

Ex 1, 1-Way ANOVA with a random effect factor

Y vs. One random effect term

Four leaves were randomly sampled from a larger population of leaves, and four discs were taken from each leaf. The CALCIUM Contents were measured.

<< Minitab printout >>

```
MTB > print c1 c2 c3
Row  LEAF  DISC  CACONC
 1      1      1  0.4567
 2      1      2  0.6033
 3      1      3  0.3918
 4      1      4  0.3016
 5      2      1  0.7291
 6      2      2  0.6527
 7      2      3  0.5165
 8      2      4  0.5704
 9      3      1  0.7727
10      3      2  0.6171
11      3      3  0.7330
12      3      4  0.8495
13      4      1  0.3526
14      4      2  0.4497
15      4      3  0.3459
16      4      4  0.5446
```

```
MTB > glm CACONC = LEAF;
SUBC> random LEAF;
SUBC> ems.
```

General Linear Model: CACONC versus LEAF

Factor	Type	Levels	Values
LEAF	random	4	1, 2, 3, 4

Analysis of Variance for CACONC, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
LEAF	3	0.28086	0.28086	0.09362	8.70	0.002
Error	12	0.12910	0.12910	0.01076		
Total	15	0.40996				

S = 0.103722 R-Sq = 68.51% R-Sq(adj) = 60.64%

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 LEAF	(2) + 4.0000 (1)
2 Error	(2)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 LEAF	12.00	0.01076	(2)

Variance Components, using Adjusted SS

Source	Estimated Value
LEAF	0.02072
Error	0.01076

1). H_0 : no significant variance difference among LEAVES. $F = \frac{0.09362}{0.01076} = 8.7 \sim F_{3,12}$, p -value=0.002

2). $\hat{\sigma}_e^2 = 0.01076$ (= variance among discs), and $\hat{\sigma}_{leaf}^2 = 0.02072$ (= variance among leaves). This means that the variation among leaves is greater than the variation within leaves. Future experiments could be improved by using more leaves.

<< R printout: lmer>>

```
datal <-
read.csv("D:\\STAT999\\Minitab_p239.csv")
```

<< R printout: lme>>

```
modell_lme <- lme(CACONC ~ 1, random=~1|LEAF/DISC)
```

```
dim(data1)
[1] 16 3
data1[1:3,]
  LEAF DISC CACONC
1     1    1 0.4567
2     1    2 0.6033
3     1    3 0.3918
attach(data1)
library(lme4)
modell1 <-
lmer(CACONC ~ (1|LEAF))
# You get exactly identical printout from
# modell1 <- lmer(CACONC ~ 1 + (1|LEAF))
summary(modell1)

Linear mixed model fit by REML ['lmerMod']
Formula: CACONC ~ (1 | LEAF)

REML criterion at convergence: -16.1496

Random effects:
Groups   Name              Variance Std.Dev.
LEAF     (Intercept)    0.02072   0.1439
Residual                    0.01076   0.1037
Number of obs: 16, groups: LEAF, 4

Fixed effects:
              Estimate Std. Error t value
(Intercept)   0.55545    0.07649    7.261
# 0.55545=mean(CACONC)

modell1_lmer <- lmer(CACONC ~ (1|LEAF))
modell2_lmer <- lmer(CACONC ~
(1|LEAF)+(1|LEAF:DISC))
Error in checkNlevels(reTrms$flist, n = n,
control) :
number of levels of each grouping factor
must be < number of observations
modell2_lmer <- lmer(CACONC ~
(1|LEAF)+(DISC|LEAF))
summary(modell2_lmer)

Linear mixed model fit by REML ['lmerMod']
Formula: CACONC ~ (1 | LEAF) + (DISC | LEAF)

REML criterion at convergence: -16.5942

Random effects:
Groups   Name              Variance Std.Dev. Corr
LEAF     (Intercept)    0.011491 0.10719
LEAF.1    (Intercept)    0.011213 0.10589
          DISC           0.001182 0.03438 -0.49
Residual                    0.008788 0.09375
Number of obs: 16, groups: LEAF, 4

Fixed effects:
              Estimate Std. Error t value
(Intercept)   0.56007    0.07555    7.413
```

To compare 2 models:

```
modell1_lmer <- lmer(CACONC ~
(1|LEAF),REML=F)
modell2_lmer <- lmer(CACONC ~
(1|LEAF)+(DISC|LEAF),REML=F)
anova(modell1_lmer,model2_lmer)
Data:
Models:
modell1_lmer: CACONC ~ (1 | LEAF)
modell2_lmer: CACONC ~ (1 | LEAF) + (DISC |
LEAF)
   Df    AIC     BIC logLik deviance Chisq ChiDf p-val
mod1 3 -13.604 -11.286  9.802  -19.604
mod2 6  -8.073  -3.438 10.037  -20.073 0.4694   3 0.9256
```

```
summary(modell1_lme)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC   logLik
-8.149635 -5.317434 8.074817

Random effects:
Formula: ~1 | LEAF
(Intercept)
StdDev:    0.1439292

Formula: ~1 | DISC %in% LEAF
(Intercept) Residual
StdDev: 0.09362071 0.04464883

Fixed effects: CACONC ~ 1
              Value Std.Error DF t-value p-value
(Intercept) 0.55545 0.07649378 12 7.261374      0

Number of Observations: 16
Number of Groups:
      LEAF DISC %in% LEAF
      4      16

modell2_lme <-
lme(CACONC ~ 1,random=~1|LEAF)
summary(modell2_lme)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC   logLik
-10.14963 -8.025484 8.074817

Random effects:
Formula: ~1 | LEAF
(Intercept) Residual
StdDev:    0.1439292 0.1037225

Fixed effects: CACONC ~ 1
              Value Std.Error DF t-value p-value
(Intercept) 0.55545 0.07649378 12 7.261374      0

Number of Observations: 16
Number of Groups: 4
```

To compare 2 models:

```
> modell1_lme <- lme(CACONC ~ 1,random=~1|LEAF/DISC,
method="ML")
> modell2_lme <- lme(CACONC ~ 1,random=~1|LEAF,
method="ML")
> anova(modell1_lme, modell2_lme)
Model df    AIC      BIC logLik Test L.Ratio p-value
m1_lme 1  4 -11.6036  -8.5132 9.8018
m2_lme 2  3 -13.6036 -11.2858 9.8018 1 vs 2 <0.001   1
```

Ex 2, 2-level nested ANOVA with random effect factors

Y vs. random1 + random2 (nested in random1)

Four PLANTS were randomly sampled from a larger population of PLANTS, and three leaves were randomly taken from each plant, and two discs were taken from each leaf. The CALCIUM Contents were measured.

<< Minitab printout >>

MTB > **print c1 c2 c3 c4**

Row	CACONC2	DISC2	LEAF2	PLANT2
1	1.6647	1	1	1
2	2.0276	2	1	1
3	2.2036	1	2	1
4	2.2238	2	2	1
5	1.9945	1	3	1
6	1.9045	2	3	1
7	1.9579	1	1	2
8	1.7441	2	1	2
9	2.4959	1	2	2
10	2.5570	2	2	2
11	1.8732	1	3	2
12	1.9046	2	3	2
13	2.2992	1	1	3
14	2.6597	2	1	3
15	2.7374	1	2	3
16	2.9178	2	2	3
17	2.8447	1	3	3
18	2.4489	2	3	3
19	2.6918	1	1	4
20	2.7402	2	1	4
21	3.3845	1	2	4
22	2.8510	2	2	4
23	2.2435	1	3	4
24	2.7079	2	3	4

MTB > **glm CACONC2 = PLANT2 + LEAF2(PLANT2);**

SUBC> **random PLANT2 LEAF2;**

SUBC> **ems.**

General Linear Model: CACONC2 versus PLANT2, LEAF2

Factor	Type	Levels	Values
PLANT2	random	4	1, 2, 3, 4
LEAF2(PLANT2)	random	12	1, 2, 3, 1, 2, 3, 1, 2, 3, 1, 2, 3

Analysis of Variance for CACONC2, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
PLANT2	3	2.71433	2.71433	0.90478	5.74	0.022
LEAF2(PLANT2)	8	1.26200	1.26200	0.15775	3.74	0.020
Error	12	0.50621	0.50621	0.04218		
Total	23	4.48254				

S = 0.205389 R-Sq = 88.71% R-Sq(adj) = 78.36%

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 PLANT2	(3) + 2.0000 (2) + 6.0000 (1)
2 LEAF2(PLANT2)	(3) + 2.0000 (2)
3 Error	(3)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 PLANT2	8.00	0.15775	(2)
2 LEAF2(PLANT2)	12.00	0.04218	(3)

Variance Components, using Adjusted SS

Source	Estimated Value
PLANT2	0.12450

LEAF2 (PLANT2)	0.05778
Error	0.04218

1). H_0 : no significant variance difference among PLANTS. $F = \frac{0.90478}{0.15775} = 5.74 \sim F_{3,8}$, p -value=0.022

H_0 : no significant variance difference among LEAVES. $F = \frac{0.15775}{0.04218} = 3.74 \sim F_{8,12}$, p -value=0.020

2). $\hat{\sigma}_e^2 = 0.04218$ (= variance among discs within a leaf), $\hat{\sigma}_{leaf}^2 = 0.05778$ (= variance among leaves within a plant). and $\hat{\sigma}_{plant}^2 = 0.12450$ (= variance among leaves). This means that the variation among plants is much greater than the variation among leaves. Future experiments could be improved by using more plants.

<< R printout >>

```
data1 <- read.csv("D:\\STAT999\\Minitab_p243.csv")
dim(data1)
[1] 24 4
data1
  CACONC2 DISC2 LEAF2 PLANT2
1  1.6647     1     1      1
2  2.0276     2     1      1
3  2.2036     1     2      1
4  2.2238     2     2      1
5  1.9945     1     3      1
6  1.9045     2     3      1
7  1.9579     1     1      2
8  1.7441     2     1      2
9  2.4959     1     2      2
10 2.5570     2     2      2
11 1.8732     1     3      2
12 1.9046     2     3      2
13 2.2992     1     1      3
14 2.6597     2     1      3
15 2.7374     1     2      3
16 2.9178     2     2      3
17 2.8447     1     3      3
18 2.4489     2     3      3
19 2.6918     1     1      4
20 2.7402     2     1      4
21 3.3845     1     2      4
22 2.8510     2     2      4
23 2.2435     1     3      4
24 2.7079     2     3      4
attach(data1)

model2 <- lmer(CACONC2 ~ (1 | PLANT2) + (1 | PLANT2:LEAF2))
summary(model2)
Linear mixed model fit by REML ['lmerMod']
Formula: CACONC2 ~ (1 | PLANT2) + (1 | PLANT2:LEAF2)

REML criterion at convergence: 15.3866

Random effects:
Groups          Name              Variance Std.Dev.
PLANT2:LEAF2 (Intercept)  0.05778  0.2404
PLANT2         (Intercept)  0.12450  0.3529
Residual                        0.04218  0.2054
Number of obs: 24, groups: PLANT2:LEAF2, 12; PLANT2, 4

Fixed effects:
              Estimate Std. Error t value
(Intercept)    2.3783    0.1942    12.25

model2 <- lmer(CACONC2 ~ (1 | PLANT2) + (1 | PLANT2:LEAF2), REML=F)
model3 <- lmer(CACONC2 ~ (1 | PLANT2:LEAF2), REML=F)
anova(model2, model3)
Data:
Models:
model3: CACONC2 ~ (1 | PLANT2:LEAF2)
model2: CACONC2 ~ (1 | PLANT2) + (1 | PLANT2:LEAF2)
      Df    AIC    BIC logLik deviance Chisq Chi Df Pr(>Chisq)
model3  3 22.866 26.400 -8.4330  16.866
```

```

model2  4 21.796 26.508 -6.8978 13.796 3.0704 1 0.07973 .
---

library(nlme)

model4 <- lme(CACONC2 ~ 1, random=~1|PLANT2/LEAF2) # lmer is better.
summary(model4)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC    logLik
23.38662 27.92859 -7.693308

Random effects:
Formula: ~1 | PLANT2
(Intercept)
StdDev: 0.3528517

Formula: ~1 | LEAF2 %in% PLANT2
(Intercept) Residual
StdDev: 0.2403798 0.2053887

Fixed effects: CACONC2 ~ 1
              Value Std.Error DF t-value p-value
(Intercept) 2.37825 0.1941622 12 12.24878 0

Number of Observations: 24
Number of Groups:
      PLANT2 LEAF2 %in% PLANT2
      4      12
> model4 <- lme(CACONC2 ~ 1, random=~LEAF2|PLANT2)
Error in lme.formula(CACONC2 ~ 1, random = ~LEAF2 | PLANT2) :
nlminb problem, convergence error code = 1
message = iteration limit reached without convergence (10)

```

Ex 3, Mixed effect model

Y vs. fixed + random (nested in fixed)

A pig farmer wants to compare the genetic quality of male pigs. He has five sires and 10 dams, and allocates two dams to each sire. The weight of two piglets per litter is then monitored and their weight gain over a 2-week period was measured.

- **SIRE** = fixed effect factor
- **DAM** = random effect factor nested within SIRE

<< Minitab printout >>

```

MTB > print c1 c2 c3
Row  SIRE  DAM  WGAIN
  1    1    1  0.671
  2    1    1  0.875
  3    1    2  1.948
  4    1    2  1.668
  5    2    1  0.797
  6    2    1  0.954
  7    2    2  0.927
  8    2    2  1.098
  9    3    1  1.301
 10    3    1  1.844
 11    3    2  1.413
 12    3    2  1.426
 13    4    1  0.938
 14    4    1  0.490
 15    4    2  0.604
 16    4    2  1.478
 17    5    1  1.298
 18    5    1  0.855
 19    5    2  1.415
 20    5    2  1.262

MTB > glm WGAIN = SIRE + DAM(SIRE);
SUBC> random DAM;
SUBC> ems.

```

General Linear Model: WGAIN versus SIRE, DAM

Factor	Type	Levels	Values
--------	------	--------	--------

```

SIRE      fixed      5  1, 2, 3, 4, 5
DAM(SIRE) random    10  1, 2, 1, 2, 1, 2, 1, 2, 1, 2

Analysis of Variance for WGAIN, using Adjusted SS for Tests

Source      DF      Seq SS      Adj SS      Adj MS      F      P
SIRE        4      1.03439    1.03439    0.25860    1.00   0.484
DAM(SIRE)   5      1.28898    1.28898    0.25780    3.12   0.059
Error       10      0.82658    0.82658    0.08266
Total       19      3.14994

```

S = 0.287503 R-Sq = 73.76% R-Sq(adj) = 50.14%

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 SIRE	(3) + 2.0000 (2) + Q[1]
2 DAM(SIRE)	(3) + 2.0000 (2)
3 Error	(3)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 SIRE	5.00	0.25780	(2)
2 DAM(SIRE)	10.00	0.08266	(3)

Variance Components, using Adjusted SS

Source	Estimated Value
DAM(SIRE)	0.08757
Error	0.08266

1). H_0 : no significant mean difference among SIREs. $F = \frac{0.25860}{0.25780} = 1.00 \sim F_{4,5}$, p -value=0.484

H_0 : no significant variance difference among DAMS. $F = \frac{0.25780}{0.08266} = 3.12 \sim F_{5,10}$, p -value=0.059

2). $\hat{\sigma}_e^2 = 0.08266$ (= variance among piglets from the same DAD), and $\hat{\sigma}_{dam(sire)}^2 = 0.08757$ (= variance among DAMS mated to the same SIRE).

<< R printout >>

```
data1 <- read.csv("D:\\STAT999\\Minitab_p245.csv")
```

```
data1
```

```

      SIRE DAM WGAIN
1       1   1 0.671
2       1   1 0.875
3       1   2 1.948
4       1   2 1.668
5       2   1 0.797
6       2   1 0.954
7       2   2 0.927
8       2   2 1.098
9       3   1 1.301
10      3   1 1.844
11      3   2 1.413
12      3   2 1.426
13      4   1 0.938
14      4   1 0.490
15      4   2 0.604
16      4   2 1.478
17      5   1 1.298
18      5   1 0.855
19      5   2 1.415
20      5   2 1.262

```

```
attach(data1)
```

```
modell <- lmer(WGAIN~SIRE+(1 | SIRE:DAM), REML=F)
```

```
summary(modell)
```

```
Linear mixed model fit by maximum likelihood ['lmerMod']
```

```
Formula: WGAIN ~ SIRE + (1 | SIRE:DAM)
```

	AIC	BIC	logLik	deviance
	25.1379	29.1208	-8.5690	17.1379

Random effects:

Groups	Name	Variance	Std.Dev.
	SIRE:DAM (Intercept)	0.07376	0.2716
	Residual	0.08266	0.2875

Number of obs: 20, groups: SIRE:DAM, 10

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	1.23285	0.25159	4.900
SIRE	-0.02325	0.07586	-0.306

Correlation of Fixed Effects:

	(Intr)
SIRE	-0.905

Different from Minitab

```

model2 <- lmer(WGAIN ~ (1|SIRE:DAM), REML=F)
anova(model1, model2)
Data:
Models:
model2: WGAIN ~ (1 | SIRE:DAM)
model1: WGAIN ~ SIRE + (1 | SIRE:DAM)

```

	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
model2	3	23.231	26.219	-8.6157	17.231				
model1	4	25.138	29.121	-8.5690	17.138	0.0935	1		0.7598

```

> model1 <- lmer(WGAIN~SIRE+(1 | SIRE:DAM) )
> summary(model1)
Linear mixed model fit by REML ['lmerMod']
Formula: WGAIN ~ SIRE + (1 | SIRE:DAM)

REML criterion at convergence: 22.8698

Random effects:
Groups      Name      Variance Std.Dev.
SIRE:DAM (Intercept) 0.10253  0.3202
Residual      0.08266  0.2875
Number of obs: 20, groups: SIRE:DAM, 10

Fixed effects:
              Estimate Std. Error t value
(Intercept)  1.23285    0.28129   4.383
SIRE         -0.02325    0.08481  -0.274

Correlation of Fixed Effects:
      (Intr)
SIRE  -0.905

> model2 <- lmer(WGAIN~(1 | SIRE:DAM) )
> summary(model2)
Linear mixed model fit by REML ['lmerMod']
Formula: WGAIN ~ (1 | SIRE:DAM)

REML criterion at convergence: 19.7971

Random effects:
Groups      Name      Variance Std.Dev.
SIRE:DAM (Intercept) 0.08775  0.2962
Residual      0.08266  0.2875
Number of obs: 20, groups: SIRE:DAM, 10

Fixed effects:
              Estimate Std. Error t value
(Intercept)  1.1631     0.1136   10.24

> model2_lme <- lme(WGAIN~SIRE, random=~1 | SIRE/DAM)
> summary(model2_lme)
Linear mixed-effects model fit by REML
Data: NULL

```

Same as Minitab

```

      AIC      BIC    logLik
32.80024 37.2521 -11.40012

Random effects:
Formula: ~1 | SIRE
(Intercept)
StdDev:    0.1412382

Formula: ~1 | DAM %in% SIRE
(Intercept) Residual
StdDev:      0.2959199 0.2875032 #0.2959199^2=0.08756859, 0.2875032^2=0.08265809

Fixed effects: WGAIn ~ SIRE
              Value Std.Error DF   t-value p-value
(Intercept)  1.23285  0.30469123 10   4.046227  0.0023
SIRE         -0.02325  0.09186786   3  -0.253081  0.8166
Correlation:
(Intr)
SIRE -0.905

Number of Observations: 20
Number of Groups:
      SIRE DAM %in% SIRE
      5      10
> 0.2959199^2
[1] 0.08756859

> modell_lme <- lme(WGAIn ~ SIRE, random=~DAM|SIRE)
> summary(modell_lme)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC    logLik
34.72079 40.06302 -11.36039

Random effects:
Formula: ~DAM | SIRE
Structure: General positive-definite, Log-Cholesky parametrization
              StdDev  Corr
(Intercept)  0.7352110 (Intr)
DAM          0.4184934 -0.943
Residual     0.2875038

Fixed effects: WGAIn ~ SIRE
              Value Std.Error DF   t-value p-value
(Intercept)  1.3101279  0.30068101 15   4.357202  0.0006
SIRE         -0.0376362  0.09065874   3  -0.415141  0.7059
Correlation:
(Intr)
SIRE -0.905

Number of Observations: 20
Number of Groups: 5

```

Ex 4, Repeated Measures (Mixed-effects models with temporal pseudoreplication)

Y vs. fixed + random (nested in fixed)

- **fertilizer** = fixed effect factor (added vs. control)
- **plant** = random effect factor (12 plants)
- **week** = each plant was measured at week=2,4,6,8,10.

<< R printout >>

```

> data1 <- read.table("D:\\STAT999\\RBOOK\\fertilizer.txt",header=T)
> dim(data1)
[1] 60 4
> data1[1:3,]
  root week plant fertilizer
1   1.3    2  ID1      added
2   3.5    4  ID1      added
3   7.0    6  ID1      added
> attach(data1)
> table(fertilizer, plant)

```



```

      plant
fertilizer ID1 ID10 ID11 ID12 ID2 ID3 ID4 ID5 ID6 ID7 ID8 ID9
added      5     0     0     0  5   5   5   5   5   0   0   0
control    0     5     5     5  0   0   0   0   0   5   5   5
> table(week, plant)
      plant
week ID1 ID10 ID11 ID12 ID2 ID3 ID4 ID5 ID6 ID7 ID8 ID9
  2    1    1    1    1    1    1    1    1    1    1    1    1
  4    1    1    1    1    1    1    1    1    1    1    1    1
  6    1    1    1    1    1    1    1    1    1    1    1    1
  8    1    1    1    1    1    1    1    1    1    1    1    1
 10    1    1    1    1    1    1    1    1    1    1    1    1
> WEEK <- factor(week)
> model1 <-

```

```
lme(root~fertilizer,random=~week|plant)
```

```

#week=continuous, plant=categorical. ERROR if you declare week as "factor" in advance
#ERROR if you enter "random=~1/plant/week"

```

```

> summary(model1)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC    logLik
171.0236 183.3863 -79.51181

Random effects:
Formula: ~week | plant
Structure: General positive-definite, Log-Cholesky parametrization
              StdDev      Corr
(Intercept)  2.8639832 (Intr)
week          0.9369412 -0.999
Residual      0.4966308

Fixed effects: root ~ fertilizer
              Value Std.Error DF   t-value p-value
(Intercept)   2.799710 0.1438367 48 19.464500 0e+00
fertilizercontrol -1.039383 0.2034158 10 -5.109644 5e-04
Correlation:
              (Intr)
fertilizercontrol -0.707

Number of Observations: 60
Number of Groups: 12

```

```

> model2 <-
lmer(root~fertilizer+(1|plant)+(week|plant))

```

Two printouts are identical

```

> summary(model2)
Linear mixed model fit by REML ['lmerMod']
Formula: root ~ fertilizer + (1 | plant) + (week | plant)

REML criterion at convergence: 159.0236

Random effects:
Groups      Name      Variance Std.Dev. Corr
plant      (Intercept) 0.02162  0.1470
plant.1    (Intercept) 8.18077  2.8602
           week        0.87786  0.9369  -1.00
Residual              0.24664  0.4966

Number of obs: 60, groups: plant, 12

Fixed effects:
              Estimate Std. Error t value
(Intercept)    2.7997    0.1438   19.46
fertilizercontrol -1.0394    0.2034   -5.11

Correlation of Fixed Effects:
              (Intr)
fertilizercontrol -0.707
> model3 <- lmer(root~fertilizer+(week|plant))

```

```

> summary(model3)
Linear mixed model fit by REML ['lmerMod']

```

WRONG!

```
Formula: root ~ fertilizer + (week | plant)
```

```
REML criterion at convergence: 162.4251
```

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
plant	(Intercept)	0.0000	0.0000	
	week	0.8745	0.9352	NaN
Residual		0.2246	0.4739	

```
Number of obs: 60, groups: plant, 12
```

Fixed effects:

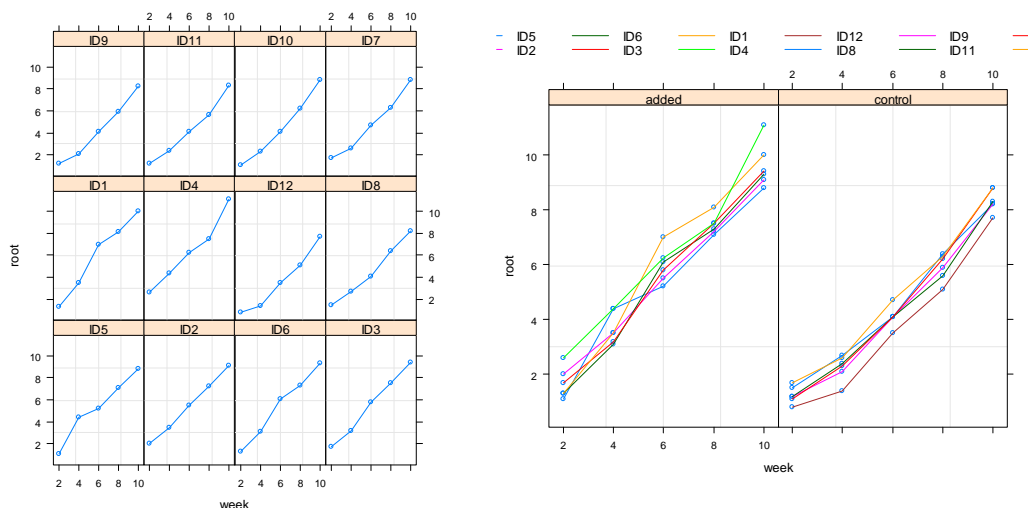
	Estimate	Std. Error	t value
(Intercept)	-0.1847	0.2024	-0.913
fertilizercontrol	-0.7612	0.2862	-2.660

Correlation of Fixed Effects:

```
(Intr)
ftrlzrcntrl -0.707
```

Benefits of “groupedData”

```
> data3 <- groupedData(root~week | plant, outer=~fertilizer)
> plot(data3)
> plot(data3, outer=T)
```



Available from “lme”

```
> plot(model1)
> plot(model1, root~fitted(.)) or > plot(model1, root~fitted(.)|plant)
> qqnorm(model1, ~resid(.))
> qqnorm(model1, ~resid(.)|plant)
```

Available from “lmer”

```
> plot(model2)
> plot(model2, root~fitted(.)) or > plot(model2, root~fitted(.)|plant)
```

Ex 4-1, One more Example

Y vs. fixed + random1 + random2 (nested in random1)

- **power** = fixed effect factor (categorical variable)
- **subject** = random effect factor (7 people)
- **eye** = nested within each person

```
<< R printout >>
```

```
> data(vision)
> attach(vision)
> dim(vision)
[1] 56 4
> vision
  acuity power eye subject
1    116  6/6 left      1
```

```

2      119  6/18  left      1
3      116  6/36  left      1
4      124  6/60  left      1
5      120   6/6  right     1
6      117  6/18  right     1
7      114  6/36  right     1
8      122  6/60  right     1

```

```

> summary(power)
6/6 6/18 6/36 6/60
 14  14  14  14

```

```
> is.factor(power)
```

```
[1] TRUE
```

```
> modell <-
```

```
lmer(acuity~power+(1 | subject)+(1 | subject:eye))
```

```
> summary(modell)
```

Linear mixed model fit by REML ['lmerMod']

Formula: acuity ~ power + (1 | subject) + (1 | subject:eye)

REML criterion at convergence: 328.7098

Random effects:

Groups	Name	Variance	Std.Dev.
subject:eye	(Intercept)	10.27	3.205
subject	(Intercept)	21.53	4.640
Residual		16.60	4.075

Number of obs: 56, groups: subject:eye, 14; subject, 7

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	112.6429	2.2349	50.40
power6/18	0.7857	1.5400	0.51
power6/36	-1.0000	1.5400	-0.65
power6/60	3.2857	1.5400	2.13

Correlation of Fixed Effects:

	(Intr)	pw6/18	pw6/36
power6/18	-0.345		
power6/36	-0.345	0.500	
power6/60	-0.345	0.500	0.500

```
> model2 <-
```

```
lme(acuity~power,random=~1 | subject/eye)
```

```
> summary(model2)
```

Linear mixed-effects model fit by REML

Data: NULL

	AIC	BIC	logLik
	342.7098	356.3686	-164.3549

Random effects:

Formula: ~1 | subject
(Intercept)

StdDev: 4.639616

Formula: ~1 | eye %in% subject
(Intercept) Residual

StdDev: 3.205242 4.074568

Fixed effects: acuity ~ power

	Value	Std.Error	DF	t-value	p-value
(Intercept)	112.64286	2.234914	39	50.40143	0.0000
power6/18	0.78571	1.540042	39	0.51019	0.6128
power6/36	-1.00000	1.540042	39	-0.64933	0.5199
power6/60	3.28571	1.540042	39	2.13352	0.0392

Correlation:

	(Intr)	pw6/18	pw6/36
power6/18	-0.345		
power6/36	-0.345	0.500	
power6/60	-0.345	0.500	0.500

Number of Observations: 56

Number of Groups:

```
subject eye %in% subject
       7       14
```

Two printouts are identical

Ex 4-2, One more Example

Repeated Measures (Mixed-effects models with spatial pseudoreplication)

Y vs. fixed + random (nested in fixed)

- `items` = fixed effect factor (ten grocery items)
- `store` = random effect factor (4 stores)

<< R printout >>

```
data1 <- read.table("http://ww2.coastal.edu/kingw/statistics/R-
tutorials/text/groceries.txt",header=T)
dim(data1)
[1] 10 5
head(data1)
  subject storeA storeB storeC storeD
1 lettuce  1.17  1.78  1.29  1.29
2 potatoes 1.77  1.98  1.99  1.99
3  milk    1.49  1.69  1.79  1.59
4  eggs    0.65  0.99  0.69  1.09
5  bread   1.58  1.70  1.89  1.89
6  cereal  3.13  3.15  2.99  3.09
data2 <- stack(data1)
Warning message:
In stack.data.frame(data1) : non-vector columns will be ignored
data2[3] <- rep(data1$subject,4)
head(data2)
  values ind      V3
1  1.17 storeA lettuce
2  1.77 storeA potatoes
3  1.49 storeA  milk
4  0.65 storeA  eggs
5  1.58 storeA  bread
6  3.13 storeA  cereal
colnames(data2) <- c("price","store","items")
head(data2)
  price store  items
1  1.17 storeA lettuce
2  1.77 storeA potatoes
3  1.49 storeA  milk
4  0.65 storeA  eggs
5  1.58 storeA  bread
6  3.13 storeA  cereal
```

```
modell1 <- aov(price~store+Error(items/store))
```

Cannot use "lm" because of "Error()"

```
summary(modell1)
```

```
Error: items
      Df Sum Sq Mean Sq F value Pr(>F)
Residuals  9  115.2    12.8

Error: items:store
      Df Sum Sq Mean Sq F value Pr(>F)
store   3  0.5859  0.19529   4.344 0.0127 *
Residuals 27  1.2137  0.04495
---
```

```
pairwise.t.test(price,store,p.adjust.method="holm",paired=T)
```

Pairwise comparisons using paired t tests

```
data: price and store
```

```

      storeA storeB storeC
storeB 0.17    -        -
storeC 0.17    0.69    -
storeD 0.07    0.49    0.33

```

P value adjustment method: holm

```
library(nlme)
```

```
model2 <-
```

```
lme(price~items,random=~1|items/store)
```

CORRECT

```
model2 <- lme(price~1,random=~items|store) # WRONG writing!
```

```
summary(model2)
```

Linear mixed-effects model fit by REML

Data: NULL

	AIC	BIC	logLik
	40.59027	58.80583	-7.295133

Random effects:

Formula: ~1 | items
(Intercept)

StdDev: 0.114664

Formula: ~1 | store %in% items
(Intercept) Residual

StdDev: 0.2293274 0.08599765

Fixed effects: price ~ items

	Value	Std.Error	DF	t-value	p-value
(Intercept)	4.8600	0.1677632	30	28.969405	0
itemsbread	-3.0950	0.2372530	0	-13.045146	NaN
itemscereal	-1.7700	0.2372530	0	-7.460391	NaN
itemsegg	-4.0050	0.2372530	0	-16.880715	NaN
itemsground.beef	-2.7225	0.2372530	0	-11.475092	NaN
itemslaundry.detergent	1.3550	0.2372530	0	5.711203	NaN
itemslettuce	-3.4775	0.2372530	0	-14.657350	NaN
itemsmilk	-3.2200	0.2372530	0	-13.572010	NaN
itemspotatoes	-2.9275	0.2372530	0	-12.339149	NaN
itemstomato.soup	-4.2075	0.2372530	0	-17.734234	NaN

OR

```
model22 <-
```

```
lme(price~items,random=~store|items)
```

```
summary(model22)
```

Linear mixed-effects model fit by REML

Data: NULL

	AIC	BIC	logLik
	47.1665	76.59164	-2.58325

Random effects:

Formula: ~store | items

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.1284731	(Intr) strstB strstC

storestoreB	0.2450132	0.000
-------------	-----------	-------

storestoreC	0.1939626	0.000	0.730
-------------	-----------	-------	-------

storestoreD	0.4504901	0.000	0.474	0.683
-------------	-----------	-------	-------	-------

Residual	0.0963548
----------	-----------

Fixed effects: price ~ items

	Value	Std.Error	DF	t-value	p-value
(Intercept)	4.545163	0.1553016	30	29.266683	0
itemsbread	-2.918642	0.2196296	0	-13.288925	NaN
itemscereal	-1.434368	0.2196296	0	-6.530848	NaN
itemsegg	-3.880588	0.2196296	0	-17.668781	NaN
itemsground.beef	-2.472678	0.2196296	0	-11.258400	NaN
itemslaundry.detergent	1.344670	0.2196296	0	6.122443	NaN
itemslettuce	-3.329650	0.2196296	0	-15.160295	NaN
itemsmilk	-3.002416	0.2196296	0	-13.670359	NaN
itemspotatoes	-2.736055	0.2196296	0	-12.457586	NaN

```
itemstomato.soup      -3.920490 0.2196296  0 -17.850460      Na
```

OR

```
> model3 <-
```

```
lmer(price~items+(1 | items)+(store | items))
```

```
> summary(model3)
```

```
Linear mixed model fit by REML ['lmerMod']
```

```
Formula: price ~ items + (1 | items) + (store | items)
```

```
REML criterion at convergence: 5.1665
```

```
Random effects:
```

Groups	Name	Variance	Std.Dev.	Corr
items	(Intercept)	0.003586	0.05988	
items.1	(Intercept)	0.066317	0.25752	
	storestoreB	0.066609	0.25809	-0.55
	storestoreC	0.044199	0.21024	-0.90 0.70
	storestoreD	0.209520	0.45773	-0.27 0.47 0.65
Residual		0.005995	0.07743	

```
Number of obs: 40, groups: items, 10
```

```
Fixed effects:
```

	Estimate	Std. Error	t value
(Intercept)	4.9664	0.1284	38.67
itemsbread	-3.0729	0.1816	-16.92
itemscereal	-1.9974	0.1816	-11.00
itemseggs	-4.3529	0.1816	-23.97
itemsground.beef	-2.9675	0.1816	-16.34
itemslaundry.detergent	0.8125	0.1816	4.47
itemslettuce	-3.6602	0.1816	-20.15
itemsmilk	-3.1280	0.1816	-17.22
itemspotatoes	-2.9706	0.1816	-16.36
itemstomato.soup	-4.3241	0.1816	-23.81

Ex 4-3, Yet another Example

Repeated Measures (Mixed-effects models with spatial pseudoreplication)

Y vs. fixed1 (nested in random) + fixed2 + random
i.e., fixed-effect (“test”) is nested in random-effect (“subject”)

- **items** = fixed effect factor (ten grocery items)
- **store** = random effect factor (4 stores)

```
<< R printout >>
```

```
data1 <- read.csv("C:\\STAT999\\hill.csv")
```

```
dim(data1)
```

```
[1] 36 4
```

```
data1
```

	subject	diet	test	SSS
1	A	chicken	pre	18
2	B	chicken	pre	13
3	C	chicken	pre	18
4	D	chicken	pre	15
5	E	chicken	pre	22
6	F	chicken	pre	32
7	G	chicken	pre	31
8	H	chicken	pre	24
9	I	chicken	pre	15
10	A	chicken	post	15
11	B	chicken	post	13
12	C	chicken	post	17
13	D	chicken	post	15
14	E	chicken	post	24
15	F	chicken	post	31
16	G	chicken	post	31

```

17      H chicken post 25
18      I chicken post 17
19      J  pasta  pre 17
20      K  pasta  pre 30
21      L  pasta  pre 18
22      M  pasta  pre 13
23      N  pasta  pre 23
24      O  pasta  pre 27
25      P  pasta  pre 27
26      Q  pasta  pre 24
27      R  pasta  pre 23
28      J  pasta post 19
29      K  pasta post 31
30      L  pasta post 18
31      M  pasta post 13
32      N  pasta post 24
33      O  pasta post 27
34      P  pasta post 26
35      Q  pasta post 28
36      R  pasta post 26

```

```
attach(data1)
```

```
modell1 <-
```

```
aov(SSS~diet*test+Error(subject/test))
```

```
summary(modell1)
```

```
Error: subject
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
diet	1	40.1	40.11	0.509	0.486
Residuals	16	1259.8	78.74		

```
Error: subject:test
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
test	1	2.778	2.778	2.174	0.16
diet:test	1	2.778	2.778	2.174	0.16
Residuals	16	20.444	1.278		

```
modell2 <-
```

```
lme(SSS~test+diet,random=~1|subject)
```

```
summary(modell2)
```

```
Linear mixed-effects model fit by REML
```

```
Data: NULL
```

	AIC	BIC	logLik
	186.7878	194.2703	-88.39389

```
Random effects:
```

```
Formula: ~1 | subject
```

```
(Intercept) Residual
```

```
StdDev: 6.219726 1.168766
```

```
Fixed effects: SSS ~ test + diet
```

	Value	Std.Error	DF	t-value	p-value
(Intercept)	21.166667	2.1005158	17	10.076890	0.0000
testpre	-0.555556	0.3895888	17	-1.426005	0.1720
dietpasta	2.111111	2.9577768	16	0.713749	0.4857

#0.713749²=0.5094376, same as F value in AOV

```
Correlation:
```

```
(Intr) testpr
```

```
testpre -0.093
```

```
dietpasta -0.704 0.000
```

```
Standardized Within-Group Residuals:
```

	Min	Q1	Med	Q3	Max
	-1.586221065	-0.416570591	0.005211599	0.426747228	1.518070736

```
Number of Observations: 36
```

```
Number of Groups: 18
```

```
modell3 <- lmer(SSS~test+diet+(1|subject))
```

```
summary(modell3)
```

```
Linear mixed model fit by REML ['lmerMod']
```

```
Formula: SSS ~ test + diet + (1 | subject)
```

```
REML criterion at convergence: 176.7878
```

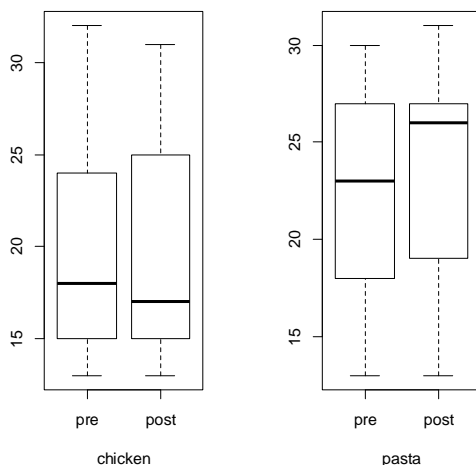
```
Random effects:
Groups      Name      Variance Std.Dev.
subject    (Intercept) 38.685    6.220
Residual                    1.366    1.169
Number of obs: 36, groups: subject, 18
```

```
Fixed effects:
              Estimate Std. Error t value
(Intercept)  21.1667    2.1005   10.077
testpre      -0.5556    0.3896   -1.426
dietpasta     2.1111    2.9578    0.714
```

```
Correlation of Fixed Effects:
              (Intr) testpr
testpre     -0.093
dietpasta   -0.704  0.000
```

To show there is no significant difference

```
> test=factor(test,levels=c("pre","post"))
> table(test)
test
pre post
18   18
> tapply(SSS,list(diet,test),mean)
              pre      post
chicken 20.88889 20.88889
pasta   22.44444 23.55556
> par(mfrow=c(1,2))
> boxplot(SSS[diet=="chicken"]~test[diet=="chicken"],xlab="chicken")
> boxplot(SSS[diet=="pasta"]~test[diet=="pasta"],xlab="pasta")
```



Ex 5, Longitudinal Data

Y vs. several fixed + random (nested in fixed1)

i.e., fixed-effect ("year") is nested in random-effect ("person")

- **person** = random effect factor
- **year** = fixed effect factor (68~90)

<< R printout >>

```
> data(psid)
> dim(psid)
[1] 1661 6
> attach(psid)
> head(psid)
  age educ sex income year person
1  31   12  M   6000   68      1
```



```

2 31 12 M 5300 69 1
3 31 12 M 5200 70 1
4 31 12 M 6900 71 1
5 31 12 M 7500 72 1
6 31 12 M 8000 73 1
> library(lattice)
> xyplot(income~year | person, type="l",subset=(person < 21))
> xyplot(income~year | person, type="l",subset=(person < 21),strip=F)
> slopes <- numeric(85); intercepts <- numeric(85)
for (i in 1:85) {
  lmod <- lm(log(income)~I(year-78),subset=(person==i))
  intercepts[i] <- coef(lmod)[1]
  slopes[i] <- coef(lmod)[2]}
> plot(intercepts,slopes)
> xyplot(log(income+100)~year|sex, type="l", groups=person)
> year <- year-78

> model1 <- lmer(log(income)~year*sex +
age+educ+(year | person))
> summary(model1)
Linear mixed model fit by REML ['lmerMod']
Formula: log(income) ~ year * sex + age + educ + (year | person)

REML criterion at convergence: 3819.776

Random effects:
Groups Name Variance Std.Dev. Corr
person (Intercept) 0.2817 0.53071
year 0.0024 0.04899 0.19
Residual 0.4673 0.68357
Number of obs: 1661, groups: person, 85

Fixed effects:
Estimate Std. Error t value
(Intercept) 6.67420 0.54332 12.284
year 0.08531 0.00900 9.480
sexM 1.15031 0.12129 9.484
age 0.01093 0.01352 0.808
educ 0.10421 0.02144 4.861
year:sexM -0.02631 0.01224 -2.150

Correlation of Fixed Effects:
(Intr) year sexM age educ
year 0.020
sexM -0.104 -0.098
age -0.874 0.002 -0.026
educ -0.597 0.000 0.008 0.167
year:sexM -0.003 -0.735 0.156 -0.010 -0.011
> qqmath(~resid(model1)|sex)

> model2 <- lme(log(income)~year*sex +
age+educ ,random=~year | person)
> anova(model2)
numDF denDF F-value p-value
(Intercept) 1 1574 21626.097 <.0001
year 1 1574 131.760 <.0001
sex 1 81 97.915 <.0001
age 1 81 0.000 0.9843
educ 1 81 23.409 <.0001
year:sex 1 1574 4.621 0.0317
> summary(model2)
Linear mixed-effects model fit by REML
Data: NULL
AIC BIC logLik
3839.776 3893.892 -1909.888

Random effects:
Formula: ~year | person
Structure: General positive-definite, Log-Cholesky parametrization
StdDev Corr
(Intercept) 0.53071321 (Intr)
year 0.04898952 0.187

```

```
Residual      0.68357323

Fixed effects: log(income) ~ year * sex + age + educ
              Value Std.Error   DF   t-value p-value
(Intercept)  6.674204  0.5433252 1574  12.283995  0.0000
year         0.085312  0.0089996 1574   9.479521  0.0000
sexM         1.150313  0.1212925   81   9.483790  0.0000
age          0.010932  0.0135238   81   0.808342  0.4213
educ         0.104210  0.0214366   81   4.861287  0.0000
year:sexM    -0.026307  0.0122378 1574  -2.149607  0.0317
```

Two printouts are identical

Ex 6, Split plots (with different spatial scales) **Y vs. several fixed + several random (nested in several layers)**

i.e., fixed-effect (“density”) is nested in fixed-effect (“irrigation”), which is nested in random-effect (“block”)

- **block** = random effect factor (A, B, C, D)
- **irrigation** = fixed effect factor (control vs. irrigated)
- **density** = fixed effect factor (low, medium, high)
- **fertilizer** = fixed effect factor (N, P, NP)

<< R printout >>

```
> data1 <- read.table("D:\\STAT999\\RBOOK\\splityield.txt",header=T)
> dim(data1)
[1] 72 5
> data1[1:7,]
  yield block irrigation density fertilizer
1    90     A   control     low          N
2    95     A   control     low          P
3   107     A   control     low         NP
4    92     A   control  medium          N
5    89     A   control  medium          P
6    92     A   control  medium         NP
7    81     A   control    high          N
> attach(data1)
> modell <-
```

```
lme(yield~irrigation*density*fertilizer,ran  
dom=~1|block/irrigation/density)
```

```
#ERROR if you enter yield~bolck*irrigation*density*fertilizer"
> summary(modell)
```

```
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC    logLik
481.6212 525.3789 -218.8106
```

Random effects:

```
Formula: ~1 | block
(Intercept)
StdDev: 0.000661967
```

```
Formula: ~1 | irrigation %in% block
(Intercept)
StdDev: 1.982465
```

```
Formula: ~1 | density %in% irrigation %in% block
(Intercept) Residual
StdDev: 6.975552 9.292805
```

```
Fixed effects: yield ~ irrigation * density * fertilizer
              Value Std.Error DF   t-value p-value
(Intercept)    80.50  5.893741 36 13.658558 0.0000
irrigationirrigated    31.75  8.335008  3  3.809235 0.0318
densitylow           5.50  8.216281 12  0.669403 0.5159
densitymedium        14.75  8.216281 12  1.795216 0.0978
fertilizerNP         5.50  6.571006 36  0.837010 0.4081
fertilizerP          4.50  6.571006 36  0.684827 0.4978
irrigationirrigated:densitylow   -39.00 11.619576 12 -3.356405 0.0057
irrigationirrigated:densitymedium -22.25 11.619576 12 -1.914872 0.0796
irrigationirrigated:fertilizerNP  13.00  9.292805 36  1.398932 0.1704
irrigationirrigated:fertilizerP   5.50  9.292805 36  0.591856 0.5576
densitylow:fertilizerNP   3.25  9.292805 36  0.349733 0.7286
densitymedium:fertilizerNP -6.75  9.292805 36 -0.726368 0.4723
densitylow:fertilizerP   -5.25  9.292805 36 -0.564953 0.5756
densitymedium:fertilizerP -5.50  9.292805 36 -0.591856 0.5576
irrigationirrigated:densitylow:fertilizerNP  7.75 13.142011 36  0.589712 0.5591
irrigationirrigated:densitymedium:fertilizerNP  3.75 13.142011 36  0.285344 0.7770
irrigationirrigated:densitylow:fertilizerP  20.00 13.142011 36  1.521837 0.1368
irrigationirrigated:densitymedium:fertilizerP  4.00 13.142011 36  0.304367 0.7626
```

Proceed to model simplification

```
> model2 <- lme(yield~(irrigation+density+fertilizer)^2,random=~1|block/irrigation/density)
> model3 <- lme(yield~irrigation*density+irrigation*fertilizer,random=~1|block/irrigation/density)
> summary(model3)
```

Linear mixed-effects model fit by REML

```
Data: NULL
      AIC      BIC    logLik
519.9035 549.6834 -245.9517
```

Random effects:

```
Formula: ~1 | block
(Intercept)
StdDev: 0.0005551647
```

```
Formula: ~1 | irrigation %in% block
(Intercept)
StdDev: 1.982612
```

```
Formula: ~1 | density %in% irrigation %in% block
(Intercept) Residual
StdDev: 7.057132 9.105995
```

```
Fixed effects: yield ~ irrigation * density + irrigation * fertilizer
              Value Std.Error DF   t-value p-value
(Intercept)    82.08333  4.994999 44 16.433103 0.0000
irrigationirrigated    27.80556  7.063995  3  3.936236 0.0292
densitylow           4.83333  6.222653 12  0.776732 0.4524
densitymedium        10.66667  6.222653 12  1.714167 0.1122
fertilizerNP         4.33333  3.717507 44  1.165656 0.2500
fertilizerP          0.91667  3.717507 44  0.246581 0.8064
irrigationirrigated:densitylow   -29.75000  8.800161 12 -3.380620 0.0055
irrigationirrigated:densitymedium -19.66667  8.800161 12 -2.234808 0.0452
irrigationirrigated:fertilizerNP  16.83333  5.257349 44  3.201867 0.0025
irrigationirrigated:fertilizerP  13.50000  5.257349 44  2.567834 0.0137
```

```
> model1r <-
```

```
lmer(yield~irrigation*density*fertilizer+(1|block)+(1|block:irrigation)+(1|block:irrigation:density))
```

```
> summary(model1r)
```

Linear mixed model fit by REML ['lmerMod']

```
Formula: yield ~ irrigation * density * fertilizer + (1 | block) + (1 | block:irrigation) + (1 | block:irrigation:density)
```

REML criterion at convergence: 437.6212

Random effects:

Groups	Name	Variance	Std.Dev.
block:irrigation:density	(Intercept)	48.657	6.975
block:irrigation	(Intercept)	3.931	1.983
block	(Intercept)	0.000	0.000
Residual		86.356	9.293

Number of obs: 72, groups: block:irrigation:density, 24; block:irrigation, 8; block, 4

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	80.500	5.894	13.659
irrigationirrigated	31.750	8.335	3.809
densitylow	5.500	8.216	0.669
densitymedium	14.750	8.216	1.795
fertilizerNP	5.500	6.571	0.837
fertilizerP	4.500	6.571	0.685
irrigationirrigated:densitylow	-39.000	11.620	-3.356
irrigationirrigated:densitymedium	-22.250	11.620	-1.915
irrigationirrigated:fertilizerNP	13.000	9.293	1.399
irrigationirrigated:fertilizerP	5.500	9.293	0.592
densitylow:fertilizerNP	3.250	9.293	0.350
densitymedium:fertilizerNP	-6.750	9.293	-0.726
densitylow:fertilizerP	-5.250	9.293	-0.565
densitymedium:fertilizerP	-5.500	9.293	-0.592
irrigationirrigated:densitylow:fertilizerNP	7.750	13.142	0.590
irrigationirrigated:densitymedium:fertilizerNP	3.750	13.142	0.285
irrigationirrigated:densitylow:fertilizerP	20.000	13.142	1.522
irrigationirrigated:densitymedium:fertilizerP	4.000	13.142	0.304

Proceed to model simplification

```
> model2r <-
lmer(yield~(irrigation+density+fertilizer)^2+(1|block)+(1|block:irrigation)+(1|block:irrigation:de
nsity))
> model3r <-
lmer(yield~irrigation*density+irrigation*fertilizer+(1|block)+(1|block:irrigation)+(1|block:irriga
tion:density))
```

These printouts are identical to “lme”

Ex 7, Hierarchical sampling

Y vs. fixed + several random (nested in several layers)

i.e., random-effect (“house”) is nested in random-effect (“street”), which is nested in random-effect (“district”), which is nested in random-effect (“town”)

- **gender** = fixed effect factor (male vs. female)
- **town** = random effect factor (Coventry, Derby, Leeds, Norwich)
- **district** = random effect factor (A,B,C,D,E,F)
- **street** = random effect factor (1,2,3,4,5,6,7,8,9,10)
- **house** = random effect factor (door1, door2, door3, door4)

<< R printout >>

```
> data1 <- read.table("D:\\STAT999\\RBOOK\\childfull.txt",header=T)
> dim(data1)
[1] 2972 8
> head(data1)
  childID child house street district town response gender
1      1      1 door1      1      A Leeds 83.88773  male
2      1      1 door2      1      A Leeds 99.96294  male
3      1      3 door3      1      A Leeds 87.20253 female
4      2      3 door3      1      A Leeds 89.37665  male
5      3      3 door3      1      A Leeds 92.01751 female
6      1      5 door4      1      A Leeds 87.12672 female
> attach(data1)
> model1 <-
```

```
lmer(response~gender+(1|town)+(1|town:district)+(
1|town:district:street)+(1|town:district:street:h
ouse))
```

You can also enter "factor(street)" "in place of "street". You get the same printout.

```
> summary(model1)
```

Linear mixed model fit by REML ['lmerMod']

Formula: response ~ gender + (1 | town) + (1 | town:district) + (1 | town:district:street) + (1 | town:district:street:house)

REML criterion at convergence: 19864.29

Random effects:

Groups	Name	Variance	Std.Dev.
town:district:street:house	(Intercept)	4.082	2.020
town:district:street	(Intercept)	15.675	3.959
town:district	(Intercept)	168.350	12.975
town	(Intercept)	36.974	6.081
Residual		36.241	6.020

Number of obs: 2972, groups: town:district:street:house, 960; town:district:street, 240; town:district, 24; town, 4

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	97.8965	4.0442	24.206
gendermale	0.5368	0.2363	2.272

Correlation of Fixed Effects:

(Intr)
gendermale -0.030

```
> model2 <- lme(response~gender,  
random=~1 | town/district/street/house)
```

Two printouts are identical.

```
> summary(model2)
```

Linear mixed-effects model fit by REML

Data: NULL

AIC	BIC	logLik
19878.29	19920.26	-9932.144

Random effects:

Formula: ~1 | town
(Intercept)

StdDev: 6.080077

Formula: ~1 | district %in% town
(Intercept)

StdDev: 12.9751

Formula: ~1 | street %in% district %in% town
(Intercept)

StdDev: 3.959135

Formula: ~1 | house %in% street %in% district %in% town
(Intercept) Residual

StdDev: 2.020298 6.020018

Fixed effects: response ~ gender

	Value	Std.Error	DF	t-value	p-value
(Intercept)	97.89646	4.044044	2011	24.207565	0.0000
gendermale	0.53679	0.236271	2011	2.271908	0.0232

Correlation:

(Intr)
gendermale -0.03

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-4.807246312	-0.631781548	-0.009618952	0.651702748	3.145238160

Number of Observations: 2972

Number of Groups:

	town	district %in% town
	4	24
street %in% district %in% town	house %in% street %in% district %in% town	

To test for “gender” effect:

```
> model11 <-
lmer(response~gender+(1|town)+(1|town:district)+(1|town:district:street)+(1|town:district:street:h
ouse),REML=F)
> model11 <-
lmer(response~(1|town)+(1|town:district)+(1|town:district:street)+(1|town:district:street:house),R
EML=F)
> anova(model11, model11)
Data:
Models:
model11: response ~ (1 | town) + (1 | town:district) + (1 | town:district:street) +
model11:      (1 | town:district:street:house)
model1: response ~ gender + (1 | town) + (1 | town:district) + (1 | town:district:street) +
model1:      (1 | town:district:street:house)
      Df    AIC    BIC logLik deviance   Chisq Chi Df Pr(>Chisq)
model11  6 19885 19921 -9936.4   19873
model1   7 19882 19924 -9933.9   19868  5.1552     1 0.02318 *
```

OR

```
> model2 <- lme(response~gender, random=~1|town/district/street/house, method="ML")
> model22 <- lme(response~1, random=~1|town/district/street/house, method="ML")
> anova(model2, model22)
      Model df      AIC      BIC    logLik    Test  L.Ratio p-value
model2      1  7 19881.72 19923.70 -9933.861
model22     2  6 19884.88 19920.86 -9936.439 1 vs 2  5.155196 0.0232
```

Ex 8, More Nested Effects

Y vs. several random (nested in several layers)

i.e., random-effect (“sample”) is nested in random-effect (“technician”), which is nested in random-effect (“lab”)

- **Lab** = random effect factor (I,II,III,IV,V,VI)
- **Technician** = random effect factor (one & two)
- **Sample** = random effect factor (G vs. H)

<< R printout >>

```
> library(faraway)
> data(eggs)
> dim(eggs)
[1] 48 4
> attach(eggs)
> table(Lab, Technician)
      Technician
Lab    one two
I       4  4
II      4  4
III     4  4
IV      4  4
V       4  4
VI      4  4
> table(Sample, Technician)
      Technician
Sample one two
G      12 12
H      12 12
> eggs
      Fat Lab Technician Sample
1  0.62  I       one      G
2  0.55  I       one      G
3  0.34  I       one      H
4  0.24  I       one      H
5  0.80  I       two      G
6  0.68  I       two      G
```

```

7 0.76 I two H
8 0.65 I two H
> library(lme4)
> library(nlme)
> modell <-
lmer(Fat~1+(1|Lab)+(1|Lab:Technician)+(1|La
b:Technician:Sample))
> summary(modell)
Linear mixed model fit by REML ['lmerMod']
Formula: Fat ~ 1 + (1 | Lab) + (1 | Lab:Technician) + (1 | Lab:Technician:Sample)

REML criterion at convergence: -64.2351

Random effects:
Groups Name Variance Std.Dev.
Lab:Technician:Sample (Intercept) 0.003065 0.05536
Lab:Technician (Intercept) 0.006980 0.08355
Lab (Intercept) 0.005920 0.07694
Residual 0.007196 0.08483
Number of obs: 48, groups: Lab:Technician:Sample, 24; Lab:Technician, 12; Lab, 6

Fixed effects:
Estimate Std. Error t value
(Intercept) 0.38750 0.04296 9.019
> modell <- lmer(Fat~1+(1|Lab)+(1|Lab:Technician)+(1|Lab:Technician:Sample),REML=F)
> summary(modell)
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: Fat ~ 1 + (1 | Lab) + (1 | Lab:Technician) + (1 | Lab:Technician:Sample)

AIC BIC logLik deviance
-58.7859 -49.4299 34.3929 -68.7859

Random effects:
Groups Name Variance Std.Dev.
Lab:Technician:Sample (Intercept) 0.003065 0.05536
Lab:Technician (Intercept) 0.006980 0.08355
Lab (Intercept) 0.004074 0.06383
Residual 0.007196 0.08483
Number of obs: 48, groups: Lab:Technician:Sample, 24; Lab:Technician, 12; Lab, 6

Fixed effects:
Estimate Std. Error t value
(Intercept) 0.38750 0.03922 9.88

# Proceed to model simplification

> model2 <- lmer(Fat~1+(1|Lab)+(1|Lab:Technician),REML=F)
> summary(model2)
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: Fat ~ 1 + (1 | Lab) + (1 | Lab:Technician)

AIC BIC logLik deviance
-59.1825 -51.6977 33.5912 -67.1825

Random effects:
Groups Name Variance Std.Dev.
Lab:Technician (Intercept) 0.008002 0.08945
Lab (Intercept) 0.004074 0.06383
Residual 0.009239 0.09612
Number of obs: 48, groups: Lab:Technician, 12; Lab, 6

Fixed effects:
Estimate Std. Error t value
(Intercept) 0.38750 0.03922 9.88
> VarCorr(model2)
Groups Name Std.Dev.
Lab:Technician (Intercept) 0.089453
Lab (Intercept) 0.063827
Residual 0.096119
> anova(modell, model2)
Data:

```

```
Models:
model2: Fat ~ 1 + (1 | Lab) + (1 | Lab:Technician)
model1: Fat ~ 1 + (1 | Lab) + (1 | Lab:Technician) + (1 | Lab:Technician:Sample)
      Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
model2  4 -59.182 -51.698 33.591 -67.182
model1  5 -58.786 -49.430 34.393 -68.786 1.6034      1      0.2054
```

Variation due to “sample” is NOT significant.

```
> model2 <- lmer(Fat~1+(1|Lab)+(1|Lab:Technician),REML=F)
> model3 <- lmer(Fat~1+(1|Lab),REML=F)
> anova(model2, model3)
```

Data:

Models:

```
model3: Fat ~ 1 + (1 | Lab)
model2: Fat ~ 1 + (1 | Lab) + (1 | Lab:Technician)
      Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
model3  3 -53.272 -47.659 29.636 -59.272
model2  4 -59.182 -51.698 33.591 -67.182 7.91      1      0.004916 **
```

Variation due to “technician” is significant.

Permutation test of logLikelihood

```
> lrstat <- numeric(500)
> for (i in 1:500) {
  sFat <- unlist(simulate(model2))
  model_2s <- lmer(sFat~1+(1|Lab)+(1|Lab:Technician))
  model_1s <- lmer(sFat~1+(1|Lab)+(1|Lab:Technician)+(1|Lab:Technician:Sample))
  lrstat[i] <- 2*(logLik(model_1s)-logLik(model_2s))
}
> hist(lrstat)
> (2*(logLik(model1)-logLik(model2)))
'log Lik.' 6.154244 (df=5)
> model1 <- lmer(sFat~1+(1|Lab)+(1|Lab:Technician)+(1|Lab:Technician:Sample))
> model2 <- lmer(sFat~1+(1|Lab)+(1|Lab:Technician))
> (2*(logLik(model1)-logLik(model2)))
'log Lik.' -4.312994e-11 (df=5)
> mean(lrstat<0.0001)
[1] 0.508 # Variation due to “sample” is NOT significant.
```

```
> model1 <- lme(Fat~1, random=~1|Lab/Technician/Sample)
```

```
> summary(model1)
```

Linear mixed-effects model fit by REML

Data: NULL

```
      AIC      BIC    logLik
-54.23508 -44.98434 32.11754
```

Random effects:

```
Formula: ~1 | Lab
(Intercept)
```

```
StdDev: 0.07694
```

```
Formula: ~1 | Technician %in% Lab
(Intercept)
```

```
StdDev: 0.08355
```

```
Formula: ~1 | Sample %in% Technician %in% Lab
(Intercept) Residual
```

```
StdDev: 0.05536 0.08483
```

Fixed effects: Fat ~ 1

```
      Value Std.Error DF t-value p-value
(Intercept) 0.3875 0.04296437 24 9.0191      0
```

Number of Observations: 48

Number of Groups:

```
      Lab      Technician %in% Lab
      6      12
Sample %in% Technician %in% Lab
      24
```

```
> model1 <- lme(Fat~1, random=~1|Lab/Technician/Sample, method="ML")
> model2 <- lme(Fat~1, random=~1|Lab/Technician, method="ML")
```



```
> anova(model1, model2)
      Model df      AIC      BIC    logLik    Test  L.Ratio p-value
model1     1   5 -58.78590 -49.42989 34.39295
model2     2   4 -59.18247 -51.69767 33.59124 1 vs 2 1.603423 0.2054
```

These printouts are identical to “lmer”

Ex 9, Crossed Effects (i.e., Factorial Experiment, i.e., NOT nested) This example is “Latin Square”
Y vs. fixed + two random
 i.e., two random-effect terms (“run”) and (“position”) are crossed.

- **material** = fixed effect factor (A,B,C,D)
- **run** = random effect factor (1,2,3,4)
- **position** = random effect factor (1,2,3,4)

<< R printout >>

```
> data(abrasion)
> dim(abrasion)
[1] 16 4
> abrasion
  run position material wear
1   1         1       C   235
2   1         2       D   236
3   1         3       B   218
4   1         4       A   268
5   2         1       A   251
6   2         2       B   241
7   2         3       D   227
8   2         4       C   229
9   3         1       D   234
10  3         2       C   273
11  3         3       A   274
12  3         4       B   226
13  4         1       B   195
14  4         2       A   270
15  4         3       C   230
16  4         4       D   225
> matrix(material,4,4)
      [,1] [,2] [,3] [,4]
[1,] "C"  "A"  "D"  "B"
[2,] "D"  "B"  "C"  "A"
[3,] "B"  "D"  "A"  "C"
[4,] "A"  "C"  "B"  "D"
> attach(abrasion)
```

Regular ANOVA for fun

```
> model1 <- aov(wear~material+run+position)
> summary(model1)
              Df Sum Sq Mean Sq F value    Pr(>F)
material      3   4622   1540.5    25.151 0.00085 ***
run            3    986    328.8     5.369 0.03901 *
position      3   1468    489.5     7.992 0.01617 *
Residuals    6    368     61.3
> model1r <-
```

```
lmer(wear~material+(1|run)+(1|position))
```

```
> summary(model1r)
Linear mixed model fit by REML ['lmerMod']
Formula: wear ~ material + (1 | run) + (1 | position)

REML criterion at convergence: 100.2563

Random effects:
Groups      Name      Variance Std.Dev.
run         (Intercept) 66.90    8.179
```

```

position (Intercept) 107.06 10.347
Residual             61.25  7.826
Number of obs: 16, groups: run, 4; position, 4

```

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	265.750	7.668	34.66
materialB	-45.750	5.534	-8.27
materialC	-24.000	5.534	-4.34
materialD	-35.250	5.534	-6.37

Correlation of Fixed Effects:

	(Intr)	matr1B	matr1C
materialB	-0.361		
materialC	-0.361	0.500	
materialD	-0.361	0.500	0.500

Do NOT use “lme” for this example!

```

> modelle <- lme(wear~material,random=~run|position)
Error in lme.formula(wear ~ material, random = ~run | position) :
  nlminb problem, convergence error code = 1
  message = iteration limit reached without convergence (10)
> modelle <- lme(wear~material,random=~position|run)
Error in lme.formula(wear ~ material, random = ~position | run) :
  nlminb problem, convergence error code = 1
  message = iteration limit reached without convergence (10)
> modelle <- lme(wear~material,random=~1|position:run)
Error in getGroups.data.frame(dataMix, groups) :
  invalid formula for groups
> modelle <-

```

```
lme(wear~material,random=~1|position/run)
```

WRONG!

```

> summary(modelle)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC    logLik
117.3022 120.6965 -51.6511

Random effects:
Formula: ~1 | position
(Intercept)
StdDev:    9.206731

Formula: ~1 | run %in% position
(Intercept) Residual
StdDev:    11.34609 4.659484

Fixed effects: wear ~ material
      Value Std.Error DF   t-value p-value
(Intercept) 265.75   7.668252   9 34.65588  0.0000
materialB   -45.75   8.673075   9 -5.27495  0.0005
materialC   -24.00   8.673075   9 -2.76718  0.0219
materialD   -35.25   8.673075   9 -4.06430  0.0028
Correlation:
      (Intr) matr1B matr1C
materialB -0.566
materialC -0.566  0.500
materialD -0.566  0.500  0.500

Number of Observations: 16
Number of Groups:
      position run %in% position
           4           16

```

Ex 10, Mixed Effects

Y vs. fixed + two random (one is nested in the other, and they are nested in fixed)
i.e., fixed (“treatment”) + random (“liver”) is nested in another random (“rat”), which is nested in fixed (“treatment”).

- **Treatment** = fixed effect factor (1,2,3)
- **Rat** = random effect factor (2 rats were allocated per Treatment)
- **Liver** = random effect factor (Each rat’s liver was cut into 3 pieces and each piece was further divided into two)

<< R printout >>

```
> data1 <- read.table("D:\\STAT999\\RBOOK\\rats.txt",header=T)
```

```
> dim(data1)
```

```
[1] 36 4
```

```
> data1
```

	Glycogen	Treatment	Rat	Liver
1	131	1	1	1
2	130	1	1	1
3	131	1	1	2
4	125	1	1	2
5	136	1	1	3
6	142	1	1	3
7	150	1	2	1
8	148	1	2	1
9	140	1	2	2
10	143	1	2	2
11	160	1	2	3
12	150	1	2	3
13	157	2	1	1
14	145	2	1	1
15	154	2	1	2
16	142	2	1	2
17	147	2	1	3
18	153	2	1	3
19	151	2	2	1
20	155	2	2	1
21	147	2	2	2
22	147	2	2	2
23	162	2	2	3
24	152	2	2	3
25	134	3	1	1
26	125	3	1	1
27	138	3	1	2
28	138	3	1	2
29	135	3	1	3
30	136	3	1	3
31	138	3	2	1
32	140	3	2	1
33	139	3	2	2
34	138	3	2	2
35	134	3	2	3
36	127	3	2	3

```
> attach(data1)
```

```
> Treatment <- factor(Treatment)
```

```
> Rat <- factor(Rat)
```

```
> Liver <- factor(Liver)
```

```
> model11 <- aov(Glycogen~Treatment) # WRONG!
```

```
> summary(model11)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Treatment	2	1558	778.8	14.5	3.03e-05 ***
Residuals	33	1773	53.7		

Six rats received TREATMENT, so df should be (2, 3), NOT (2, 33)

```
> model1 <-
```

```
aov(Glycogen~Treatment+Error(Treatment/Rat/Liver))
```

```
> summary(model1)
```

Error: Treatment

	Df	Sum Sq	Mean Sq
Treatment	2	1558	778.8

Error: Treatment:Rat

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	3	797.7	265.9		

#(265.9-49.5)+6 = 36.06

Error: Treatment:Rat:Liver

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	12	594	49.5		

#(49.5-21.17)+2 = 14.165

Error: Within

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	18	381	21.17		

```
> library(lme4)
```

```
> modellr <-
```

```
lmer(Glycogen~Treatment+(1|Treatment:Rat)+(1|Treatment:Rat:Liver))
```

```
> summary(modellr)
```

Linear mixed model fit by REML ['lmerMod']

Formula: Glycogen ~ Treatment + (1 | Treatment:Rat) + (1 | Treatment:Rat:Liver)

REML criterion at convergence: 219.6213

Random effects:

Groups	Name	Variance	Std.Dev.
Treatment:Rat:Liver	(Intercept)	14.17	3.764
Treatment:Rat	(Intercept)	36.06	6.005
Residual		21.17	4.601

Number of obs: 36, groups: Treatment:Rat:Liver, 18; Treatment:Rat, 6

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	140.500	4.707	29.848
Treatment2	10.500	6.657	1.577
Treatment3	-5.333	6.657	-0.801

Correlation of Fixed Effects:

	(Intr)	Trtmn2
Treatment2	-0.707	
Treatment3	-0.707	0.500

```
> anova(modellr)
```

Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value
Treatment	2	123.99	61.997	2.929

```
> library(nlme)
```

```
> modelle <-
```

```
lme(Glycogen~Treatment,random=~1|Treatment/Rat/Liver)
```

```
> summary(modelle)
```

Linear mixed-effects model fit by REML

Data: NULL

	AIC	BIC	logLik
	233.6213	244.0968	-109.8106

Random effects:

Formula: ~1 | Treatment
(Intercept)

StdDev: 3.541638

Formula: ~1 | Rat %in% Treatment
(Intercept)

StdDev: 6.005399

Formula: ~1 | Liver %in% Rat %in% Treatment

```

      (Intercept) Residual
StdDev:    3.763863  4.600725

Fixed effects: Glycogen ~ Treatment
              Value Std.Error DF   t-value p-value
(Intercept) 140.50000  5.890722 18 23.851065     0
Treatment2   10.50000  8.330739  0  1.260392    NaN
Treatment3   -5.33333  8.330739  0 -0.640199    NaN

Correlation:
      (Intr) Trtmn2
Treatment2 -0.707
Treatment3 -0.707  0.500

Number of Observations: 36
Number of Groups:
      Treatment      Rat %in% Treatment
           3              6
Liver %in% Rat %in% Treatment
           18

```

Available plots

```

> plot(modelle)
> plot(modelle,Glycogen~fitted(.))
> qqnorm(modelle,~resid(.))
> qqnorm(modelle,~resid(.)|Treatment)

> plot(model1r)
> plot(model1r,Glycogen~fitted(.))
> qqmath(~resid(model1r))
> qqmath(~resid(model1r)|Treatment)
> xyplot(resid(model1r)~fitted(model1r)|Treatment, layout=c(3,1))

```

Ex 11, Mixed effect model

Y vs. fixed + two random (one is nested in in fixed)
i.e., fixed (“melanic”) + one random (“male”) +
another random (“female”), which is nested in
fixed (“melanic”).

The aim of this experiment is to measure the “vigor of courtship” using pheasants. The study involves placing a male and a female in separate cages, screened from each other, in the evening. The screen is removed during the hours of darkness. The male’s behavior is then recorded as the fraction of that hour in which the male’s wattles are erect. The main constraint is that only 36 measurement can be taken, owing to the number of days of the breeding season and the likely frequency of rain.

PLAN 1: Six males and six females, with each possible combination of pairing.

- **MELANIC** = fixed effect factor (1=melanic phenotype, 2=normal phenotype)
- **MALE** = random effect factor
- **FEMALE** = random effect factor nested within MELANIC

<< Minitab printout >>

```

MTB > print c1 c2 c3

Row  MALE  FEMALE  MELANIC  WATTLE
  1     1      1      1      0.598
  2     2      1      1      0.579
  3     3      1      1      0.513
  4     4      1      1      0.574
  5     5      1      1      0.455
  6     6      1      1      0.436
  7     1      2      1      0.620
  8     2      2      1      0.682
  9     3      2      1      0.596
 10     4      2      1      0.739
 11     5      2      1      0.573
 12     6      2      1      0.498
 13     1      3      1      0.517

```

14	2	3	1	0.537
15	3	3	1	0.427
16	4	3	1	0.464
17	5	3	1	0.507
18	6	3	1	0.518
19	1	1	2	0.596
20	2	1	2	0.495
21	3	1	2	0.421
22	4	1	2	0.579
23	5	1	2	0.462
24	6	1	2	0.465
25	1	2	2	0.399
26	2	2	2	0.586
27	3	2	2	0.427
28	4	2	2	0.671
29	5	2	2	0.701
30	6	2	2	0.583
31	1	3	2	0.547
32	2	3	2	0.479
33	3	3	2	0.437
34	4	3	2	0.529
35	5	3	2	0.517
36	6	3	2	0.571

```
MTB > glm WATTLE = MELANIC + MALE + FEMALE(MELANIC);
SUBC> random MALE FEMALE(MELANIC);
SUBC> ems.
```

General Linear Model: WATTLE versus MELANIC, MALE, FEMALE

Factor	Type	Levels	Values
MELANIC	fixed	2	1, 2
MALE	random	6	1, 2, 3, 4, 5, 6
FEMALE(MELANIC)	random	6	1, 2, 3, 1, 2, 3

Analysis of Variance for WATTLE, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
MELANIC	1	0.003762	0.003762	0.003762	0.25	0.645
MALE	5	0.052756	0.052756	0.010551	2.06	0.105
FEMALE(MELANIC)	4	0.060705	0.060705	0.015176	2.96	0.039
Error	25	0.128077	0.128077	0.005123		
Total	35	0.245300				

S = 0.0715757 R-Sq = 47.79% R-Sq(adj) = 26.90%

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 MELANIC	(4) + 6.0000 (3) + Q[1]
2 MALE	(4) + 6.0000 (2)
3 FEMALE(MELANIC)	(4) + 6.0000 (3)
4 Error	(4)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 MELANIC	4.00	0.015176	(3)
2 MALE	25.00	0.005123	(4)
3 FEMALE(MELANIC)	25.00	0.005123	(4)

Variance Components, using Adjusted SS

Source	Estimated Value
MALE	0.00090
FEMALE(MELANIC)	0.00168
Error	0.00512

<< R printout >>

```

data1 <- read.csv("D:\\STAT999\\Minitab_p248.csv")
head(data1)
  MALE FEMALE MELANIC WATTLE
1     1      1      1  0.598
2     2      1      1  0.579
3     3      1      1  0.513
4     4      1      1  0.574
5     5      1      1  0.455
6     6      1      1  0.436
attach(data1)
MELATIC <- factor(MELANIC)
MALE <- factor(MALE)
FEMALE <- factor(FEMALE)
modell <-
lmer(WATTLE~MELANIC+(1|MALE)+(1|MELANIC:FEMALE))
summary(modell)
Linear mixed model fit by REML ['lmerMod']
Formula: WATTLE ~ MELANIC + (1 | MALE) + (1 | MELANIC:FEMALE)

REML criterion at convergence: -69.0911

Random effects:
Groups          Name          Variance Std.Dev.
MALE             (Intercept)  0.0009047 0.03008
MELANIC:FEMALE   (Intercept)  0.0016755 0.04093
Residual                    0.0051231 0.07158
Number of obs: 36, groups: MALE, 6; MELANIC:FEMALE, 6

Fixed effects:
              Estimate Std. Error t value
(Intercept)  0.56672    0.06608   8.577
MELANIC      -0.02044    0.04106  -0.498

Correlation of Fixed Effects:
          (Intr)
MELANIC  -0.932

modell <- lmer(WATTLE~MELANIC + (1|MALE) + (1|MELANIC:FEMALE),REML=F)
model2 <- lmer(WATTLE~ (1|MALE) + (1|MELANIC:FEMALE),REML=F)
anova(modell, model2)
Data:
Models:
model2: WATTLE ~ (1 | MALE) + (1 | MELANIC:FEMALE)
model1: WATTLE ~ MELANIC + (1 | MALE) + (1 | MELANIC:FEMALE)
      Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
model2  4 -71.283 -64.949 39.642  -79.283
model1  5 -69.627 -61.710 39.814  -79.627 0.3441      1      0.5575

model3 <- lmer(WATTLE~ MELANIC + (1|MELANIC:FEMALE),REML=F)
anova(modell, model3)
Data:
Models:
model3: WATTLE ~ MELANIC + (1 | MELANIC:FEMALE)
model1: WATTLE ~ MELANIC + (1 | MALE) + (1 | MELANIC:FEMALE)
      Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
model3  4 -70.739 -64.405 39.369  -78.739
model1  5 -69.627 -61.710 39.814  -79.627 0.8886      1      0.3459

```

PLAN 1 -1: If you include MELANIC*MALE interaction

<< Minitab printout >>

```

MTB > glm WATTLE = MELANIC|MALE + FEMALE(MELANIC);
SUBC> random MALE FEMALE;
SUBC> ems.

```

General Linear Model: WATTLE versus MELANIC, MALE, FEMALE

Factor	Type	Levels	Values
MELANIC	fixed	2	1, 2
MALE	random	6	1, 2, 3, 4, 5, 6
FEMALE(MELANIC)	random	6	1, 2, 3, 1, 2, 3

Analysis of Variance for WATTLE, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P	
MELANIC	1	0.003762	0.003762	0.003762	0.23	0.657	x NOT exact test!
MALE	5	0.052756	0.052756	0.010551	1.73	0.282	
MELANIC*MALE	5	0.030540	0.030540	0.006108	1.25	0.322	
FEMALE(MELANIC)	4	0.060705	0.060705	0.015176	3.11	0.038	
Error	20	0.097537	0.097537	0.004877			
Total	35	0.245300					

x Not an exact F-test.

S = 0.0698344 R-Sq = 60.24% R-Sq(adj) = 30.42%

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 MELANIC	(5) + 6.0000 (4) + 3.0000 (3) + Q[1]
2 MALE	(5) + 3.0000 (3) + 6.0000 (2)
3 MELANIC*MALE	(5) + 3.0000 (3)
4 FEMALE(MELANIC)	(5) + 6.0000 (4)
5 Error	(5)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 MELANIC	4.06	0.016407	(3) + (4) - (5)
2 MALE	5.00	0.006108	(3)
3 MELANIC*MALE	20.00	0.004877	(5)
4 FEMALE(MELANIC)	20.00	0.004877	(5)

Variance Components, using Adjusted SS

Source	Estimated Value
MALE	0.00074
MELANIC*MALE	0.00041
FEMALE(MELANIC)	0.00172
Error	0.00488

<< R printout >>

```

> model11 <- lmer(WATTLE~MELANIC + (1|MALE) + (1|MELANIC:FEMALE),REML=F)

> model15 <-
lmer(WATTLE~MELANIC*MALE+(1|MALE)+(1|MELANIC:FEMALE),REML=F)
> summary(model15)
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: WATTLE ~ MELANIC * MALE + (1 | MALE) + (1 | MELANIC:FEMALE)

      AIC      BIC    logLik deviance
-67.2590 -43.5062  48.6295 -97.2590

Random effects:
Groups      Name          Variance Std.Dev.
MALE        (Intercept)  0.000000  0.00000
MELANIC:FEMALE (Intercept) 0.001144  0.03383
Residual                    0.003251  0.05702
Number of obs: 36, groups: MALE, 6; MELANIC:FEMALE, 6

Fixed effects:
              Estimate Std. Error t value
(Intercept)   0.64267    0.08559   7.508
MELANIC       -0.06433    0.05413  -1.188
MALE2         0.03600    0.10410   0.346
MALE3        -0.04700    0.10410  -0.451
MALE4        -0.05100    0.10410  -0.490
MALE5        -0.17933    0.10410  -1.723
MALE6        -0.21433    0.10410  -2.059
MELANIC:MALE2 -0.01500    0.06584  -0.228
MELANIC:MALE3 -0.01933    0.06584  -0.294
MELANIC:MALE4  0.06500    0.06584   0.987
MELANIC:MALE5  0.11267    0.06584   1.711
MELANIC:MALE6  0.12000    0.06584   1.823

> AIC(model11, model15)
      df      AIC

```



```

modell  5 -69.62732
model5 15 -67.25901

> anova(modell1, model5)
Data:
Models:
modell1: WATTLE ~ MELANIC + (1 | MALE) + (1 | MELANIC:FEMALE)
model5: WATTLE ~ MELANIC * MALE + (1 | MALE) + (1 | MELANIC:FEMALE)
      Df      AIC      BIC logLik deviance   Chisq Chi Df Pr(>Chisq)
modell1  5 -69.627 -61.710 39.814  -79.627
model5 15 -67.259 -43.506 48.630  -97.259 17.632    10    0.0615 .
---
```

For your thoughts

PLAN 2: Six males and thirty six females, with each male used six times and each female used only once

- **MELANIC** = fixed effect factor (1=melanic phenotype, 2=normal phenotype)
- **MALE** = random effect factor
- **FEMALE** = random effect factor nested within MELANIC

<< Minitab printout >>

```

MTB > glm WATTLE = MELANIC | MALE;
SUBC> random MALE;
SUBC> ems.
```

General Linear Model: WATTLE versus MELANIC, MALE

Factor	Type	Levels	Values
MELANIC	fixed	2	1, 2
MALE	random	6	1, 2, 3, 4, 5, 6

Analysis of Variance for WATTLE, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
MELANIC	1	0.003762	0.003762	0.003762	0.62	0.468
MALE	5	0.052756	0.052756	0.010551	1.73	0.282
MELANIC*MALE	5	0.030540	0.030540	0.006108	0.93	0.481
Error	24	0.158242	0.158242	0.006593		
Total	35	0.245300				

S = 0.0811999 R-Sq = 35.49% R-Sq(adj) = 5.92%

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 MELANIC	(4) + 3.0000 (3) + Q[1]
2 MALE	(4) + 3.0000 (3) + 6.0000 (2)
3 MELANIC*MALE	(4) + 3.0000 (3)
4 Error	(4)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 MELANIC	5.00	0.006108	(3)
2 MALE	5.00	0.006108	(3)
3 MELANIC*MALE	24.00	0.006593	(4)

Variance Components, using Adjusted SS

Source	Estimated Value
MALE	0.00074
MELANIC*MALE	-0.00016
Error	0.00659

<< R printout >>

```

model6 <- lmer(WATTLE~MELANIC + (1|MALE) +(1|MELANIC:MALE), REML=F)
model7 <- lmer(WATTLE~MELANIC + (1|MALE) +(1|MALE:MELANIC), REML=F)
> summary(model7)
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: WATTLE ~ MELANIC + (1 | MALE) + (1 | MALE:MELANIC)

      AIC      BIC    logLik deviance
```

```

-68.2902 -60.3726 39.1451 -78.2902

Random effects:
Groups      Name      Variance Std.Dev.
MALE:MELANIC (Intercept) 0.0000000 0.00000
MALE        (Intercept) 0.0004166 0.02041
Residual                    0.0062927 0.07933
Number of obs: 36, groups: MALE:MELANIC, 12; MALE, 6

Fixed effects:
              Estimate Std. Error t value
(Intercept)  0.56672    0.04263  13.294
MELANIC      -0.02044    0.02644  -0.773

Correlation of Fixed Effects:
      (Intr)
MELANIC -0.930

> summary(model16)
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: WATTLE ~ MELANIC + (1 | MALE) + (1 | MELANIC:MALE)

      AIC      BIC    logLik deviance
-68.2902 -60.3726 39.1451 -78.2902

Random effects:
Groups      Name      Variance Std.Dev.
MELANIC:MALE (Intercept) 0.0000000 0.00000
MALE        (Intercept) 0.0004166 0.02041
Residual                    0.0062927 0.07933
Number of obs: 36, groups: MELANIC:MALE, 12; MALE, 6

Fixed effects:
              Estimate Std. Error t value
(Intercept)  0.56672    0.04263  13.294
MELANIC      -0.02044    0.02644  -0.773

Correlation of Fixed Effects:
      (Intr)
MELANIC -0.930
> model18 <- lmer(WATTLE~ (1|MALE) +(1|MALE:MELANIC), REML=F)
> anova(model17, model18)
Data:
Models:
model18: WATTLE ~ (1 | MALE) + (1 | MALE:MELANIC)
model17: WATTLE ~ MELANIC + (1 | MALE) + (1 | MALE:MELANIC)
      Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
model18  4 -69.698 -63.364 38.849  -77.698
model17  5 -68.290 -60.373 39.145  -78.290 0.5919      1      0.4417

> model19 <- lmer(WATTLE~MELANIC + (1|MALE:MELANIC), REML=F)
> anova(model17, model19)
Data:
Models:
model19: WATTLE ~ MELANIC + (1 | MALE:MELANIC)
model17: WATTLE ~ MELANIC + (1 | MALE) + (1 | MALE:MELANIC)
      Df      AIC      BIC logLik deviance Chisq Chi Df Pr(>Chisq)
model19  4 -70.00 -63.666 39.000  -78.00
model17  5 -68.29 -60.373 39.145  -78.29 0.2903      1      0.5901

```

For more thoughts

PLAN 3: Six females and thirty six males, with each female used six times and each male used only once

- **MELANIC** = fixed effect factor (1=melanic phenotype, 2=normal phenotype)
- **MALE** = random effect factor
- **FEMALE** = random effect factor nested within MELANIC

<< Minitab printout >>

```

MTB > glm WATTLE = MELANIC + FEMALE(MELANIC);
SUBC> random FEMALE;
SUBC> ems.

```

General Linear Model: WATTLE versus MELANIC, FEMALE

Factor	Type	Levels	Values
--------	------	--------	--------

```

MELANIC          fixed          2 1, 2
FEMALE(MELANIC)  random         6 1, 2, 3, 1, 2, 3

Analysis of Variance for WATTLE, using Adjusted SS for Tests

Source            DF      Seq SS      Adj SS      Adj MS      F      P
MELANIC           1    0.003762    0.003762    0.003762    0.25   0.645
FEMALE(MELANIC)   4    0.060705    0.060705    0.015176    2.52   0.062
Error            30    0.180833    0.180833    0.006028
Total            35    0.245300

```

S = 0.0776387 R-Sq = 26.28% R-Sq(adj) = 13.99%

Expected Mean Squares, using Adjusted SS

Source	Expected Mean Square for Each Term
1 MELANIC	(3) + 6.0000 (2) + Q[1]
2 FEMALE(MELANIC)	(3) + 6.0000 (2)
3 Error	(3)

Error Terms for Tests, using Adjusted SS

Source	Error DF	Error MS	Synthesis of Error MS
1 MELANIC	4.00	0.015176	(2)
2 FEMALE(MELANIC)	30.00	0.006028	(3)

Variance Components, using Adjusted SS

Source	Estimated Value
FEMALE(MELANIC)	0.00152
Error	0.00603

<< R printout >>

```

> modell0 <- lmer(WATTLE~MELANIC+(1|MELANIC:FEMALE))
> summary(modell0)
Linear mixed model fit by REML ['lmerMod']
Formula: WATTLE ~ MELANIC + (1 | MELANIC:FEMALE)

REML criterion at convergence: -67.8249

Random effects:
 Groups          Name              Variance Std.Dev.
MELANIC:FEMALE (Intercept)  0.001525  0.03905
Residual              0.006028  0.07764
Number of obs: 36, groups: MELANIC:FEMALE, 6

Fixed effects:
              Estimate Std. Error t value
(Intercept)   0.56672    0.06493   8.729
MELANIC       -0.02044    0.04106  -0.498

Correlation of Fixed Effects:
      (Intr)
MELANIC -0.949

> modell0 <- lmer(WATTLE~MELANIC + (1|MELANIC:FEMALE), REML=F)
> modell1 <- lmer(WATTLE~ (1|MELANIC:FEMALE), REML=F)
> anova(modell0, modell1)
Data:
Models:
modell1: WATTLE ~ (1 | MELANIC:FEMALE)
modell0: WATTLE ~ MELANIC + (1 | MELANIC:FEMALE)
              Df      AIC      BIC logLik deviance  Chisq Chi Df Pr(>Chisq)
modell1      3 -72.378 -67.627 39.189  -78.378
modell0      4 -70.739 -64.405 39.369  -78.739  0.3607      1    0.5481

> AIC(modell0, modell1)
              df      AIC
modell0      4 -70.73871
modell1      3 -72.37796

```

```

> modell1 <-
lme(WATTLE~MELANIC, random=~1 | MELANIC/FEMALE)
> summary(modell1)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC    logLik
-57.82488 -50.19308 33.91244

Random effects:
Formula: ~1 | MELANIC
      (Intercept)
StdDev:      0.048799

Formula: ~1 | FEMALE %in% MELANIC
      (Intercept) Residual
StdDev:  0.03904806 0.0776387 #0.03904806^2=0.001524751 0.0776387^2=0.0060278

Fixed effects: WATTLE ~ MELANIC
              Value Std.Error DF   t-value p-value
(Intercept)  0.5667222 0.12697379 30   4.463301 1e-04
MELANIC      -0.0204444 0.08030527  0  -0.254584   NaN
Correlation:
      (Intr)
MELANIC -0.949

Number of Observations: 36
Number of Groups:
      MELANIC FEMALE %in% MELANIC
           2             6

```