876 \\ -5.566 \\ -3.721 \\ 10.116 \\ 0.892 \\ 5.504 \\ -0.492 \\ 8.733 \end{bmatrix}$$

3.7 ML estimates

We can also estimate the variances using maximum likelihood. The MLE's are

$$MLE(\sigma^2) = \frac{SSE}{N} = 5.246605$$

$$MLE(\sigma_p^2) = \frac{\frac{SSP(G)}{N} - \frac{SSE}{N}}{2} = 31.19013$$

We get the same estimates for β and \mathbf{u} because we see that the G -matrix should be the only thing that changes. The G -matrix is $G = \frac{\hat{\sigma}_p^2}{\hat{\sigma}^2} \cdot \mathbf{I}$. But the REML estimates were the same as the estimates from the Anova table. So we see that

$$\begin{aligned} \frac{MLE(\sigma_p^2)}{MLE(\sigma^2)} &= \frac{2 \cdot (SSP(G) - SSE)}{N} \cdot \frac{N}{SSE} \\ &= \frac{2 \cdot (SSP(G) - SSE)}{N - 2} \cdot \frac{N - 2}{SSE} \\ &= \frac{\hat{\sigma}_p^2}{\hat{\sigma}^2} \end{aligned}$$

When the REML estimate of variance and the estimate from the Anova table are not same this will not hold. But it does in our incident.

3.8 LRT of hypothesis

We can test hypothesis H_1 and H_2 by likelihood ratio test. We do this by comparing our full model

$$fm : y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \rho_{k(j)} + \epsilon_{ijk}$$

To reduced models. For H_1 we compare to

$$rm1 : y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \epsilon_{ijk}$$

and for H_2 to

$$rm2 : y_{ijk} = \mu + \tau_i + \gamma_j + \rho_{k(j)} + \epsilon_{ijk}$$

The likelihood ratio test statistic is

$$D = -2(\log(L_{rm}) - \log(L_{fm})).$$

This statistic follows approximately a χ^2 -distribution with one degree of freedom. For H_1 we get $D = 44.48906$ and a p-value, $p = 2.557828 \cdot 10^{-11}$. For H_2 we get $D = 7.70081$ and $p = 0.005519605$.

We therefor reject both hypothesis like we did with the previous F-test. We can't really simplify the model.

3.9 Estimates with all effects fixed

If we estimate assuming all effects are fixed. We have to have restrictions on the ρ 's. We will use

$$\begin{aligned} \rho_{1(0)} &= 0 \\ \rho_{13(1)} &= 0 \end{aligned}$$

Our model in matrix form we be

$$\mathbf{y} = X\beta + \epsilon.$$

Where we have

$$X\beta = \begin{bmatrix}
 1 & 0 & 0 & 0 & 0 & \cdots & \cdot & \cdot & \cdots & 0 \\
 1 & 0 & 0 & 0 & 1 & \ddots & \cdot & \cdot & \ddots & \cdot \\
 \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
 \cdot & \cdot & 0 & \cdot & \cdot & \ddots & 1 & \cdot & \ddots & \cdot \\
 \cdot & \cdot & 1 & \cdot & \cdot & \ddots & \cdot & 0 & \ddots & \cdot \\
 \cdot & \cdot & 1 & \cdot & \cdot & \ddots & \cdot & 1 & \ddots & \cdot \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
 \cdot & 0 & 1 & \cdot & 0 & \cdots & \cdot & \cdot & 0 & 1 \\
 \cdot & 1 & 0 & \cdot & 0 & \cdots & \cdot & \cdot & \cdots & 0 \\
 \cdot & 1 & 0 & \cdot & 1 & \ddots & \cdot & \cdot & \ddots & \cdot \\
 \cdot & \cdot & \cdot & \cdot & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
 \cdot & \cdot & 0 & 0 & \cdot & \ddots & 1 & \cdot & \ddots & \cdot \\
 \cdot & \cdot & 1 & 1 & \cdot & \ddots & \cdot & 0 & \ddots & \cdot \\
 \cdot & \cdot & 1 & 1 & \cdot & \ddots & \cdot & 1 & \ddots & \cdot \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
 1 & 1 & 1 & 1 & 0 & \cdots & \cdot & \cdot & 0 & 1
 \end{bmatrix}
 \begin{bmatrix}
 \mu \\
 \tau_2 \\
 \gamma_1 \\
 (\tau\gamma)_{21} \\
 \rho_{2(0)} \\
 \vdots \\
 \rho_{12(0)} \\
 \rho_{14(1)} \\
 \vdots \\
 \rho_{27(1)}
 \end{bmatrix}$$

The result we get are

$$\hat{\beta} = (X'X)^{-1}X'y = \begin{bmatrix} 100.95833 \\ 0.08333 \\ -13.05833 \\ 3.11667 \\ -9.50000 \\ -12.50000 \\ -1.00000 \\ 8.00000 \\ -12.50000 \\ -5.00000 \\ -6.00000 \\ -4.00000 \\ -10.50000 \\ -4.50000 \\ 6.00000 \\ 10.50000 \\ 9.00000 \\ 3.50000 \\ 4.50000 \\ 4.50000 \\ 9.50000 \\ 6.00000 \\ 2.00000 \\ 4.00000 \\ 19.00000 \\ 9.00000 \\ 14.00000 \\ 7.50000 \\ 17.50000 \end{bmatrix}$$

We though have to keep in mind the the interpretation of the parameters is quit different than in the random effect model.

3.10 Compare estimates of fixed effects REML vs ML

The estimates of the fixed effects are the same for both REML and ML estimates as we showed in the part about the ML estimate of variance.

3.11 Compare variance and SS estimates

The estimates we got for the variances can be seen in the following table.

Table 13: Variance estimates

Variance	SS-estimate	REML	ML
σ^2	5.66633	5.666334	5.246605
σ_p^2	33.68533	33.68533	31.19013

We see that the SS-estimate and the REML estimate are almost identical. The ML estimate is however lower. It's also well known the ML estimates for variances are biased.

3.12 Compare results from assuming fixed vs random effects

We can't directly compare the estimates from the model with all effect fixed vs. the original random effect model. But by examining the interpretation of the parameters we can compare between the models. The comparison can be seen in table 14.

Table 14: Comparison of fixed vs random effect model

Parameter FM	Parameters RM	Estimate FM	Estimate RM	Difference FM-RM
μ	$\mu + \rho_{1(0)}$	100.9583	100.6256	0.3328
τ_2	τ_2	0.0833	0.0833	0.0000
γ_1	$\gamma_1 + \rho_{13(1)} - \rho_{1(0)}$	-13.0583	-12.1027	-0.9557
$(\tau\gamma)_{21}$	$(\tau\gamma)_{21}$	3.1167	3.1167	0.0000
$\rho_{2(0)}$	$\rho_{2(0)} - \rho_{1(0)}$	-9.5000	-8.7634	-0.7366
$\rho_{3(0)}$	$\rho_{3(0)} - \rho_{1(0)}$	-12.5000	-11.5308	-0.9692
$\rho_{4(0)}$	$\rho_{4(0)} - \rho_{1(0)}$	-1.0000	-0.9225	-0.0775
$\rho_{5(0)}$	$\rho_{5(0)} - \rho_{1(0)}$	8.0000	7.3797	0.6203
$\rho_{6(0)}$	$\rho_{6(0)} - \rho_{1(0)}$	-12.5000	-11.5308	-0.9692
$\rho_{7(0)}$	$\rho_{7(0)} - \rho_{1(0)}$	-5.0000	-4.6123	-0.3877
$\rho_{8(0)}$	$\rho_{8(0)} - \rho_{1(0)}$	-6.0000	-5.5348	-0.4652
$\rho_{9(0)}$	$\rho_{9(0)} - \rho_{1(0)}$	-4.0000	-3.6898	-0.3102
$\rho_{10(0)}$	$\rho_{10(0)} - \rho_{1(0)}$	-10.5000	-9.6858	-0.8142
$\rho_{11(0)}$	$\rho_{11(0)} - \rho_{1(0)}$	-4.5000	-4.1511	-0.3489
$\rho_{12(0)}$	$\rho_{12(0)} - \rho_{1(0)}$	6.0000	5.5348	0.4652
$\rho_{14(1)}$	$\rho_{14(1)} - \rho_{13(1)}$	10.5000	9.6858	0.8142
$\rho_{15(1)}$	$\rho_{15(1)} - \rho_{13(1)}$	9.0000	8.3021	0.6979
$\rho_{16(1)}$	$\rho_{16(1)} - \rho_{13(1)}$	3.5000	3.2286	0.2714
$\rho_{17(1)}$	$\rho_{17(1)} - \rho_{13(1)}$	4.5000	4.1511	0.3489
$\rho_{18(1)}$	$\rho_{18(1)} - \rho_{13(1)}$	4.5000	4.1511	0.3489
$\rho_{19(1)}$	$\rho_{19(1)} - \rho_{13(1)}$	9.5000	8.7634	0.7366
$\rho_{20(1)}$	$\rho_{20(1)} - \rho_{13(1)}$	6.0000	5.5348	0.4652
$\rho_{21(1)}$	$\rho_{21(1)} - \rho_{13(1)}$	2.0000	1.8449	0.1551
$\rho_{22(1)}$	$\rho_{22(1)} - \rho_{13(1)}$	4.0000	3.6898	0.3102
$\rho_{23(1)}$	$\rho_{23(1)} - \rho_{13(1)}$	19.0000	17.5267	1.4733
$\rho_{24(1)}$	$\rho_{24(1)} - \rho_{13(1)}$	9.0000	8.3021	0.6979
$\rho_{25(1)}$	$\rho_{25(1)} - \rho_{13(1)}$	14.0000	12.9144	1.0856
$\rho_{26(1)}$	$\rho_{26(1)} - \rho_{13(1)}$	7.5000	6.9185	0.5815
$\rho_{27(1)}$	$\rho_{27(1)} - \rho_{13(1)}$	17.5000	16.1431	1.3569

We see that the estimates are similar but we notice that the absolute value of the ρ 's estimates are smaller for the random effect model.

3.13 T-test and comparison

To test hypothesis 3-6 we can use a t-test for the parameters in the model. It's can be shown that

$$\text{Var}[\hat{\beta}] = \sigma^2 \left(X' (R + ZGZ')^{-1} X \right)^{-1}$$

and we can calculate this matrix using the previous estimates for the variances. We can also just do this t-tests on the raw data. We can test H_3 and H_4 with a regular t-test comparing the group before (H_3) and after (H_4) the training period. We did this tests assuming equal variance. We can test H_5 and H_6 with a paired t-test or with regular t-test using the change in throwing velocity for each group. The results from this t-tests are in table 15.

Table 15: T-tests from full model and raw data

Hypothesis	Hypothesis with restriction	T-value from the full model	T-value from raw data
$H_3 : \mu_{10} = \mu_{11}$	$\gamma_1 = 0$	0.3018522	0.2936248
$H_4 : \mu_{20} = \mu_{21}$	$\gamma_1 + (\tau\gamma)_{21} = 0$	0.9810197	1.010041
$H_5 : \mu_{10} = \mu_{20}$	$\tau_2 = 0$	0.0857518	0.08304548
$H_6 : \mu_{21} = \mu_{11}$	$\tau_2 + (\tau\gamma)_{21} = 0$	3.68154	3.781253

We see that the t-values are similar for the tests from the full model and from the raw data. This implies that our assumption of equal variance is appropriate. We also see that H_6 is the only null-hypothesis we can reject ($\alpha = 0.05$). That is there is a significant change in throwing velocity for the players in the training group.

3.14 Equaliance of F-tests in the full model and the Δ -model

Now we will show that we could have tested H_2 (the test for the interaction) with a simple t-test comparing players change in throwing velocity by group. We do this by showing that the F-test for the interaction in the full model is the same as the F-test for the group effect in the Δ -model. Then we take for granted that the F-test in a simple anova with two groups is the same as doing a simple t-test. To show that the F-tests are the same we have to show that

$$\frac{MSG_{\Delta}}{MSE_{\Delta}} = \frac{MSTG_{fm}}{MSE_{fm}}.$$

To show this it's enough to show that

$$\begin{aligned} SSE_{\Delta} &= 2 \cdot SSE_{fm} \\ SSG_{\Delta} &= 2 \cdot SSTG_{fm} \end{aligned}$$

Some formulas we need for this are

$$\begin{aligned} \bar{y}_{\cdot jk} &= \frac{y_{2jk} + y_{1jk}}{2} \\ \bar{y}_{\cdot j\cdot} &= \frac{\bar{y}_{2j\cdot} + \bar{y}_{1j\cdot}}{2} \\ \bar{y}_{\dots} &= \frac{\bar{y}_{2\cdot\cdot} + \bar{y}_{1\cdot\cdot}}{2} \\ \Delta\bar{y}_{j\cdot} &= \Delta\bar{y}_{j\cdot} = \bar{y}_{2j\cdot} - \bar{y}_{1j\cdot} \\ \Delta\bar{y}_{\cdot\cdot} &= \Delta\bar{y}_{\cdot\cdot} = \bar{y}_{2\cdot\cdot} - \bar{y}_{1\cdot\cdot} \end{aligned}$$

We'll begin by showing this for the SSEs

$$\begin{aligned} SSE_{fm} &= \sum_i \sum_j \sum_k (y_{ijk} + \bar{y}_{\cdot j\cdot} - \bar{y}_{ij\cdot} - \bar{y}_{\cdot jk})^2 \\ &= \sum_j \sum_k (y_{2jk} + \bar{y}_{\cdot j\cdot} - \bar{y}_{2j\cdot} - \bar{y}_{\cdot jk})^2 + (y_{1jk} + \bar{y}_{\cdot j\cdot} - \bar{y}_{1j\cdot} - \bar{y}_{\cdot jk})^2 \\ &= \sum_j \sum_k \left(y_{2jk} + \frac{\bar{y}_{2j\cdot} + \bar{y}_{1j\cdot}}{2} - \bar{y}_{2j\cdot} - \frac{y_{2jk} + y_{1jk}}{2} \right)^2 + \left(y_{1jk} + \frac{\bar{y}_{2j\cdot} + \bar{y}_{1j\cdot}}{2} - \bar{y}_{1j\cdot} - \frac{y_{2jk} + y_{1jk}}{2} \right)^2 \\ &= \sum_j \sum_k \frac{1}{2^2} (y_{2jk} - y_{1jk} - (\bar{y}_{2j\cdot} - \bar{y}_{1j\cdot}))^2 + \frac{1}{2^2} (y_{1jk} - y_{2jk} - (\bar{y}_{1j\cdot} - \bar{y}_{2j\cdot}))^2 \\ &= \frac{1}{2} \sum_j \sum_k (\Delta y_{jk} - \Delta \bar{y}_{j\cdot})^2 \\ &= \frac{1}{2} SSE_{\Delta}. \end{aligned}$$

The proof is similar for the second part

$$\begin{aligned}
SSTG_{fm} &= \sum_i \sum_j n_j (\bar{y}_{ij} + \bar{y}_{i..} - \bar{y}_{i..} - \bar{y}_{.j})^2 \\
&= \sum_j n_j (\bar{y}_{2j} + \bar{y}_{..} - \bar{y}_{2..} - \bar{y}_{.j})^2 + n_j (\bar{y}_{1j} + \bar{y}_{..} - \bar{y}_{1..} - \bar{y}_{.j})^2 \\
&= \sum_j n_j \left(\bar{y}_{2j} + \frac{\bar{y}_{2..} + \bar{y}_{1..}}{2} - \bar{y}_{2..} - \frac{\bar{y}_{2j} + \bar{y}_{1j}}{2} \right)^2 + n_j \left(\bar{y}_{1j} + \frac{\bar{y}_{2..} + \bar{y}_{1..}}{2} - \bar{y}_{1..} - \frac{\bar{y}_{2j} + \bar{y}_{1j}}{2} \right)^2 \\
&= \sum_j \frac{n_j}{2} (\bar{y}_{2j} - \bar{y}_{1j} - (\bar{y}_{2..} - \bar{y}_{1..}))^2 + \frac{n_j}{2} (\bar{y}_{1j} - \bar{y}_{2j} - (\bar{y}_{1..} - \bar{y}_{2..}))^2 \\
&= \frac{1}{2} \sum_j n_j (\Delta \bar{y}_j - \Delta \bar{y}_{..})^2 \\
&= \frac{1}{2} SSG_{\Delta}
\end{aligned}$$

3.15 Discussion

We have shown what hypothesis we can reject and which not. We found that we could reject H_1 so the individual players effect is significant. We also found that we could reject H_2 so that either the groups were either different or the training had an effect on their throwing velocity. By testing hypothesis $H_3 - H_6$ we saw that the change in throwing velocity for the training group was the only significant thing. So we see that the training has an effect. The players throwing velocity increases.

Another interesting thing is that we could have tested all our hypothesis except the first one with t-tests on the raw data. And gotten the same or simmlar results.

4 Machines

4.1 Discussion

In terms of significant effects, we cannot decide which model is better. It appears that all the effects in the proposed model (be it mixed or fixed) have effect on the productivity score.

However since we are really interested in checking whether there is a difference in terms of productivity score of the machines, but not between individual workers in this sample, it is more natural to got with the mixed effect model.

The conclusion is, there is a significant difference of ovarall productivity score between the machines.

4.2 The data

Data on an experiment to compare three brands of machines used in an industrial process are presented. Six workers were chosen randomly among the employees of a factory to operate each machine three times. The response is an overall productivity score taking into account the number and quality of components produced. These data are shown in Figure 1.

4.3 The model and hypothesis

The proposed model is

$$y_{ijk} = \mu + \alpha_i + b_j + (\alpha b)_{ij} + \varepsilon_{ijk}, \quad i = 1, \dots, 3, \quad j = 1, \dots, 6, \quad k = 1 \dots, 3 \quad (17)$$

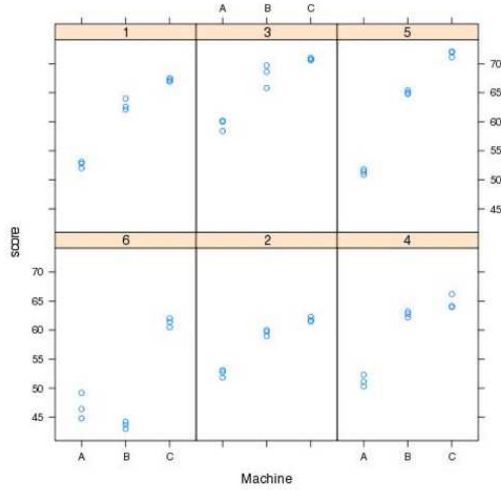


Figure 3: Overall productivity score for machines A, B and C for each worker.

where

- μ is a constant which describes the overall effect
- the α_i are constants to which we impose the restrictions $\sum \alpha_i = 0$
- b_i are independent $\mathcal{N}(0, \sigma_b^2)$ random variables
- $(\alpha b)_{ij}$ are pairwise independent $\mathcal{N}(0, \frac{I-1}{I} \sigma_{\alpha b}^2)$ random variables.
- ϵ_{ijk} are independent $\mathcal{N}(0, \sigma^2)$ random variables
- b_j , $(\alpha b)_{ij}$, and ϵ_{ijk} are pairwise independent for all i, j, k .

This model has a fixed main effect for each type of machine and a random main effects for every worker, namely α_i and b_j respectively. It also incorporates an interaction term between the two main effects which must be a random effect, namely $(\alpha b)_{ij}$.

We want to check the hypothesis whether there is any difference in productivity between machines. That is, we have the null hypothesis

$$H_{0A} : \alpha_i = 0 \quad \forall i \quad \text{v.s.} \quad H_{1A} : \exists i \text{ s.t. } \alpha_i \neq 0 \quad (18)$$

The hypothesis whether the workers have effect on productivity may be stated as follows

$$H_{0B} : \sigma_b^2 = 0 \quad \text{v.s.} \quad H_{1B} : \sigma_b^2 > 0. \quad (19)$$

To test whether or not machine and workers interact we test the following

$$H_{0(AB)} : \sigma_{\alpha b}^2 = 0 \quad \text{v.s.} \quad H_{1(AB)} : \sigma_{\alpha b}^2 > 0. \quad (20)$$

4.4 Matrix formulation

Now we want to find matrices X and Z which will satisfy

$$y = X\alpha + Zu + \epsilon.$$

We construct the following

$$y = \begin{pmatrix} y_{111} \\ y_{112} \\ y_{113} \\ y_{121} \\ \vdots \\ y_{163} \\ y_{211} \\ \vdots \\ y_{363} \end{pmatrix} + \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 1 \end{pmatrix} \alpha + \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ddots & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \ddots & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} u$$

and

$$\alpha = \begin{pmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}, \quad u = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_6 \\ (\alpha b)_{11} \\ \vdots \\ (\alpha b)_{16} \\ (\alpha b)_{21} \\ \vdots \\ (\alpha b)_{36} \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{113} \\ \varepsilon_{121} \\ \vdots \\ \varepsilon_{363} \end{pmatrix}.$$

To include the restraints in mentioned above we will introduce $\mu^* = \mu + \alpha_1$ and delete the second column in X .

4.5 Anova tables, tests and estimates of variance components

Let $N = 54$ denote the total number of observation, $I = 3$ denote the number of machines, $J = 6$ number of workers and $K = 3$ the totals numbers of trials of each worker on each machine. The total sum of squares may be partitioned into

$$\begin{aligned} \text{SSA} &= JK \sum_{i=1}^I (\bar{y}_{i..} - \bar{y}_{...})^2 && : \text{machine, fixed effect} \\ \text{SSB} &= IK \sum_{j=1}^J (\bar{y}_{.j.} - \bar{y}_{...})^2 && : \text{workers, random effect} \\ \text{SSAB} &= K \sum_{i=1}^I \sum_{j=1}^J (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 && : \text{interaction, random effect} \\ \text{SSE} &= \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (y_{ijk} - \bar{y}_{ij.})^2 && : \text{residuals} \end{aligned} \tag{21}$$

We also calculate the expected values of the mean sums of squares and the results can be seen in the following table

Source	SS	DF	MSS	E[MSS]
Machine	SSA	$I - 1$	$\frac{SSA}{I-1}$	$\sigma^2 + JK \frac{\sum \alpha_i^2}{I-1} + K \sigma_{\alpha b}^2$
Workers	SSB	$J - 1$	$\frac{SSB}{J-1}$	$\sigma^2 + IK \sigma_b^2$
Machine:Worker	SSAB	$(I - 1)(J - 1)$	$\frac{SSB(A)}{I(J-1)}$	$\sigma^2 + K \sigma_{\alpha b}^2$
Residuals	SSE	$IJ(K - 1)$	$\frac{SSE}{IJ(K-1)}$	σ^2
Total	SST	$IJK - 1$	$\frac{SST}{IJK-1}$	

To test hypothesis (18) we calculate

$$\frac{MSA}{MSAB}$$

and make use of the fact that this ratio follows a F -distribution. Similarly we check hypotheses (19) and (20) by calculating the ratios

$$\frac{MSB}{MSE} \quad \text{and} \quad \frac{MSAB}{MSE}$$

respectively and compare them to corresponding quantiles from the F -distribution.

From the ANOVA-table we obviously get the estimators

$$\hat{\sigma}^2 = \text{MSE}, \quad \hat{\sigma}_b^2 = \frac{MSB - \text{MSE}}{IK} \quad \text{and} \quad \hat{\sigma}_{\alpha b}^2 = \frac{MSAB - \text{MSE}}{K} \quad (22)$$

With a little help from R we get the following

	Df	SS	MSS	Ratio F-test	F-Value
Machines	2.00	1755.263	877.631	20.576	4.103
Workers	5.00	1241.895	248.379		2.477
Machine:Worker	10	426.530	42.653		2.106
Residuals	36.00	33.287	0.925		
Total	53.00	3456.975	7.66		

Table 16: ANOVA table

and the estimates of the variances become

$$\hat{\sigma}^2 = 0.925, \quad \hat{\sigma}_b^2 = 27.495 \quad \text{and} \quad \hat{\sigma}_{\alpha b}^2 = 13.909 \quad (23)$$

4.6 Estimating equations and solutions

When estimating α and u , we use the following equations

$$\hat{\alpha} = (X^T(R + ZGZ^T)^{-1}X)^{-1}X^T(ZGZ^T + R)^{-1}y$$

$$\hat{u} = (Z^T R^{-1}Z + G^{-1})^{-1}(Z^T R^{-1} - Z^T R^{-1}X(X^T(R + ZGZ^T)^{-1}X)^{-1}X^T(R + ZGZ^T)^{-1})y$$

where the matrices G and R fulfill the following equation

$$\text{Var} \begin{bmatrix} u \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \sigma^2 = \begin{bmatrix} \begin{bmatrix} \frac{\sigma_b^2}{\sigma^2} \mathbf{I}_6 & 0 \\ 0 & \frac{(I-1)\sigma_{\alpha b}^2}{I\sigma^2} \mathbf{I}_{18} \end{bmatrix} & 0 \\ 0 & \mathbf{I}_{54} \end{bmatrix}$$

That is $G = \begin{bmatrix} \frac{\sigma_b^2}{\sigma^2} \mathbf{I}_6 & 0 \\ 0 & \frac{(I-1)\sigma_{\alpha b}^2}{I\sigma^2} \mathbf{I}_{18} \end{bmatrix}$ and $R = \mathbf{I}_{18}$. We use our variance estimates in (23) and get the following estimates for α

and the estimates for u

	α_i
1	52.36
2	7.97
3	13.92

j	b_j	$(\alpha b)_{1j}$	$(\alpha b)_{2j}$	$(\alpha b)_{3j}$
1	0.24	0.01	0.54	0.16
2	-0.31	0.12	-0.10	-0.95
3	1.22	1.37	1.49	0.76
4	-0.01	-0.26	0.56	-0.34
5	0.58	-0.36	0.96	1.12
6	-1.71	-0.88	-3.45	-0.75

4.7 REML estimates

In order to find REML estimates for the variances σ^2 , σ_b^2 and $\sigma_{\alpha b}^2$ let K be a matrix such that

$$KX = 0.$$

To that end, we use a QR -decomposition for X , namely write X such that

$$X = QR$$

where R is an upper triangular $n \times p$ matrix and Q is an orthonormal $n \times n$ matrix, meaning the columns $[q_1, \dots, q_n]$ in Q form an orthonormal basis in \mathbb{R}^n . Then it is known that

$$[q_{p+1}, \dots, q_n]^T X = 0.$$

Consequently we choose $K = [q_{p+1}, \dots, q_n]^T$.

It is now easy to see that

$$\tilde{y} := Ky \sim \mathcal{N}(0, K\Sigma_y K^T)$$

where $\Sigma_y = \sigma^2(ZGZ' + R)$. The REML estimates can now be found by finding $(\sigma^2, \sigma_b^2, \sigma_{\alpha b}^2)$ which maximizes the log-likelihood function of $\mathcal{N}(0, K\Sigma_y K^T)$.

We use R to approximate this maxima. In order to use the optimizing functions in R we must specify a starting point. A sensible choice is using our estimates in (23). We get the REML estimates

$$\hat{\sigma}^2 = 0.924, \quad \hat{\sigma}_b^2 = 22.858 \quad \text{and} \quad \hat{\sigma}_{\alpha b}^2 = 13.90933$$

4.8 LRT of hypothesis

We can test hypothesis (18), (19) and (20) with a likelihood ratio test, which done by comparing the full model with the following models respectively

$$\text{reduced}_A : y_{ijk} = \mu + b_j + (\alpha b)_{ij} + \varepsilon_{ijk}$$

$$\text{reduced}_B : y_{ijk} = \mu + \alpha_i + (\alpha b)_{ij} + \varepsilon_{ijk}$$

$$\text{reduced}_{AB} : y_{ijk} = \mu + \alpha_i + b_j + \varepsilon_{ijk}$$

The likelihood ratio test statistic is given by

$$D = -2(\log(L_{\text{reduced}}) - \log(L_{\text{full}})).$$

This statistic follows approximately a χ^2 -distribution with one degree of freedom. The following table shows the D -values for each model respectively and corresponding p -values.

Thus we reject all the null hypothesis, just like we did earlier with the F -test.

	H_{0A}	H_{0B}	$H_{0(AB)}$
D	19.475	6.681	68.288
p -value	$5.9 \cdot 10^{-5}$	0.009	$1.4 \cdot 10^{-16}$

4.9 Estimates with all effects fixed

We assume all the effects, except ε_{ijk} , in the model

$$y_{ijk} = \mu + \alpha_i + b_j + (\alpha b)_{ij} + \varepsilon_{ijk} \quad (24)$$

are fixed. The corresponding matrix formulation becomes

$$y = \tilde{X}\tilde{\alpha} + \varepsilon$$

where

$$\tilde{X} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 & 0 & 0 & \ddots & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \ddots & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 1 & 0 & 0 & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & 1 & 0 \end{pmatrix}$$

and

$$\tilde{\alpha} = \begin{pmatrix} \mu^* \\ \alpha_2 \\ \alpha_3 \\ b_1 \\ b_2 \\ \vdots \\ b_6 \\ (\alpha b)_{11} \\ \vdots \\ (\alpha b)_{16} \\ (\alpha b)_{21} \\ \vdots \\ (\alpha b)_{36} \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{113} \\ \varepsilon_{121} \\ \vdots \\ \varepsilon_{363} \end{pmatrix}.$$

Because of linear dependence between columns in \tilde{X} we cannot estimate all the effects in β . Using the same restrictions as R we will exclude the effects of the first machine and the first worker, namely by removing columns 4,10 to 15, 16 and 22 from the matrix \tilde{X} . We get the following estimates

4.10 The SS-formulas in the ANOVA tables

Using the same method to partition the sum of squares as before we get a similar table; the only difference is the expected mean square since we are using fixed effects.

	α_i
μ^*	52.63
α_2	10.27
α_3	14.57
b_2	-0.07
b_3	6.90
b_4	-1.40
b_5	-1.27
b_6	-5.83
$(\alpha_b)_{22}$	-3.27
$(\alpha_b)_{23}$	-1.77
$(\alpha_b)_{24}$	1.23
$(\alpha_b)_{25}$	3.43
$(\alpha_b)_{26}$	-13.43
$(\alpha_b)_{32}$	-5.30
$(\alpha_b)_{33}$	-3.30
$(\alpha_b)_{34}$	-1.03
$(\alpha_b)_{35}$	5.80
$(\alpha_b)_{36}$	-0.07

Source	SS	DF	MSS	E[MSS]
Machine	SSA	$I - 1$	$\frac{SSA}{I-1}$	$\sigma^2 + JK \frac{\sum \alpha_i^2}{I-1}$
Workers	SSB	$J - 1$	$\frac{SSB}{J-1}$	$\sigma^2 + IK \frac{\sum b_i^2}{J-1}$
Machine:Worker	SSAB	$(I - 1)(J - 1)$	$\frac{SSB(A)}{I(J-1)}$	$\sigma^2 + K \frac{\sum \sum (\alpha b)_{ij}^2}{(I-1)(J-1)}$
Residuals	SSE	$IJ(K - 1)$	$\frac{SSE}{IJ(K-1)}$	σ^2
Total	SST	$IJK - 1$	$\frac{SST}{IJK-1}$	

We can check whether the machine, worker and the interaction effects are significant by calculating the ratios

$$\frac{MSB}{MSE}, \quad \frac{MSB}{MSE} \quad \text{and} \quad \frac{MSAB}{MSE}.$$

respectively and comparing them to corresponding quantiles from the F -distribution. Using the results from table 4.10 we get the values in table 4.10. There we clearly see that all the effects are significant.

Effects	Ratio F-test	F-Value
Machines	948.79	$> F_{2,36} = 3.26$
Workers	268.52	$> F_{5,36} = 2.478$
Machine:Worker	46.11	$> F_{10,36} = 2.106$
Residuals		

Table 17: ANOVA table

4.11 SS computed by 'hand'

Calculating the sums of squares in the ANOVA table (4.5) by 'hand' yields the following results by using the formulation in (21).

i	$\bar{y}_{i..}$	$(\bar{y}_{i..} - \bar{y}_{...})^2$
1	52.36	53.22
2	60.32	0.45
3	66.27	43.85
Total	178.950	97.515
SSA		1755.27

j	$\bar{y}_{.j.}$	$(\bar{y}_{.j.} - \bar{y}_{...})^2$
1	60.91	1.59
2	57.99	2.76
3	66.12	41.89
4	59.58	0.01
5	62.72	9.44
6	50.58	82.31
Total	357.90	137.9883
SSB		1241.895

j	$\bar{y}_{1i.}$	$\bar{y}_{2i.}$	$\bar{y}_{3i.}$	$(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$		
				$i = 1$	$i = 2$	$i = 3$
1	52.63	62.90	67.20	0.967	1.734	0.111
2	52.57	59.57	61.83	3.505	0.820	7.71
3	59.53	68.03	70.80	0.498	1.535	3.781
4	51.23	62.73	64.77	1.102	6.1669	2.054
5	51.37	65.07	71.73	16.493	2.796	5.707
6	46.80	43.63	61.30	12.367	58.014	16.810
Total				34.932	71.065	36.179
SSAB				426.53		

4.12 Maximum likelihood estimates

The ML estimates for α are the same as before. We may find the MLE for σ^2 with

$$\hat{\sigma}^2 = \frac{1}{IJK - 1} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (y_{ijk} - \hat{y}_{ijk})^2$$

where \hat{y}_{ijk} are the fitted values from the fixed effect model 24. With a little help from R we get $\hat{\sigma}^2 = 0.6164$

4.13 Compare variance and SS estimates

In the mixed effect model we got the following estimates for the fixed effects

$$\hat{\alpha} = \begin{pmatrix} \mu^* \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 52.36 \\ 7.97 \\ 13.92 \end{pmatrix}$$

and in the fixed effect models we got the following estimates for the fixed effects

$$\hat{\alpha} = \begin{pmatrix} \mu^* \\ \alpha_2 \\ \alpha_3 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ (\alpha_b)_{22} \\ (\alpha_b)_{23} \\ (\alpha_b)_{24} \\ (\alpha_b)_{25} \\ (\alpha_b)_{26} \\ (\alpha_b)_{32} \\ (\alpha_b)_{33} \\ (\alpha_b)_{34} \\ (\alpha_b)_{35} \\ (\alpha_b)_{36} \end{pmatrix} = \begin{pmatrix} 52.63 \\ 10.27 \\ 14.57 \\ -0.07 \\ 6.90 \\ -1.40 \\ -1.27 \\ -5.83 \\ -3.27 \\ -1.77 \\ 1.23 \\ 3.43 \\ -13.43 \\ -5.30 \\ -3.30 \\ -1.03 \\ 5.80 \\ -0.07 \end{pmatrix}.$$

We note that estimates for μ^* are the same in the mixed and fixed effect models. And the estimates for α_2 and α_3 are differ only slightly in the mixed and fixed effects models.

4.14 Compare variance and SS estimates

In the mixed effect model, the estimates of the variances by using the ANOVA table were

$$\hat{\sigma}^2 = 0.925, \quad \hat{\sigma}_b^2 = 27.495 \quad \text{and} \quad \hat{\sigma}_{\alpha_b}^2 = 13.909.$$

The REML estimates however were

$$\hat{\sigma}^2 = 0.924, \quad \hat{\sigma}_b^2 = 22.858 \quad \text{and} \quad \hat{\sigma}_{\alpha b}^2 = 13.90933.$$

We see that the estimates for σ^2 and $\sigma_{\alpha b}^2$ are nearly identical and the estimates for σ_b^2 differ slightly.

In the fixed effect model, the ANOVA estimate for σ^2 is found with the same method as in the mixed effect model, yielding the estimate

$$\hat{\sigma}^2 = 0.925.$$

The ML estimate for σ^2 is however

$$\hat{\sigma}^2 = 0.6164.$$

We see that the ML estimate is slightly lower than the ANOVA estimate.

Compare results from assuming fixed vs random effects In the mixed effect model we saw that the fixed effect terms were significant, and we rejected the null hypotheses that $\sigma_b^2 = 0$ and $\sigma_{\alpha b}^2 = 0$. Meaning that all the proposed effects in model 17 should be included since they appear to have a significant effect on the overall productivity score.

In the fixed effect model we saw that all the effects were significant and thus should not be excluded from model 24.

We also note that the ANOVA estimate of σ^2 for both models are the same but there is slight difference between the REML estimate for the mixed effect model and the ML estimate for the fixed effect model.

5 Rail

5.1 The Data

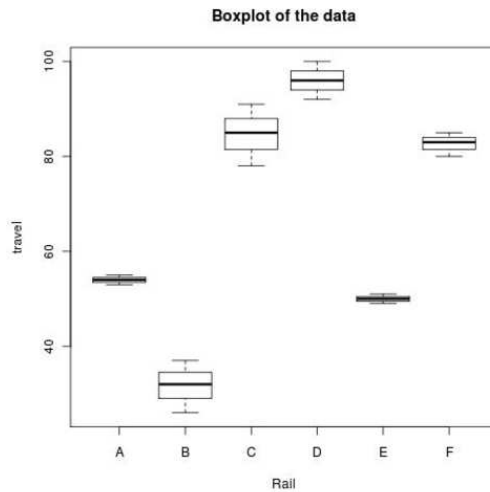


Figure 4: Boxplot of the data

The data (Rail) are from an experiment in nondestructive testing for longitudinal stress in railway. Six rails were chosen at random and tested three times each by measuring the time it took for a certain type of ultrasonic wave to travel the length of the rail. The following table contains the data

	Rail	Travel
1	A	55
2	A	53
3	A	54
4	B	26
5	B	37
6	B	32
7	C	78
8	C	91
9	C	85
10	D	92
11	D	100
12	D	96
13	E	49
14	E	51
15	E	50
16	F	80
17	F	85
18	F	83

We can see in the boxplot that there is more variability between groups than within groups.

5.2 The Model and Hypothesis

The model that is proposed is

$$y_{ij} = \mu + \beta_i + \varepsilon_{ij}$$

where $\beta_i \sim n(0, \sigma_A^2)$, $\varepsilon_{ij} \sim n(0, \sigma^2)$ and β and ε are independent. Here μ is the mean travel time across the population of rails being sampled and β_i is the random variable that represents the deviation from the population mean of the mean travel time for the i th rail. ε_{ij} is a random variable representing the deviation in travel time for observation j on rail i from the mean travel time for rail i .

The null hypothesis is

$$H_0 : \sigma_A^2 = 0$$

5.3 Matrix formulation

We can set up the model with the matrix formulation $Y = X\mu + Z\beta + \varepsilon$. Here does X contain the fixed effects and Z the random effects.

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \\ y_{16} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{24} \\ y_{25} \\ y_{26} \\ y_{31} \\ y_{32} \\ y_{33} \\ y_{34} \\ y_{35} \\ y_{36} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [\mu] + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \\ \epsilon_{15} \\ \epsilon_{16} \\ \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \\ \epsilon_{24} \\ \epsilon_{25} \\ \epsilon_{26} \\ \epsilon_{31} \\ \epsilon_{32} \\ \epsilon_{33} \\ \epsilon_{34} \\ \epsilon_{35} \\ \epsilon_{36} \end{bmatrix}$$

5.4 Anova tables, tests and estimates of variance components.

The anova table is almost the same as it were for the fixed effect model, that is

Source	df	SS	MS	E{MS}	F
Between groups	I-1	SSA= $\sum n_i(\bar{y}_i - \bar{y}_{..})^2$	MSA=SSA/(I-1)	$\sigma^2 + n\sigma_A^2$	MSA/MSE
Whithin group	n-I	SSE= $\sum \sum (y_{ij} - \bar{y}_i)^2$	MSE=SSE/(n-I)	σ^2	
Total	n-1	SSTOT= $\sum \sum (y_{ij} - \bar{y}_{..})^2$			

When we calculate these sums we get

	df	SS	MSE	F*	F
A	5	9319.5	1862.1	115.18	3.11
Error	12	194	16.2		
Total	17	9504.5			

Here we can see that $F^* > F$ so we can reject the hypothesis that $\sigma_A^2 = 0$.

From the anova table we can calculate the variances of compounds, that is $\hat{\sigma}^2 = MSE = 16.2$ and $\hat{\sigma}_A^2 = \frac{MSA - MSE}{n_i} = 615.3$.

5.5 Estimating equations (for mu and beta) and solutions.

When we estimate μ and β , we can use the equations

$$\hat{\mu} = (X^T(R + ZGZ^T)^{-1}X)^{-1}X^T(ZGZ^T + R)^{-1}y$$

$$\hat{\beta} = (Z^T R^{-1}Z + G^{-1})^{-1}(Z^T R^{-1} - Z^T R^{-1}X(X^T(R + ZGZ^T)^{-1}X)^{-1}X^T(R + ZGZ^T)^{-1})y$$

where we have that

$$\text{Var} \begin{bmatrix} \beta \\ \epsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \sigma^2 = \begin{bmatrix} \sigma_A^2/\sigma^2 \mathbf{I}_6 & 0 \\ 0 & \mathbf{I}_{18} \end{bmatrix}$$

That is $G = \sigma_A^2 / \sigma^2 \mathbf{I}_6$ and $R = \mathbf{I}_{18}$. Then we have everything to calculate $\hat{\mu}$ and $\hat{\beta}$ (we use σ^2 and σ_A^2 from slide 25). We then obtain

$$\hat{\mu} = 66.5$$

and

$$\hat{\beta} = \begin{bmatrix} -12.39125 \\ -34.53029 \\ 18.00862 \\ 29.24335 \\ -16.35645 \\ 16.02602 \end{bmatrix}$$

5.6 REML estimates (of variances).

We can use the REML method to estimate the variances. At first we need to find a matrix K such that

$$KX = 0$$

We can find that matrix with the QR-decomposition. We get

$$X = QR = [q_1, q_2, \dots, q_p, \dots, q_n]R$$

Here q_1, q_n form an orthonormal basis, Q is an $n \times n$ matrix and R is an $n \times p$ matrix. Then we have that

$$[q_{p+1}, \dots, q_n]'X = 0$$

so we can choose $K = [q_{p+1}, \dots, q_n]'$. Now we know that $Ky = \tilde{y} \sim n(0, K\Sigma_y K')$ so the likelihood function of \tilde{y} is

$$L(\theta) = \frac{e^{-1/2\tilde{y}'K\Sigma_y K'\tilde{y}}}{(2\pi)^{n/2} |K\Sigma_y K'|^{1/2}}$$

Then we find $\hat{\theta} = (\hat{\sigma}_A^2, \hat{\sigma}^2)$ which solves $\max_{\theta} \ln(L(\theta))$.

By programing this in R we can obtain the following estimates $\hat{\theta} = (\hat{\sigma}_A^2, \hat{\sigma}^2) = (615.72530, 14.92307)$. These estimates are similar to those that the built in function `lmer` gives us but not exactly the same (although it uses the REML method). Those estimates are (615.311, 16.167).

5.7 LRT of hypothesis

We can use likelihood ratio test to compare the fit of two models where one is nested in the other. We can compare our random effect model

$$y_{ij} = \mu + \beta_i + \epsilon_{ij}$$

to the model

$$y_{ij} = \mu + \epsilon_{ij}$$

Lets call the likelihood function for the more complex model L_1 and the likelihood function for the other one L_0 . We need to calculate the likelihood ratio test statistic, that is

$$D = -2(\log(L_0) - \log(L_1))$$

The distribution of this test statistic is approximately chi squared with degrees of freedom equal to the difference in the dimension of the two parameters spaces. With calculation in R we get that $D = 28.30046$ and we get the p-value $1.038714e - 07$. The p value is well below 5% so we can reject the null hypothesis $H_0 : \sigma_A^2 = 0$.

5.8 Estimates with all effects fixed

If we assume that all the effects are fixed we get the model

$$y_{ij} = \mu + \beta_i + \varepsilon_{ij}$$

where β_i are constants ($\beta_1 = 0$) and $\varepsilon \sim n(0, \sigma^2)$. The matrix form of the model is $y = X\beta + \varepsilon$ (now there is no Z). That is

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \\ y_{16} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{24} \\ y_{25} \\ y_{26} \\ y_{31} \\ y_{32} \\ y_{33} \\ y_{34} \\ y_{35} \\ y_{36} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{15} \\ \varepsilon_{16} \\ \varepsilon_{21} \\ \varepsilon_{22} \\ \varepsilon_{23} \\ \varepsilon_{24} \\ \varepsilon_{25} \\ \varepsilon_{26} \\ \varepsilon_{31} \\ \varepsilon_{32} \\ \varepsilon_{33} \\ \varepsilon_{34} \\ \varepsilon_{35} \\ \varepsilon_{36} \end{bmatrix}$$

Now it is easy to estimate β . We use simple linear regression methods (OLS and MLE give the same estimate), that is

$$\hat{\beta} = (X^T X)^{-1} X^T y = \begin{bmatrix} 54.00000 \\ -22.33333 \\ 30.66667 \\ 42.00000 \\ -4.00000 \\ 28.66667 \end{bmatrix}$$

5.9 The SS - formulas in the ANOVA table

The anova table is almost the same as the one for the random effect model. The only difference is in the $E\{MS\}$.

Source	df	SS	MS	$E\{MS\}$	F
Between groups	I-1	$SSA = \sum n_i (\bar{y}_i - \bar{y}_{..})^2$	$MSA = SSA / (I-1)$	$\sigma^2 + \frac{\sum n_i (\beta_i - \beta_{..})^2}{I-1}$	MSA / MSE
Whithin group	n-I	$SSE = \sum \sum (y_{ij} - \bar{y}_i)^2$	$MSE = SSE / (n-I)$	σ^2	
Total	n-1	$SSTOT = \sum \sum (y_{ij} - \bar{y}_{..})^2$			

5.10 SS computed "by hand"

Now we can compute

Factor	i	y_{i1}, y_{i2}, y_{i3}	$\sum_j y_{ij}$	\bar{y}_i	$\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$	$n(\bar{y}_i - \bar{y}_{..})$
A	1	55, 53, 54	162	54	2	468.75
B	2	26, 37, 32	95	31.67	60.67	3640.083
C	3	78, 91, 85	254	84.67	84.67	990.083
D	4	92, 100, 96	288	96	32	2610.75
E	5	49, 51, 50	150	50	2	816.75
F	6	80, 85, 83	248	82.67	12.67	784.083
Σ			1197		194	9310.5

where we have that $\bar{y}_{..} = \frac{1197}{18} = 66.5$. We also know that $n = 18$, $I = 6$ and $n_i = 3$. Now we can fill in the anova table

	df	SS	MSE	F*	F
A	5	9319.5	1862.1	115.18	3.11
Error	12	194	16.2		
Total	17	9504.5			

Here we can see that $F^* > F$ so we can reject the hypothesis that $\beta = 0$. From the anova table we get the following estimate $\hat{\sigma}^2 = MSE = 16.2$.

5.11 ML estimates

We have the likelihood function

$$L(\beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left[-\frac{1}{2\sigma^2} \sum_i \sum_j (y_{ij} - (\mu + \beta_i))^2 \right]$$

Maximizing the likelihood function is equivalent to minimizing the sum $\sum_i \sum_j (y_{ij} - (\mu + \beta_i))^2$ in the exponent. The i th component of this sum is $Q_i = \sum_j (y_{ij} - (\mu + \beta_i))^2$. When we differentiate with respect to β_i , we obtain

$$\frac{dQ_i}{d\beta_i} = \sum_j -2(y_{ij} - (\mu + \beta_i))$$

If we set the derivative equal to zero and isolate β we obtain the MLE estimator. That is $\hat{\beta} = \bar{y}_i - \mu$. Because $\beta_1 = 0$ we obtain $\mu = \bar{y}_{1..}$. Then we get the estimate

$$\hat{\beta}_i = \begin{bmatrix} 54.00000 \\ -22.33333 \\ 30.66667 \\ 42.00000 \\ -4.00000 \\ 28.66667 \end{bmatrix}$$

Which is the same as we got in slide 45.
The logarithm of the likelihood function is

$$\log L(\hat{\beta}, \sigma^2) = -\frac{n}{2} \log(2\pi\sigma) - \frac{1}{2\sigma^2} \sum_i \sum_j (y_{ij} - \hat{\beta}_i)^2$$

When we differentiate with respect to σ^2 we obtain

$$-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_i \sum_j (y_{ij} - \hat{\beta}_i)^2$$

If we set the derivative equal to zero and isolate σ^2 we obtain the MLE estimator. That is

$$\hat{\sigma}^2 = \frac{1}{n} \sum_i \sum_j (y_{ij} - \hat{\beta}_i)^2$$

Then we get the estimate

$$\hat{\sigma}^2 = 194/18 = 10.78$$

which is smaller than we got from the anova table. (The ML estimate for σ^2 is biased).

5.12 Compare estimates of fixed effects [sl30 vs sl60]

In the random effect model we got the following estimate for the fixed effect

$$\hat{\beta} = 66.5$$

and the estimate for the random effect was

$$\hat{u} = \begin{bmatrix} -12.39125 \\ -34.53029 \\ 18.00862 \\ 29.24335 \\ -16.35645 \\ 16.02602 \end{bmatrix}$$

In the fixed effect model we got the following estimates

$$\hat{\beta} = \begin{bmatrix} 54.00000 \\ -22.33333 \\ 30.66667 \\ 42.00000 \\ -4.00000 \\ 28.66667 \end{bmatrix}$$

We need to remember how the random effect model and the fixed were built up to understand how these estimates are related. The models were shown in matrix form in slides 20 and 45. It can be seen that μ in the fixed effect model is related to $\mu - \beta_1$ in the random effect model, that is 54 should be not so far from $66.5 - 12.39 = 54.11$, which is true.

5.13 Compare variance and SS estimates [sl25, sl 35, sl55, sl60]

The SS estimates are the same for both models, but there is a difference in the estimates for the variance.

Variance estimates for the random effect model were the following:

In slide 25 we got an estimate for $\hat{\sigma}^2$ and $\hat{\sigma}_A^2$ from the anova table. That is

$$\hat{\sigma}^2 = MSE = 16.2 \text{ and } \hat{\sigma}_A^2 = \frac{MSA - MSE}{n_i} = 615.3.$$

The MSE estimate for σ^2 is the same one as in the fixed effect model. In slide 35 we used the REML method to get an estimate for $\hat{\sigma}^2$ and $\hat{\sigma}_A^2$. They were close to those we got in slide 25 but not exactly the same. The estimate for σ^2 was closer to the one we got for σ^2 on slide 25 than the estimate for $\hat{\sigma}_A^2$. We obtained

$$\hat{\sigma}_A^2 = 615.72530 \text{ and } \hat{\sigma}^2 = 14.92307$$

Variance estimates for the fixed effect model were the following:

When we used the anova table in slide 55 we got an estimate for σ^2 (in the fixed effect model there is no σ_A^2)

$$\hat{\sigma}^2 = MSE = 16.2$$

When we used the ML method in slide 60 we got another estimate (that estimate is biased), which was lower than the one from the anova table

$$\hat{\sigma}^2 = 194/18 = 10.78$$

5.14 Compare results from assuming fixed vs random effects

In the fixed effect model all effects were assumed to be fixed. All the parameters in the model were significant, and we estimated them with MLE. We also estimated σ^2 with two methods, both with an anova table and maximum likelihood method. Those estimates were quite different.

In the random effect model we assumed that $\beta \sim n(0, \sigma_A^2)$, so we had to estimate σ_A^2 and σ^2 . They were estimated with two methods, REML and ANOVA. The estimates for σ_A^2 were the same in both methods and estimates for σ^2 were similar in both cases, but not the same. The ANOVA estimate for σ^2 was the same as the one for σ^2 estimated with ANOVA in the fixed effect model. We rejected the hypothesis that $\sigma_A^2 = 0$, so the parameters in the model were significant. Because we rejected the hypothesis that $\sigma_A^2 = 0$ we assume that the random effect model is a better model than the fixed effect model.

5.15 Discussion

The random effect model that we started with was the model

$$y_{ij} = \mu + \beta_i + \epsilon_{ij}$$

We rejected the hypothesis $\sigma_A^2 = 0$, so we did not reduce the model. Because we rejected the hypothesis it was decided that the random effect model was better than the fixed effect model. The conclusion was that the rail was a factor that had random effect on the travel time.