Introduction to Linear Mixed-Effect Models

Niels Hagenbuch

March 1, 2010

Linear regression model:

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \varepsilon_i, \quad (i = 1, \ldots, n \textit{ observations}) \\
&= \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \quad (p \textit{ parameters}) \\
&= \mathbf{x}_i^{\mathsf{T}} \boldsymbol{\beta} + \varepsilon_i, \qquad \textit{with} \quad \varepsilon_i \sim \mathcal{N}\langle\, 0, \sigma^2 \,\rangle
\end{aligned}
$$

or in matrix notation:

$$
\underset{(n \times 1)}{\mathbf{y}} = \underset{(n \times p)}{\mathbf{X}} \times \underset{(p \times 1)}{\boldsymbol{\beta}} + \underset{(n \times 1)}{\boldsymbol{\varepsilon}}
$$

R-output of a simple linear regression model (intercept and one covariate x1):

```
Call:
lm(formula = y1 ~ x1, data = anscombe)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.92127 -0.45577 -0.04136  0.70941  1.83882 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)   
(Intercept)   3.0001     1.1247   2.667  0.02573 * 
x1            0.5001     0.1179   4.241  0.00217 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 1.237 on 9 degrees of freedom
Multiple R-squared: 0.6665, Adjusted R-squared: 0.6295 
F-statistic: 17.99 on 1 and 9 DF,  p-value: 0.002170
```

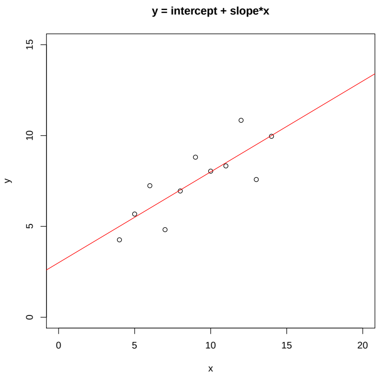


Figure: Plot of 11 data points and a corresponding fit of a linear regression model according to $y = \beta_0 + \beta_1 x + \varepsilon$ (data from anscombe).

Analysis of Variance (ANOVA) model for two (crossed) factors:

$$y_{ijk} = \mu + \beta_{i1} + \beta_{j2} + \varepsilon_{ijk}$$

with:

μ population mean

β_{i1} main effect of factor level β_{i1}

β_{j2} main effect of factor level β_{j2}

$\varepsilon_{ijk} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$ (error term)

R-output of a ANOVA model:

```
Analysis of Variance Table

Response: score
          Df  Sum Sq Mean Sq F value    Pr(>F)
Rater      2 3048.78 1524.39 50.3505 2.072e-11 ***
Manager   19   51.55    2.71  0.0896         1
Residuals 38 1150.47   30.28
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Classical Analysis of Variance (ANOVA)

- Classical ANOVA developed for agricultural experiments
- Control over
  - crop varieties
  - type of fertilizers used
  - amount of fertilizer
  - plots, subplots
  - greenhouse: light, irrigation, soil type, etc.
- Need for randomization in order to avoid confounding
- Different designs

Classical ANOVA: Randomization

TABLE 2.03. ORTHOGONAL 10×10 SQUARE AND ONE OF A SET OF FIVE 12×12 SQUARES

A_1	B_2	C_3	D_4	E_5	F_6	G_7	H_8	I_9	J_{10}
B_3	C_1	A_2	G_9	H_{10}	I_4	J_5	D_6	E_7	F_8
C_2	A_3	B_1	F_5	G_6	H_7	I_8	J_9	D_{10}	E_4
D_8	H_4	I_{10}	A_7	B_9	E_1	C_6	F_3	J_2	G_5
E_9	I_5	J_4	H_6	A_8	B_{10}	F_1	C_7	G_3	D_2
F_{10}	J_6	D_5	E_2	I_7	A_9	B_4	G_1	C_8	H_3
G_4	D_7	E_6	I_3	F_2	J_8	A_{10}	B_5	H_1	C_9
H_5	E_8	F_7	C_{10}	J_3	G_2	D_9	A_4	B_6	I_1
I_6	F_9	G_8	J_1	C_4	D_3	H_2	E_{10}	A_5	B_7
J_7	G_{10}	H_9	B_8	D_1	C_5	E_3	I_2	F_4	A_6

1	2	3	4	5	6	7	8	9	10	11	12
2	3	4	5	6	1	8	9	10	11	12	7
3	4	5	6	1	2	9	10	11	12	7	8
4	5	6	1	2	3	10	11	12	7	8	9
5	6	1	2	3	4	11	12	7	8	9	10
6	1	2	3	4	5	12	7	8	9	10	11
7	8	9	10	11	12	1	2	3	4	5	6
8	9	10	11	12	7	2	3	4	5	6	1
9	10	11	12	7	8	3	4	5	6	1	2
10	11	12	7	8	9	4	5	6	1	2	3
11	12	7	8	9	10	5	6	1	2	3	4
12	7	8	9	10	11	6	1	2	3	4	5

Figure: FISHER, Ronald A ; YATES, Frank: *Statistical Tables for Biological, Agricultural and Medical Research.* 6^{th} edition, revised and enlarged. Edinburgh/London: Oliver and Boyd, 1963, p. 25.

Classical ANOVA: Randomization

TABLE 2.2. YOUDEN SQUARE SOLUTION ($b = v = 31$, $r = k = 10$)

a b c d e f g h i j	i m t g u j B q v z	q r k n j D b c x u	y e j B k z r a E s
b l q u p a t y e o	j y h f t n m o w r	r f l i B C v b D y	z w v e q k f d l n
c k d y x o i t s v	k o b s h w q C B g	s t n D b v z A a h	A x e m r s l i h q
d q m w A c D B y a	l C r t a g w z c x	t p i k D e C j A w	B i a x s p u w n f
e B x b m E d n t C	m z f p g y A x b k	u h B c E A k l f t	C a u v d r h k p m
f s C o c m e u z D	n A s C y u j g d l	v j w E l b s p m c	D E g q f t p s r d
g D E a n l o m k i	o n p r v B c e g A	w u z A i d E r o b	E g D h w x y v u e
h c y z C i n E q p	p d o l z h x D j B	x v A j o q a f C E	

Figure: FISHER, Ronald A ; YATES, Frank: *Statistical Tables for Biological, Agricultural and Medical Research.* 6^{th} edition, revised and enlarged. Edinburgh/London: Oliver and Boyd, 1963, p. 28.

Experimental Design: Crossed vs. Nested

- Two factors are said to be **crossed** if all levels of the first factor appear in combination with all levels of the other factor.
- A crossed design is also called *factorial*.
- Example: five enzymes are tested in two solutions with different pH (four measurements each).

Two way factorial ANOVA

	1	2	3	4	5
pH level 1	* * * *	* * * *	* * * *	* * * *	* * * *
pH level 2	* * * *	* * * *	* * * *	* * * *	* * * *

Nested ANOVA

1		2		3		4		5		
1	2	1	2	1	2	1	2	1	2	*Sugar batches*
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	

Experimental Design: Crossed vs. Nested

- A factors is said to be **nested** under (or within) another factor if any given level of the nested factor appears at only one level of the nesting factor.
- Rationale: either given by the experiment (sample from a batch) or to avoid disproportional complexity (three raters evaluate ten managers each [30 interviews instead of 90]).
- A nested design is also called *hierarchical*.
- Example: From five production batches two samples are taken for quality control (four measurements each).

Two way factorial ANOVA

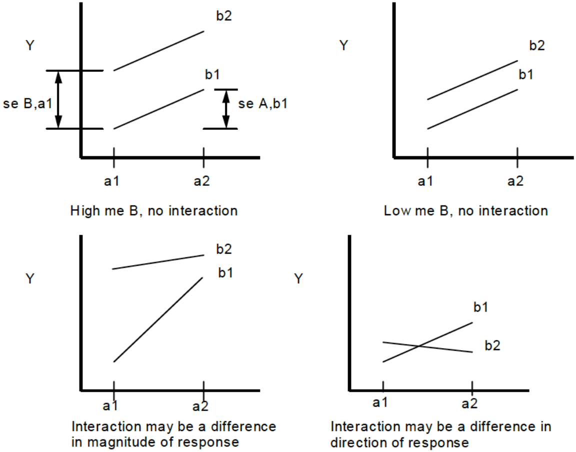
	1	2	3	4	5
pH level 1	* * * *	* * * *	* * * *	* * * *	* * * *
pH level 2	* * * *	* * * *	* * * *	* * * *	* * * *

Nested ANOVA

1		2		3		4		5		
1	2	1	2	1	2	1	2	1	2	*Sugar batches*
* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	* * * *	

Interaction

If the effect of one covariate changes considerably as the value of another covariate changes.



Interaction

The resulting effect is no longer additive and has to be modelled separately:

$$y_{ijk} = \mu + \beta_{i1} + \beta_{j2} + \beta_{ij3} + \varepsilon_{ijk}$$

with:

μ	population mean
β_{i1}	main effect of factor level β_{i1}
β_{j2}	main effect of factor level β_{j2}
β_{ij3}	interaction term $\beta_{i1} \times \beta_{j2}$

Fixed vs. Random Effects

- So far, covariates were considered **fixed**.
- Levels were chosen deliberately
- Levels had a specific meaning
- Changing levels leads to fundamental changes in the experiment

Examples:

1. Different crop varieties (oat; oat and rye; oat, rye, barley)
2. Medical test on humans (males; males and females)
3. Life expectancy (german speaking countries; german and french; german, french, italian)

Fixed vs. Random Effects

Random effects take into account that there is further variation:

- subjects tested
- batches or units sampled
- plots used to grow crops

The last example shows the shift in interest: plots can be defined as *fixed*, then the analysis provides information on the specific plots used.

Or plots are defined as *random*, considering them as arbitrary source of variability without further significance for the experiment.

Random Effects

Random effects

- assess variability
- reduce threats to validity
- increase generalizability

Random effects allow to generalize

- the results: schools in Markus' nano-introduction
- the research question: can a research method be used "universally"?

Random Effects

When should a factor be considered **random**?

- Specific levels could be replaced by other levels (e.g. subjects): the chosen levels are **arbitrary** or **substitutable**
- The conclusion of the experiment is to be **generalized** (e.g. other schools)
- Conclusions drawn for each separate level are **not of interest**: not the specific items are important but the population they are drawn from

Model

Linear mixed-effects model for one fixed and one random factor:

$$y_{ij} = \mu + \beta_j + b_i + \varepsilon_{ij}$$

or in matrix notation:

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta} + \mathbf{Z_i} b_i + \boldsymbol{\varepsilon}_\mathbf{i}$$

with:

μ population mean

β_j fixed effect, $j = 1, \ldots, n_j$ (levels of the fixed factor)

b_i random effect, $i = 1, \ldots, M$ (subjects, batches, etc.)

$b_i \sim \mathcal{N}\langle 0, \sigma_b^2 \rangle$ ("between-variability")

$\varepsilon_{ij} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$ (error term or "within-variability")

$b_i \perp \varepsilon_{ij}$

Software

The new variance term σ_b^2 needs to be estimated numerically, implemented differently in software packages:

`R:`	`lme()` from package `nlme` (Pinheiro/Bates 2000)
	`lmer()` from package `lme4` (Bates 2010)
SAS:	PROC GLM, NESTED, ANOVA or VARCOMP
SPSS:	MIXED

A first example: the `Rail` dataset

Example for a simple "one factor design" or "one-way classification".
The datset is included in the package `nlme` (`R: data(Rail)`).

Data:

`rail`	6 levels (six rails chosen at random)
`travel`	3 measurements per rail of travel time of a ultrasonic wave

Interest:

1. Average travel time of an ultrasonic sound wave in a rail (= expected travel time)
2. Variation in travel time among rails (= between-rail variability, $\hat{\sigma}_b^2$)
3. Variation in the three measurements of travel time for a single rail (= within-rail variability, $\hat{\sigma}^2$)

A first example: the Rail dataset

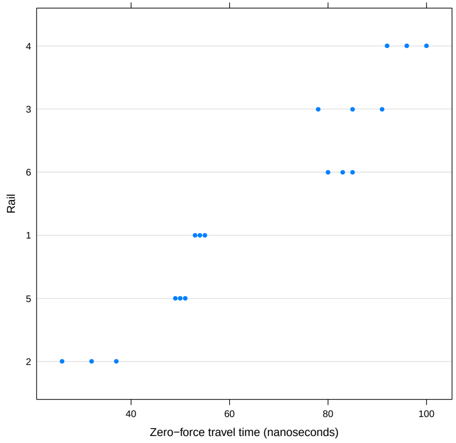


Figure: Plot of the 6 × 3 data points, grouped by Rail (lines) in ascending order of the mean travel time per rail. The three measurements are depicted as dots. Overall the between-rail variability is much larger than the within-rail variability.

A first example: the `rail` dataset

Model:

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

with

- y_{ij} response variable: travel time of rail i at j^{th} measurement
- μ mean travel time across all 6 rails (= estimate of the population mean of all rails) (fixed effect)
- b_i random effect of rail i (= deviation from μ)
- ε_{ij} error term

Assumptions:

1. $b_i \sim \mathcal{N}\langle 0, \sigma_b^2 \rangle$ ("between-rail variability")
2. $\varepsilon_{ij} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$ (error term or "within-rail variability")
3. $b_i \perp \varepsilon_{ij}$

A first example: the `Rail` dataset

Implementation in R:

```
lme(fixed= travel ~ 1, random = ~ 1 | Rail, data=Rail)
```

or

```
lmer(travel ~ 1 + (1|Rail), data=Rail)
```

Both calls model a single fixed effect for all rails ($=$ intercept or μ or estimate of the population mean), and a random effect for each rail (grouping variable is `Rail`).

A first example: the `Rail` dataset

Output (`lmer`):

```
Linear mixed model fit by REML
Formula: travel ~ 1 + (1 | Rail)
   Data: rail
 REML
122.2

Random effects:
 Groups   Name        Variance Std.Dev.
 Rail     (Intercept) 615.311  24.8055
 Residual              16.167   4.0208
Number of obs: 18, groups: Rail, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)    66.50      10.17   6.538
```

where $\hat{\mu}$ = `(Intercept): 66.50`

$\hat{\sigma}_b$ = `Rail: 24.8055` (between-rail variability)

$\hat{\sigma}$ = `Residual: 4.0208` (within-rail variability)

A simple model

Ignore grouping: $y_{ij} = \mu + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$

```
Call:
lm(formula = travel ~ 1, data = rail)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   66.500      5.573   11.93 1.10e-09 ***

Residual standard error: 23.65 on 17 degrees of freedom
```

where $\hat{\mu}$ = `(Intercept):  66.50`

$\hat{\sigma}$ = `Residual standard error:  23.65`

are the direct estimators from the data:

```
> mean(rail$travel)
[1] 66.5
> sd(rail$travel)
[1] 23.64505
```

Is it a good model?

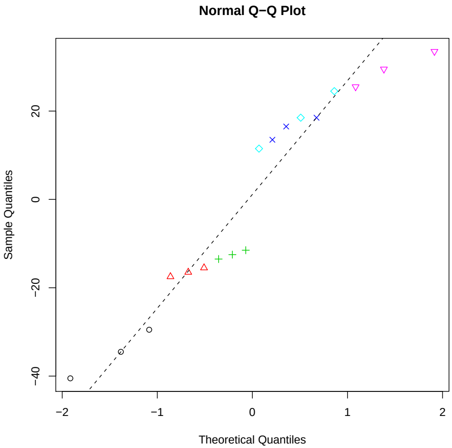


Figure: Normal-Plot of the residuals ($r_i = y_i - \hat{y}_i$) for the linear regression model. There is a repetitive pattern of deviation from the standard normal due to ignored grouping effect (`Rail`).

Add the grouping factor

Include grouping: $y_{ij} = \beta_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$

```
Call:
lm(formula = travel ~ Rail - 1, data = rail)

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
Rail1   54.000      2.321   23.26 2.37e-11 ***
Rail2   31.667      2.321   13.64 1.15e-08 ***
Rail3   84.667      2.321   36.47 1.16e-13 ***
Rail4   96.000      2.321   41.35 2.59e-14 ***
Rail5   50.000      2.321   21.54 5.86e-11 ***
Rail6   82.667      2.321   35.61 1.54e-13 ***

Residual standard error: 4.021 on 12 degrees of freedom
Multiple R-squared: 0.9978, Adjusted R-squared: 0.9967
F-statistic: 916.6 on 6 and 12 DF,  p-value: 2.971e-15
```

where the $\hat{\beta}_i$ are listed as `Rail1`,...,`Rail6` and $\hat{\sigma}$ is 4.021.

- model fit is better (residual plots not shown)
- $\hat{\sigma}$ (variability within the rails) decreased from 23.65 to 4.021
- model does not provide an estimate for the variability *between* rails.

Developing the random-effects model

Reparameterization of the linear model

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

to

$$y_{ij} = \bar{\beta} + (\beta_i - \bar{\beta}) + \varepsilon_{ij}, \qquad \bar{\beta} = \frac{1}{M}\sum_{i=1}^{M} \beta_i$$

leads to

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

where

- $\bar{\beta}$ is replaced by μ, the mean travel time across the population of rails being sampled (fixed effect)
- $(\beta_i - \bar{\beta})$ is replaced by b_i, random variables representing the deviation from the population mean of the mean travel time for the i^{th} rail; the distribution of the b_i has to be estimated
- ε_{ij} random variable representing the deviation in travel time for observation j on rail i from the mean travel time for rail i

Two factors: 1 fix / 1 random – the `Machines` dataset

Example for a "replicated, randomized block design" with an interaction term.
The dataset is included in the package `nlme` (`R: data(Machines)`).

Data:

`score`	Productivity score achieved on a machine (measured 3 times)
`Machine`	Type of machine used (3 levels)
`Worker`	Subject used in the experiment (6 levels)

Interest:

1. Expected productivity score of the three different machines
2. Is there an interdependence of worker and machine?
3. Variation among workers (= between-worker variability, $\hat{\sigma}_b^2$)
4. Variation within a worker (= within-worker variability, $\hat{\sigma}^2$)

Two factors: 1 fix / 1 random – the Machines dataset

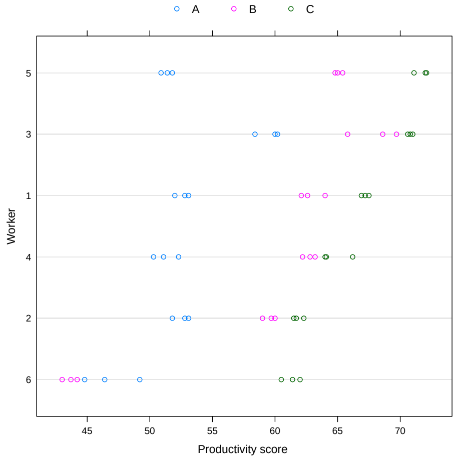


Figure: Plot of the $6 \times 3 \times 3$ data points, grouped by `Worker` (lines) in ascending order of the mean score per worker. The measurements for the three different machines (A, B, C) are depicted in different colors, the three replications as dots. Overall the between-worker variability is not that large, while there is considerable variability between machines.

Is there an interaction between worker and machines?

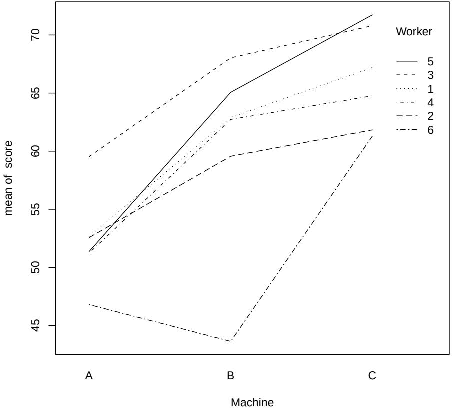


Figure: As can be seen, worker 6 scored differently on machine B compared to the others.

Two factors: 1 fix / 1 random – the Machines dataset

Model: $$y_{ijk} = \mu + \beta_j + b_i + b_{ij} + \varepsilon_{ijk}$$

with

- y_{ijk} response variable: productivity score of worker i on machine j at k^{th} measurement
- μ mean score across all 6 workers (fixed effect)
- β_j mean productivity score for each machine (fixed effect)
- b_i random effect of worker (= deviation from β_j)
- b_{ij} interaction term worker $\times$ machine (random due to worker)
- ε_{ijk} error term

Assumptions:

1. $b_i \sim \mathcal{N}\langle 0, \sigma_1^2 \rangle$ ("between-worker variability")
2. $b_{ij} \sim \mathcal{N}\langle 0, \sigma_2^2 \rangle$ (interaction)
3. $\varepsilon_{ijk} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$ (error term or "within-worker variability")
4. $b_i \perp b_{ij}, \quad b_i, b_{ij} \perp \varepsilon_{ij}$

Two factors: 1 fix / 1 random – the Machines dataset

Output (lmer):

```
Linear mixed model fit by REML
Formula: score ~ Machine + (1 | Worker) + (1 | Worker:Machine)
   Data: machine
 REML
215.7

Random effects:
 Groups         Name        Variance Std.Dev.
 Worker:Machine (Intercept) 13.90945 3.72954
 Worker         (Intercept) 22.85851 4.78106
 Residual                    0.92463 0.96158
Number of obs: 54, groups: Worker:Machine, 18; Worker, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)   52.356      2.486  21.062
MachineB       7.967      2.177   3.660
MachineC      13.917      2.177   6.393

Correlation of Fixed Effects:
         (Intr) MachnB
MachineB -0.438
MachineC -0.438  0.500
```

Two factors: 1 fix / 1 random – the Machines dataset

Output (`lmer`):

```
Random effects:
 Groups         Name        Variance Std.Dev.
 Worker:Machine (Intercept) 13.90945 3.72954
 Worker         (Intercept) 22.85851 4.78106
 Residual                    0.92463 0.96158
Number of obs: 54, groups: Worker:Machine, 18; Worker, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)   52.356      2.486  21.062
MachineB       7.967      2.177   3.660
MachineC      13.917      2.177   6.393
```

where $\hat{\beta}_j$ = Estimates under `Fixed effects:`

$\hat{\sigma}_1$ = `Worker 4.78106` (between-worker variability)

$\hat{\sigma}_2$ = `Worker:Machine 3.72954` (interaction term)

$\hat{\sigma}$ = `Residual: 0.96158` (within-worker variability)

Two factors: both random – the `manager1` dataset

Example for a model with two random effects in a **crossed** design. (The dataset `manager1` is constructed.)

Background	A behavioural researcher has devised a method for evaluating "managerial style" by observing the ordinary workday interactions of managers and rating certain kinds of occurrences. Because the evaluation method is to be applied in the field by many different evaluators, it is important to find out whether ratings vary much or little from one trained evaluator to another.
Interest:	Variation among raters (= between-rater variability) with the main question: "can the test be generalized"?
Design:	Each rater evaluated all 20 managers (crossed design).

Data:

`score`	Score on an observational rating scale (measured once)
`Rater`	(3 levels)
`Manager`	Subject evaluated in the experiment (20 levels)

Two factors: both random – the `manager1` dataset

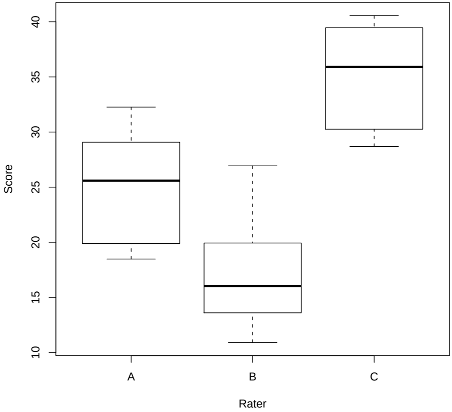


Figure: Boxplots of the 20 manager ratings separated by the three raters. The raters seem to evaluate the managers differently.

Two factors: both random – the `manager1` dataset

Model: $$y_{ij} = \mu + b_{i1} + b_{j2} + \varepsilon_{ij}$$

with

y_{ij} response variable: score given by rater i for manager j

μ mean score across all 20 managers (= estimate of the population mean of all managers) (fixed effect)

b_{i1} random effect of rater i (= deviation from μ)

b_{j2} random effect of manager j (= deviation from μ)

ε_{ij} error term

Assumptions:

1. $b_{i1} \sim \mathcal{N}\langle 0, \sigma_1^2 \rangle$ ("between-rater variability")
2. $b_{j2} \sim \mathcal{N}\langle 0, \sigma_2^2 \rangle$ ("between-manager variability")
3. $\varepsilon_{ij} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$ (error term or "within-rater variability")
4. $b_{i1} \perp b_{j2}, \quad b_{i1}, b_{j2} \perp \varepsilon_{ij}$

Two factors: both random – the manager1 dataset

Output (lmer):

```
Linear mixed model fit by REML
Formula: score ~ 1 + (1 | Rater) + (1 | Manager)
   Data: manager1

Random effects:
 Groups   Name        Variance Std.Dev.
 Manager  (Intercept)  0.000   0.0000
 Rater    (Intercept) 79.561   8.9197
 Residual             23.323   4.8293
Number of obs: 60, groups: Manager, 20; Rater, 3

Fixed effects:
            Estimate Std. Error t value
(Intercept)   25.756      5.187   4.965
```

where

$$\hat{\mu} = \texttt{(Intercept) 25.756}$$

$$\hat{\sigma}_1 = \texttt{Rater 8.9197} \qquad \text{(between-rater variability)}$$

$$\hat{\sigma}_2 = \texttt{Manager 0.0000} \qquad \text{(between-manager variability)}$$

$$\hat{\sigma} = \texttt{Residual: 4.8293}$$

Note: $\hat{\sigma}_2 = 0$ indicates that the "between-manager" variability is not sufficient to warrant incorporating Manager as a random effect into the model.

Two factors: both random – the pastes dataset

Example for a model with two random effects in a **nested** design. (The dataset pastes is included in the package `lme4a` (R: `data(Pastes)`).

Data:

`strength`	Concentration of a chemical product (measured 2 times)
`batch`	Batch from which the quality control samples were drawn (10 levels)
`sample`	Samples taken in order to measure the strength (3 levels)

Design: The samples are nested within the batches they are drawn from.

Interest:

1. Expected strength of paste
2. Variation among batches (= between-batch variability, $\hat{\sigma}^2_{b1}$)
3. Variation among samples (= between-samples variability, $\hat{\sigma}^2_{b2}$)
4. Variation within a sample (= within-sample variability, $\hat{\sigma}^2$)

Two factors: both random – the pastes dataset

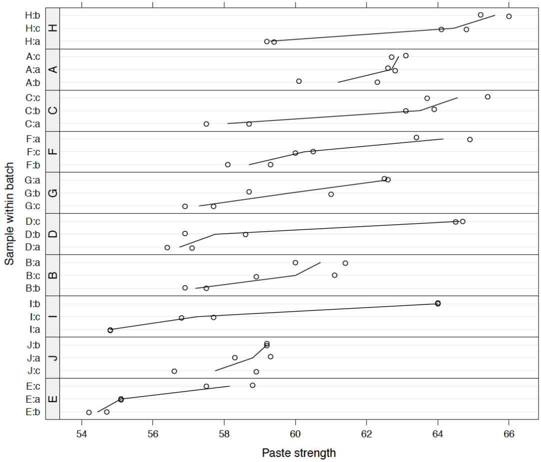


Figure: Strength of paste preparations according to `batch` and `sample` within the batch. There is large variability between the samples, the batches do not differ that much.

Two factors: both random – the pastes dataset

Model: $$y_{ijk} = \mu + b_i + b_{j(i)} + \varepsilon_{ijk}$$

with

- y_{ijk} response variable: strength of paste (k^{th} measurement) in sample j taken from batch i
- μ mean score across all 10 batches (= estimate of the population mean of all batches) (fixed effect)
- b_i random effect of batch i (= deviation from μ)
- $b_{j(i)}$ random effect of sample j taken from batch i (= deviation from μ)
- ε_{ijk} error term

Assumptions:

1. $b_i \sim \mathcal{N}\langle 0, \sigma_1^2 \rangle$ ("between-batch variability")
2. $b_{j(i)} \sim \mathcal{N}\langle 0, \sigma_2^2 \rangle$ ("between-sample variability")
3. $\varepsilon_{ijk} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$ (error term or "within-sample variability")
4. $b_i \perp b_{j(i)}, \quad b_1, b_{j(i)} \perp \varepsilon_{ijk}$

Two factors: both random – the pastes dataset

Output (lmer):

```
Linear mixed model fit by REML
Formula: strength ~ 1 + (1 | batch) + (1 | sample)
   Data: Pastes
REML
 247

Random effects:
 Groups   Name        Variance Std.Dev.
 sample   (Intercept) 8.4337   2.9041
 batch    (Intercept) 1.6573   1.2874
 Residual             0.6780   0.8234
Number of obs: 60, groups: sample, 30; batch, 10

Fixed effects:
            Estimate Std. Error t value
(Intercept)  60.0533     0.6769   88.72
```

where

$$\hat{\mu} = \texttt{(Intercept) 60.0533}$$
$$\hat{\sigma}_1 = \texttt{batch 1.2874} \quad \text{(between-batch variability)}$$
$$\hat{\sigma}_2 = \texttt{sample 2.9041} \quad \text{(between-sample variability)}$$
$$\hat{\sigma} = \texttt{Residual: 0.8234}$$

A complex model – the `oats` dataset

Example for a "split-plot experiment" with two fixed and one random effect, an interaction between the fixed effects and a fixed factor nested within a random factor. (The dataset `oates` is included in the package `lme4a` (`R: data(Pastes)`).

Data:

`yield`	yield of oats (measured once)
`nitro`	Nitrogen-fertilizer concentration used (4 levels)
`Variety`	Different crop varieties (3 levels)
`Block`	Large field that was splitted into 12 subplots for all possible combinations of crop varieties and nitrogen fertilizer (6 levels)

Design:

- Next to the fixed and random effects there was an interest whether crop variety and fertilizer concentration interact.
- The division of the block into plots and subplots was modelled with a nested design.

A complex model – the `oats` dataset

Interest:

1. Expected yield of the different oat varieties depending on fertilizer concentration
2. Variation among blocks (= between-block variability, $\hat{\sigma}_1^2$)
3. Variation among crop varieties (= between-variety variability, $\hat{\sigma}_2^2$)
4. General variation ($\hat{\sigma}^2$)

A complex model – the oats dataset

Model: $$y_{ijk} = \mu + \beta_{j1} + \beta_{k2} + \beta_{jk3} + b_i + b_{k(i)} + \varepsilon_{ijk}$$

with

- y_{ijk} response variable: yield of crop variety k in block i, treated with nitrogen-level j
- μ mean score across all 6 blocks (= estimate of the population mean of all blocks) (fixed effect)
- β_1 nitrogen-concentration (fixed effect)
- β_2 crop variety (fixed effect)
- β_3 interaction nitrogen-conc. $\times$ crop variety (fixed effect)
- b_i random effect of block i (= deviation from μ)
- $b_{k(i)}$ random effect of variety within block (= deviation from μ)
- ε_{ijk} error term

Assumptions:

1. $b_i \sim \mathcal{N}\langle 0, \sigma_1^2 \rangle$ ("between-block variability")
2. $b_{k(i)} \sim \mathcal{N}\langle 0, \sigma_2^2 \rangle$ ("between-variety variability")
3. $\varepsilon_{ijk} \sim \mathcal{N}\langle 0, \sigma^2 \rangle$ (error term)
4. $b_i \perp b_{k(i)}, \quad b_1, b_{k(i)} \perp \varepsilon_{ijk}$

A complex model – the oats dataset

Output (`lmer`):

```
Linear mixed model fit by REML
Formula: yield ~ ordered(nitro) * Variety + (1 | Block) + (1 | Block/Variety)
   Data: Oats
 REML
533.2

Random effects:
 Groups        Name        Variance Std.Dev.
 Variety:Block (Intercept) 106.06   10.299
 Block         (Intercept) 107.24   10.356
 Block         (Intercept) 107.24   10.356
 Residual                  177.08   13.307
Number of obs: 72, groups: Variety:Block, 18; Block, 6
```

where $\hat{\sigma}_1$ = `Block 10.356` (between-batch variability)

$\hat{\sigma}_2$ = `Variety:Block 10.299` (between-variety variability)

$\hat{\sigma}$ = `Residual: 13.307`

β_i fixed effects omitted