lmer for **SAS PROC MIXED** Users

Douglas Bates
Department of Statistics
University of Wisconsin – Madison
`Bates@wisc.edu`

1   Introduction

The `lmer` function from the `lme4` package for R is used to fit linear mixed-effects models. It is similar in scope to the **SAS** procedure `PROC MIXED` described in Littell et al. (1996).

A file on the SAS Institute web site (http://www.sas.com) contains all the data sets in the book and all the SAS programs used in Littell et al. (1996). We have converted the data sets from the tabular representation used for SAS to the `data.frame` objects used by `lmer`. To help users familiar with **SAS PROC MIXED** get up to speed with `lmer` more quickly, we provide transcripts of some `lmer` analyses paralleling the **SAS PROC MIXED** analyses in Littell et al. (1996).

In this paper we highlight some of the similarities and differences of `lmer` analysis and **SAS PROC MIXED** analysis.

2   Similarities between lmer and SAS PROC MIXED

Both `SAS PROC MIXED` and `lmer` can fit linear mixed-effects models expressed in the Laird-Ware formulation. For a single level of grouping Laird and Ware (1982) write the n_i-dimensional response vector $\boldsymbol{y}_i$ for the ith experimental

unit as

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \dots, M \tag{1}$$
$$\boldsymbol{b}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\epsilon}_i \sim \mathcal{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$$

where $\boldsymbol{\beta}$ is the p-dimensional vector of *fixed effects*, $\boldsymbol{b}_i$ is the q-dimensional vector of *random effects*, $\boldsymbol{X}_i$ (of size $n_i \times p$) and $\boldsymbol{Z}_i$ (of size $n_i \times q$) are known fixed-effects and random-effects regressor matrices, and $\boldsymbol{\epsilon}_i$ is the n_i-dimensional *within-group error* vector with a spherical Gaussian distribution. The assumption $\text{Var}(\boldsymbol{\epsilon}_i) = \sigma^2\boldsymbol{I}$ can be relaxed using additional arguments in the model fitting.

The basic specification of the model requires a linear model expression for the fixed effects and a linear model expression for the random effects. In `SAS PROC MIXED` the fixed-effects part is specified in the `model` statement and the random-effects part in the `random` statement. In `lmer` the fixed effects and the random effects are both specified as terms in the `formula` argument to `lmer`.

Both `SAS PROC MIXED` and `lmer` allow a mixed-effects model to be fit by maximum likelihood (`method = ml` in SAS) or by maximum residual likelihood, sometimes also called restricted maximum likelihood or REML. This is the default criterion in `SAS PROC MIXED` and in `lmer`. To get ML estimates use the optional argument `REML=FALSE` in the call to `lmer`.

3 Important differences

The output from `PROC MIXED` typically includes values of the Akaike Information Criterion (AIC) and Schwartz's Bayesian Criterion (SBC). These are used to compare different models fit to the same data. The output of the `summary` function applied to the object created by `lmer` also produces values of AIC and BIC but the definitions used in older versions of `PROC MIXED` are different from those used in more recent versions of `PROC MIXED` and in `lmer`. In `lmer` the definitions are such that "smaller is better". In some older versions of `PROC MIXED` the definitions are such that "bigger is better".

When models are fit by REML, the values of AIC, SBC (or BIC) and the log-likelihood can only be compared between models with exactly the same fixed-effects structure. When models are fit by maximum likelihood these criteria can be compared between any models fit to the same data. That is,

these quality-of-fit criteria can be used to evaluate different fixed-effects specifications or different random-effects specifications or different specifications of both fixed effects and random effects.

We encourage developing and testing the model using likelihood ratio tests or the AIC and BIC criteria. Once a form for both the random effects and the fixed effects has been determined, the model can be refit with `REML = TRUE` if the restricted estimates of the variance components are desired. Note that the `update` function provides a convenient way of refitting a model with changes to one or more arguments.

4 Data manipulation

Both `PROC MIXED` and `lmer` work with data in a tabular form with one row per observation. There are, however, important differences in the internal representations of variables in the data.

In SAS a qualitative factor can be stored either as numerical values or alphanumeric labels. When a factor stored as numerical values is used in `PROC MIXED` it is listed in the `class` statement to indicate that it is a factor. In S this information is stored with the data itself by converting the variable to a factor when it is first stored. If the factor represents an ordered set of levels, it should be converted to an `ordered` factor.

For example the SAS code

```
data animal;
 input trait animal y;
 datalines;
1 1 6
1 2 8
1 3 7
2 1 9
2 2 5
2 3 .
;
```

would require that the `trait` and `animal` variables be specified in a class statement in any model that is fit.

In R these data could be read from a file, say `animal.dat`, and converted to factors by

```
animal <- within(read.table("animal.dat", header = TRUE),
          {
```

```
              trait <- factor(trait)
              animal <- factor(animal)
        })
```

In general it is a good idea to check the types of variables in a data frame before working with it. One way of doing this is to apply the function data.class to each variable in turn using the sapply function.

```
> sapply(Animal, data.class)
        Sire          Dam AvgDailyGain
    "factor"     "factor"    "numeric"
> str(Animal)
'data.frame':         20 obs. of  3 variables:
 $ Sire        : Factor w/ 5 levels "1","2","3","4",..: 1 1 1 1 2 2 2 2 3 3 .
 $ Dam         : Factor w/ 2 levels "1","2": 1 1 2 2 1 1 2 2 1 1 ...
 $ AvgDailyGain: num  2.24 1.85 2.05 2.41 1.99 1.93 2.72 2.32 2.33 2.68 ...
 - attr(*, "ginfo")=List of 7
  ..$ formula    :Class 'formula' length 3 AvgDailyGain ~ 1 | Sire/Dam
  .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
  ..$ order.groups:List of 2
  .. ..$ Sire: logi TRUE
  .. ..$ Dam : logi TRUE
  ..$ FUN        :function (x)
  ..$ outer      : NULL
  ..$ inner      : NULL
  ..$ labels     :List of 1
  .. ..$ AvgDailyGain: chr "Average Daily Weight Gain"
  ..$ units      : list()
```

4.1   Unique levels of factors

Designs with nested grouping factors are indicated differently in the two languages. An example of such an experimental design is the semiconductor experiment described in section 2.2 of Littell et al. (1996) where twelve wafers are assigned to four experimental treatments with three wafers per treatment. The levels for the wafer factor are 1, 2, and 3 but the wafer factor is only meaningful within the same level of the treatment factor, et. There is nothing associating wafer 1 of the third treatment group with wafer 1 of the first treatment group.

In SAS this nesting of factors is denoted by wafer(et). In S the nesting is written with   ET/Wafer and read "wafer within ET". If both levels of nested

factors are to be associated with random effects then this is all you need to know. You would use an expression with a `"/"` in the grouping factor part of the formula in the call to `lmer` object. The random effects term would be either

```
(1 | ET/Wafer)
```

or, equivalently

```
(1 | ET:Wafer) + (1 | ET)
```

In this case, however, there would not usually be any random effects associated with the "experimental treatment" or `ET` factor. The only random effects are at the `Wafer` level. It is necessary to create a factor that will have unique levels for each `Wafer` within each level of `ET`. One way to do this is to assign

```
> Semiconductor <- within(Semiconductor, Grp <- factor(ET:Wafer))
```

after which we could specify a random effects term of `(1 | Grp)`. Alternatively, we can use the explicit term

```
(1 | ET:Wafer)
```

4.2 General approach

As a general approach to importing data into R for mixed-effects analysis you should:

- Create a `data.frame` with one row per observation and one column per variable.

- Use `factor` or `as.factor` to explicitly convert any ordered factors to class `ordered`.

- Use `ordered` or `as.ordered` to explicitly convert any ordered factors to class `ordered`.

- If necessary, use interaction terms to create a factor with unique levels from inner nested factors.

- Plot the data. Plot it several ways. The use of lattice graphics is closely integrated with the `lme4` library. Lattice plots can provide invaluable insight into the structure of the data. Use them.

5

5 Contrasts

When comparing estimates produced by SAS PROC MIXED and by lmer one must be careful to consider the contrasts that are used to define the effects of factors. In SAS a model with an intercept and a qualitative factor is defined in terms of the intercept and the indicator variables for all but the last level of the factor. The default behaviour in S is to use the Helmert contrasts for the factor. On a balanced factor these provide a set of orthogonal contrasts. In R the default is the "treatment" contrasts which are almost the same as the SAS parameterization except that they drop the indicator of the first level, not the last level.

When in doubt, check which contrasts are being used with the contrasts function.

To make comparisons easier, you may find it worthwhile to declare

```
> options(contrasts = c(factor = "contr.SAS", ordered = "contr.poly"))
```

at the beginning of your session.

References

Nan M. Laird and James H. Ware. Random-effects models for longitudinal data. *Biometrics*, 38:963–974, 1982.

Ramon C. Littell, George A. Milliken, Walter W. Stroup, and Russell D. Wolfinger. *SAS System for Mixed Models*. SAS Institute, Inc., 1996.

A AvgDailyGain

```
> print(xyplot(adg ~ Treatment | Block, AvgDailyGain, type = c("g", "p", "r")
+        xlab = "Treatment (amount of feed additive)",
+        ylab = "Average daily weight gain (lb.)", aspect = "xy",
+        index.cond = function(x, y) coef(lm(y ~ x))[1]))

> ## compare with output 5.1, p. 178
> (fm1Adg <- lmer(adg ~ (Treatment - 1)*InitWt + (1 | Block), AvgDailyGain))
Linear mixed model fit by REML ['lmerMod']
Formula: adg ~ (Treatment - 1) * InitWt + (1 | Block)
   Data: AvgDailyGain
```
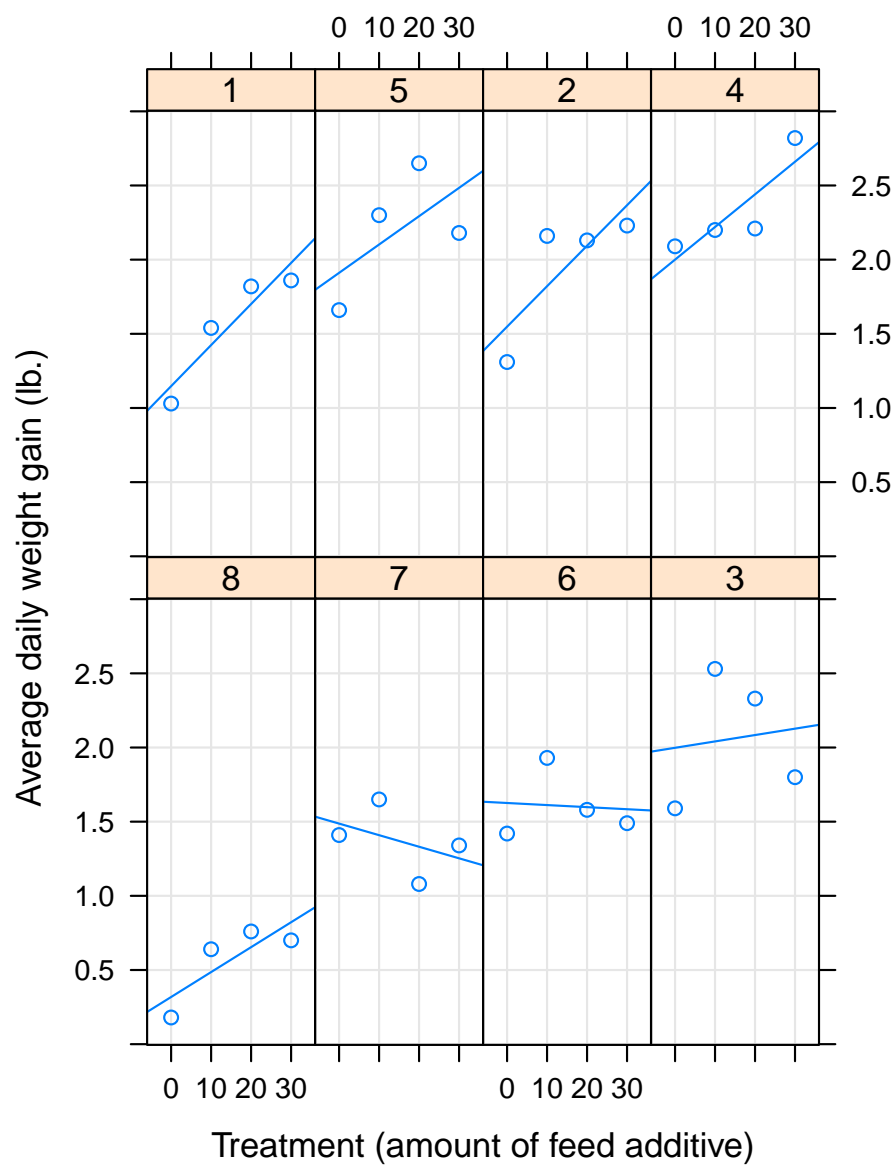
Figure 1: Average daily weight gain

7

```
REML criterion at convergence: 65.3268

Random effects:
 Groups    Name         Variance Std.Dev.
 Block     (Intercept) 0.25931  0.5092
 Residual               0.04943  0.2223
Number of obs: 32, groups: Block, 8

Fixed effects:
                    Estimate Std. Error t value
Treatment0          0.439137   0.711088    0.618
Treatment10         1.426118   0.637546    2.237
Treatment20         0.479628   0.548887    0.874
Treatment30         0.200107   0.775199    0.258
InitWt              0.004448   0.002082    2.137
Treatment0:InitWt  -0.002154   0.002786   -0.773
Treatment10:InitWt -0.003365   0.002515   -1.338
Treatment20:InitWt -0.001082   0.002488   -0.435

Correlation of Fixed Effects:
            Trtmn0 Trtm10 Trtm20 Trtm30 InitWt Tr0:IW T10:IW
Treatment10  0.039
Treatment20  0.080  0.334
Treatment30  0.011  0.097  0.043
InitWt       0.050 -0.032  0.035 -0.967
Trtmnt0:InW -0.640  0.046 -0.024  0.754 -0.780
Trtmnt10:IW -0.021 -0.534 -0.178  0.781 -0.808  0.617
Trtmnt20:IW -0.040 -0.106 -0.512  0.828 -0.856  0.666  0.775
> anova(fm1Adg)    # checking significance of terms
Analysis of Variance Table
                 Df Sum Sq Mean Sq F value
Treatment         4 5.7248 1.43119 28.9543
InitWt            1 0.5495 0.54953 11.1175
Treatment:InitWt  3 0.1381 0.04603  0.9312
> ## common slope model
> (fm2Adg <- lmer(adg ~ InitWt + Treatment + (1 | Block), AvgDailyGain))
Linear mixed model fit by REML ['lmerMod']
Formula: adg ~ InitWt + Treatment + (1 | Block)
   Data: AvgDailyGain
```

```
REML criterion at convergence: 36.3373

Random effects:
 Groups    Name          Variance Std.Dev.
 Block     (Intercept) 0.24084  0.4908
 Residual              0.05008  0.2238
Number of obs: 32, groups: Block, 8

Fixed effects:
             Estimate Std. Error t value
(Intercept)  0.8011082  0.3556610    2.252
InitWt       0.0027797  0.0008334    3.336
Treatment0  -0.5520737  0.1148131   -4.808
Treatment10 -0.0685661  0.1189690   -0.576
Treatment20 -0.0881291  0.1162878   -0.758

Correlation of Fixed Effects:
            (Intr) InitWt Trtmn0 Trtm10
InitWt      -0.844
Treatment0   0.036 -0.224
Treatment10  0.139 -0.340  0.534
Treatment20  0.079 -0.272  0.530  0.545
> anova(fm2Adg)
Analysis of Variance Table
          Df  Sum Sq Mean Sq F value
InitWt     1 0.51455 0.51455  10.275
Treatment  3 1.52670 0.50890  10.162
> (fm3Adg <- lmer(adg ~ InitWt + Treatment - 1 + (1 | Block), AvgDailyGain))
Linear mixed model fit by REML ['lmerMod']
Formula: adg ~ InitWt + Treatment - 1 + (1 | Block)
   Data: AvgDailyGain

REML criterion at convergence: 36.3373

Random effects:
 Groups    Name          Variance Std.Dev.
 Block     (Intercept) 0.24084  0.4908
 Residual              0.05008  0.2238
Number of obs: 32, groups: Block, 8
```

```
Fixed effects:
            Estimate Std. Error t value
InitWt      0.0027797  0.0008334   3.336
Treatment0  0.2490346  0.3776318   0.659
Treatment10 0.7325421  0.3903798   1.876
Treatment20 0.7129791  0.3827685   1.863
Treatment30 0.8011082  0.3556610   2.252


Correlation of Fixed Effects:
            InitWt Trtmn0 Trtm10 Trtm20
Treatment0  −0.863
Treatment10 −0.873  0.957
Treatment20 −0.867  0.957  0.958
Treatment30 −0.844  0.953  0.953  0.953
```

B BIB

```
> print(xyplot(y ~ x | Block, BIB, groups = Treatment, type = c("g", "p"),
+              aspect = "xy", auto.key = list(points = TRUE, space = "right",
+              lines = FALSE)))


> ## compare with Output 5.7, p. 188
> (fm1BIB <- lmer(y ~ Treatment * x + (1|Block), BIB))
Linear mixed model fit by REML ['lmerMod']
Formula: y ~ Treatment * x + (1 | Block)
   Data: BIB


REML criterion at convergence: 104.8945


Random effects:
 Groups   Name        Variance Std.Dev.
 Block    (Intercept) 18.25    4.272
 Residual              1.20    1.096
Number of obs: 24, groups: Block, 8


Fixed effects:
            Estimate Std. Error t value
(Intercept) 22.36784    3.10182   7.211
Treatment1   4.42949    3.36503   1.316
```
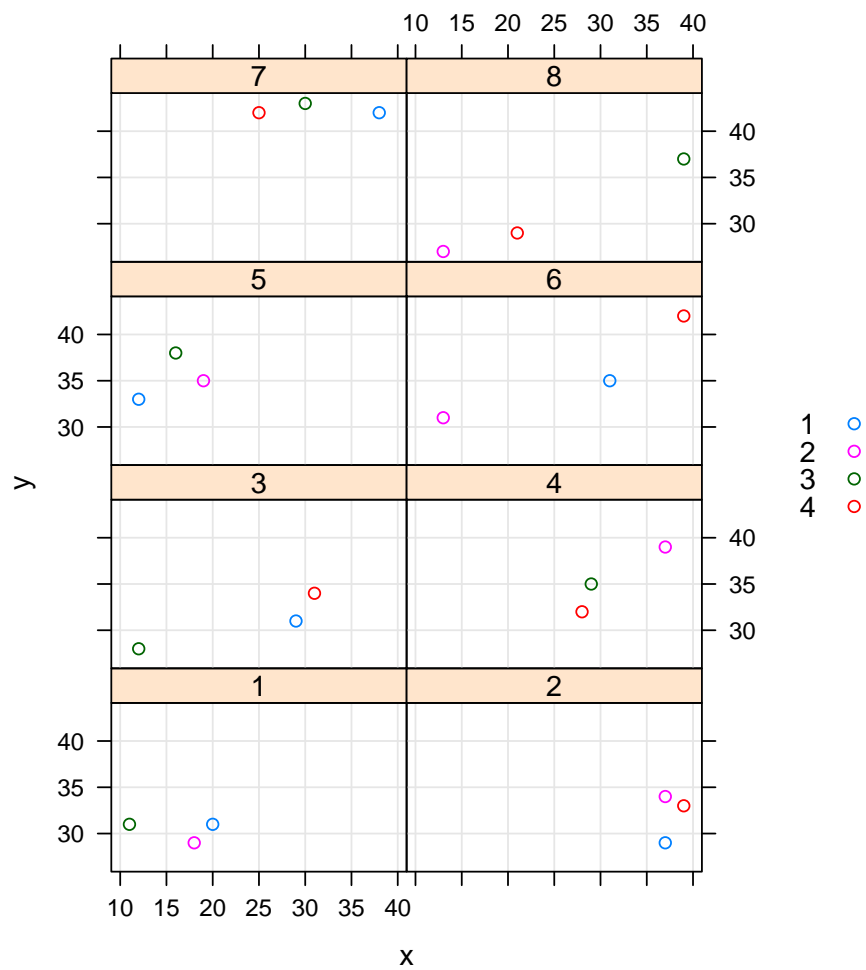
Figure 2: Balanced incomplete block design

11

```
Treatment2   -0.43737    2.93319  -0.149
Treatment3    6.27864    3.28202   1.913
x             0.44255    0.08706   5.083
Treatment1:x -0.22377    0.10608  -2.109
Treatment2:x  0.05338    0.09714   0.550
Treatment3:x -0.17918    0.11571  -1.549


Correlation of Fixed Effects:
           (Intr) Trtmn1 Trtmn2 Trtmn3 x       Trtm1: Trtm2:
Treatment1 -0.728
Treatment2 -0.778  0.797
Treatment3 -0.796  0.827  0.826
x          -0.859  0.797  0.865  0.886
Treatmnt1:x 0.709 -0.979 -0.774 -0.797 -0.799
Treatmnt2:x 0.722 -0.731 -0.965 -0.763 -0.829  0.729
Treatmnt3:x 0.769 -0.789 -0.790 -0.976 -0.879  0.777  0.748
> anova(fm1BIB)      # strong evidence of different slopes
Analysis of Variance Table
           Df  Sum Sq Mean Sq  F value
Treatment   3  23.447   7.816   6.5110
x           1 136.809 136.809 113.9695
Treatment:x 3  18.427   6.142   5.1169
> ## compare with Output 5.9, p. 193
> (fm2BIB <- lmer(y ~ Treatment + x:Grp + (1|Block), BIB))
Linear mixed model fit by REML ['lmerMod']
Formula: y ~ Treatment + x:Grp + (1 | Block)
   Data: BIB


REML criterion at convergence: 99.177


Random effects:
 Groups   Name        Variance Std.Dev.
 Block    (Intercept) 18.526   4.304
 Residual              1.038   1.019
Number of obs: 24, groups: Block, 8


Fixed effects:
            Estimate Std. Error t value
(Intercept) 20.94516    2.06230  10.156
Treatment1   5.34145    1.97570   2.704
```

```
Treatment2    1.13557    0.71399    1.590
Treatment3    8.18103    1.77010    4.622
x:Grp13       0.23952    0.04296    5.575
x:Grp24       0.48923    0.04412   11.088


Correlation of Fixed Effects:
          (Intr) Trtmn1 Trtmn2 Trtmn3 x:Gr13
Treatment1 -0.501
Treatment2 -0.431  0.559
Treatment3 -0.527  0.942  0.581
x:Grp13     0.027 -0.663 -0.165 -0.605
x:Grp24    -0.639  0.651  0.452  0.688  0.042
> anova(fm2BIB)
Analysis of Variance Table
          Df   Sum Sq Mean Sq F value
Treatment  3   23.424   7.808  7.5236
x:Grp      2  154.733  77.367 74.5471
```

C   Bond

```
> ## compare with output 1.1 on p. 6
> (fm1Bond <- lmer(pressure ~ Metal + (1|Ingot), Bond))
Linear mixed model fit by REML ['lmerMod']
Formula: pressure ~ Metal + (1 | Ingot)
   Data: Bond


REML criterion at convergence: 107.7902


Random effects:
 Groups   Name        Variance Std.Dev.
 Ingot    (Intercept) 11.45    3.383
 Residual             10.37    3.220
Number of obs: 21, groups: Ingot, 7


Fixed effects:
            Estimate Std. Error t value
(Intercept)  71.1000     1.7655   40.27
Metalc       -0.9143     1.7214   -0.53
Metali        4.8000     1.7214    2.79
```

```
Correlation of Fixed Effects:
       (Intr) Metalc
Metalc -0.488
Metali -0.488  0.500
> anova(fm1Bond)
Analysis of Variance Table
      Df Sum Sq Mean Sq F value
Metal  2  131.9   65.95  6.3587
```

D   Cultivation

```
> str(Cultivation)
'data.frame':         24 obs. of  4 variables:
 $ Block: Factor w/ 4 levels "1","2","3","4": 1 1 1 1 1 1 2 2 2 2 ...
 $ Cult : Factor w/ 2 levels "a","b": 1 1 1 2 2 2 1 1 1 2 ...
 $ Inoc : Factor w/ 3 levels "con","dea","liv": 1 2 3 1 2 3 1 2 3 1 ...
 $ drywt: num  27.4 29.7 34.5 29.4 32.5 34.4 28.9 28.7 33.4 28.7 ...
 - attr(*, "ginfo")=List of 7
  ..$ formula     :Class 'formula' length 3 drywt ~ 1 | Block/Cult
  .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
  ..$ order.groups:List of 2
  .. ..$ Block: logi TRUE
  .. ..$ Cult : logi TRUE
  ..$ FUN         :function (x)
  ..$ outer       : NULL
  ..$ inner       :List of 1
  .. ..$ Cult:Class 'formula' length 2 ~Inoc
  .. .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
  ..$ labels      :List of 1
  .. ..$ drywt: chr "Yield"
  ..$ units       : list()
> xtabs(~Block+Cult, Cultivation)
     Cult
Block a b
    1 3 3
    2 3 3
    3 3 3
    4 3 3
> (fm1Cult <- lmer(drywt ~ Inoc * Cult + (1|Block) + (1|Cult), Cultivation))
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: drywt ~ Inoc * Cult + (1 | Block) + (1 | Cult)
   Data: Cultivation

REML criterion at convergence: 68.4874

Random effects:
 Groups    Name        Variance Std.Dev.
 Block     (Intercept) 1.207    1.099
 Cult      (Intercept) 1.210    1.100
 Residual              1.196    1.094
Number of obs: 24, groups: Block, 4; Cult, 2

Fixed effects:
              Estimate Std. Error t value
(Intercept)    33.5250     1.3458  24.911
Inoccon        -5.5000     0.7734  -7.111
Inocdea        -2.8750     0.7734  -3.717
Culta          -0.3750     1.7375  -0.216
Inoccon:Culta   0.2500     1.0938   0.229
Inocdea:Culta  -1.0250     1.0938  -0.937

Correlation of Fixed Effects:
            (Intr) Inoccn Inocde Culta  Incc:C
Inoccon     -0.287
Inocdea     -0.287  0.500
Culta       -0.646  0.223  0.223
Inoccon:Clt  0.203 -0.707 -0.354 -0.315
Inocdea:Clt  0.203 -0.354 -0.707 -0.315  0.500
> anova(fm1Cult)
Analysis of Variance Table
          Df  Sum Sq Mean Sq F value
Inoc       2 118.176  59.088 49.3908
Cult       1   0.183   0.183  0.1531
Inoc:Cult  2   1.826   0.913  0.7631
> (fm2Cult <- lmer(drywt ~ Inoc + Cult + (1|Block) + (1|Cult), Cultivation))
Linear mixed model fit by REML ['lmerMod']
Formula: drywt ~ Inoc + Cult + (1 | Block) + (1 | Cult)
   Data: Cultivation
```

```
REML criterion at convergence: 73.7535

Random effects:
 Groups    Name         Variance Std.Dev.
 Block     (Intercept) 1.213     1.101
 Cult      (Intercept) 1.143     1.069
 Residual               1.163     1.078
Number of obs: 24, groups: Block, 4; Cult, 2

Fixed effects:
            Estimate Std. Error t value
(Intercept)  33.6542     1.2805  26.281
Inoccon      -5.3750     0.5392  -9.968
Inocdea      -3.3875     0.5392  -6.282
Culta        -0.6333     1.5746  -0.402

Correlation of Fixed Effects:
        (Intr) Inoccn Inocde
Inoccon -0.211
Inocdea -0.211  0.500
Culta   -0.615  0.000  0.000
> anova(fm2Cult)
Analysis of Variance Table
     Df  Sum Sq Mean Sq F value
Inoc  2 118.176  59.088 50.8069
Cult  1   0.188   0.188  0.1618
> (fm3Cult <- lmer(drywt ~ Inoc + (1|Block) + (1|Cult), Cultivation))
Linear mixed model fit by REML ['lmerMod']
Formula: drywt ~ Inoc + (1 | Block) + (1 | Cult)
   Data: Cultivation

REML criterion at convergence: 75.6778

Random effects:
 Groups    Name         Variance Std.Dev.
 Block     (Intercept) 1.2128    1.1013
 Cult      (Intercept) 0.1036    0.3219
 Residual               1.1630    1.0784
Number of obs: 24, groups: Block, 4; Cult, 2
```

```
Fixed effects:
            Estimate Std. Error t value
(Intercept)  33.3375      0.7074   47.13
Inoccon      -5.3750      0.5392   -9.97
Inocdea      -3.3875      0.5392   -6.28


Correlation of Fixed Effects:
        (Intr) Inoccn
Inoccon -0.381
Inocdea -0.381  0.500
> anova(fm3Cult)
Analysis of Variance Table
     Df Sum Sq Mean Sq F value
Inoc  2 118.18  59.088  50.807
```

E   Demand

```
> ## compare to output 3.13, p. 132
> (fm1Demand <-
+  lmer(log(d) ~ log(y) + log(rd) + log(rt) + log(rs) + (1|State) + (1|Year),
+       Demand))
Linear mixed model fit by REML ['lmerMod']
Formula: log(d) ~ log(y) + log(rd) + log(rt) + log(rs) + (1 | State) +        (
   Data: Demand

REML criterion at convergence: -240.1653

Random effects:
 Groups   Name        Variance  Std.Dev.
 Year     (Intercept) 0.0002647 0.01627
 State    (Intercept) 0.0295054 0.17177
 Residual             0.0011170 0.03342
Number of obs: 77, groups: Year, 11; State, 7

Fixed effects:
            Estimate Std. Error t value
(Intercept) -1.28382    0.72343  -1.775
log(y)       1.06978    0.10393  10.294
log(rd)     -0.29533    0.05246  -5.629
log(rt)      0.03988    0.02789   1.430
```

```
log(rs)      -0.32673    0.11438  -2.856
```

Correlation of Fixed Effects:
```
        (Intr) log(y) lg(rd) lg(rt)
log(y)  -0.976
log(rd)  0.383 -0.227
log(rt)  0.077 -0.062 -0.337
log(rs)  0.444 -0.600 -0.270 -0.323
```

F   HR

```
> ## linear trend in time
> (fm1HR <- lmer(HR ~ Time * Drug + baseHR + (Time|Patient), HR))
Linear mixed model fit by REML ['lmerMod']
Formula: HR ~ Time * Drug + baseHR + (Time | Patient)
   Data: HR

REML criterion at convergence: 767.607

Random effects:
 Groups    Name        Variance Std.Dev. Corr
 Patient   (Intercept) 60.63    7.787
           Time        37.78    6.147    -0.563
 Residual              24.36    4.936
Number of obs: 120, groups: Patient, 24

Fixed effects:
           Estimate Std. Error t value
(Intercept) 33.9776    10.2830   3.304
Time        -3.1970     3.0849  -1.036
Druga        3.5992     4.2314   0.851
Drugb        7.0912     4.2094   1.685
baseHR       0.5434     0.1161   4.679
Time:Druga  -7.5013     4.3627  -1.719
Time:Drugb  -3.9894     4.3627  -0.914

Correlation of Fixed Effects:
         (Intr) Time   Druga  Drugb  baseHR Tim:Drg
Time     -0.162
Druga    -0.308  0.394
```

18

```
Drugb        -0.244  0.396  0.501
baseHR       -0.957  0.000  0.110  0.041
Time:Druga   0.115 -0.707 -0.557 -0.280  0.000
Time:Drugb   0.115 -0.707 -0.278 -0.560  0.000  0.500
> anova(fm1HR)
Analysis of Variance Table
         Df Sum Sq Mean Sq F value
Time      1 379.23  379.23 15.5671
Drug      2  92.88   46.44  1.9064
baseHR    1 533.27  533.27 21.8905
Time:Drug 2  72.12   36.06  1.4802
> ## remove interaction
> (fm3HR <- lmer(HR ~ Time + Drug + baseHR + (Time|Patient), HR))
Linear mixed model fit by REML ['lmerMod']
Formula: HR ~ Time + Drug + baseHR + (Time | Patient)
   Data: HR

REML criterion at convergence: 779.8283

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Patient  (Intercept) 61.56    7.846
          Time        40.96    6.400    -0.570
 Residual             24.36    4.936
Number of obs: 120, groups: Patient, 24

Fixed effects:
            Estimate Std. Error t value
(Intercept)  36.0463    10.1945   3.536
Time         -7.0273     1.8179  -3.866
Druga        -0.4524     3.5146  -0.129
Drugb         4.9365     3.4881   1.415
baseHR        0.5434     0.1161   4.679

Correlation of Fixed Effects:
      (Intr) Time   Druga  Drugb
Time  -0.096
Druga -0.297  0.000
Drugb -0.219  0.000  0.502
baseHR -0.966 0.000  0.132  0.050
```

```
> anova(fm3HR)
Analysis of Variance Table
       Df Sum Sq Mean Sq F value
Time    1 364.02  364.02 14.9431
Drug    2  92.88   46.44  1.9064
baseHR  1 533.27  533.27 21.8905
> ## remove Drug term
> (fm4HR <- lmer(HR ~ Time + baseHR + (Time|Patient), HR))
Linear mixed model fit by REML ['lmerMod']
Formula: HR ~ Time + baseHR + (Time | Patient)
   Data: HR

REML criterion at convergence: 791.1481

Random effects:
 Groups    Name        Variance Std.Dev. Corr
 Patient   (Intercept) 63.03    7.939
           Time        40.96    6.400    -0.553
 Residual              24.36    4.936
Number of obs: 120, groups: Patient, 24

Fixed effects:
            Estimate Std. Error t value
(Intercept)  36.9313     9.9014   3.730
Time         -7.0273     1.8179  -3.866
baseHR        0.5508     0.1175   4.686

Correlation of Fixed Effects:
       (Intr) Time
Time   -0.098
baseHR -0.984  0.000
> anova(fm4HR)
Analysis of Variance Table
       Df Sum Sq Mean Sq F value
Time    1 364.03  364.03  14.943
baseHR  1 534.87  534.87  21.956
```

G   Mississippi

```
> ## compare with output 4.1, p. 142
> (fm1Miss <- lmer(y ~ 1 + (1 | influent), Mississippi))
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: y ~ 1 + (1 | influent)
   Data: Mississippi

REML criterion at convergence: 252.3511

Random effects:
 Groups   Name        Variance Std.Dev.
 influent (Intercept) 63.32    7.958
 Residual             42.66    6.531
Number of obs: 37, groups: influent, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)   21.223      3.429   6.189
> ## compare with output 4.2, p. 143
> (fm1MLMiss <- lmer(y ~ 1 + (1 | influent), Mississippi, REML=FALSE))
Linear mixed model fit by maximum likelihood ['lmerMod']
Formula: y ~ 1 + (1 | influent)
   Data: Mississippi

      AIC       BIC     logLik  deviance
 262.5570  267.3898 -128.2785  256.5570

Random effects:
 Groups   Name        Variance Std.Dev.
 influent (Intercept) 51.25    7.159
 Residual             42.70    6.534
Number of obs: 37, groups: influent, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)   21.217      3.122   6.796
> ranef(fm1MLMiss)          # BLUP's of random effects on p. 144
$influent
  (Intercept)
1   0.3097833
2  -6.5772239
3  -3.7862717
4   2.8826693
```

```
5   -5.8435163
6   13.0145592

attr(,"class")
[1] "ranef.mer"
> ranef(fm1Miss)                  # BLUP's of random effects on p. 142
$influent
   (Intercept)
1     0.309286
2    -6.719325
3    -3.897940
4     2.946101
5    -6.012976
6    13.374854

attr(,"class")
[1] "ranef.mer"
> VarCorr(fm1Miss)                # compare to output 4.7, p. 148
 Groups    Name        Variance Std.Dev.
 influent (Intercept) 63.323    7.9576
 Residual              42.658    6.5313
> ## compare to output 4.8 and 4.9, pp. 150-152
> (fm2Miss <- lmer(y ~ Type + (1 | influent), Mississippi))
Linear mixed model fit by REML ['lmerMod']
Formula: y ~ Type + (1 | influent)
   Data: Mississippi

REML criterion at convergence: 234.5246

Random effects:
 Groups    Name        Variance Std.Dev.
 influent (Intercept) 14.97     3.869
 Residual             42.51     6.520
Number of obs: 37, groups: influent, 6

Fixed effects:
            Estimate Std. Error t value
(Intercept)   36.400      4.845    7.513
Type1        -20.800      5.934   -3.505
Type2        -16.462      5.517   -2.984
```

```
Correlation of Fixed Effects:
      (Intr) Type1
Type1 -0.816
Type2 -0.878  0.717
> anova(fm2Miss)
Analysis of Variance Table
     Df Sum Sq Mean Sq F value
Type  2 541.75  270.88  6.3715
```

H   Multilocation

```
> str(Multilocation)
'data.frame':           108 obs. of  7 variables:
 $ obs     : num  3 4 6 7 9 10 12 16 19 20 ...
 $ Location: Factor w/ 9 levels "A","B","C","D",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ Block   : Factor w/ 3 levels "1","2","3": 1 1 1 1 2 2 2 2 3 3 ...
 $ Trt     : Factor w/ 4 levels "1","2","3","4": 3 4 2 1 2 1 3 4 1 2 ...
 $ Adj     : num  3.16 3.12 3.16 3.25 2.71 ...
 $ Fe      : num  7.1 6.68 6.83 6.53 8.25 ...
 $ Grp     : Factor w/ 27 levels "A/1","A/2","A/3",..: 1 1 1 1 2 2 2 2 3 3 ..
 - attr(*, "ginfo")=List of 7
  ..$ formula      :Class 'formula' length 3 Adj ~ 1 | Location/Block
  .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
  ..$ order.groups:List of 2
  .. ..$ Location: logi TRUE
  .. ..$ Block   : logi TRUE
  ..$ FUN          :function (x)
  ..$ outer        : NULL
  ..$ inner        :List of 1
  .. ..$ Block:Class 'formula' length 2 ~Trt
  .. .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
  ..$ labels       :List of 1
  .. ..$ Adj: chr "Adjusted yield"
  ..$ units        : list()
> ### Create a Block %in% Location factor
> Multilocation$Grp <- with(Multilocation, Block:Location)
> (fm1Mult <- lmer(Adj ~ Location * Trt + (1|Grp), Multilocation))
Linear mixed model fit by REML ['lmerMod']
Formula: Adj ~ Location * Trt + (1 | Grp)
```

```
    Data: Multilocation

REML criterion at convergence: 10.6462

Random effects:
 Groups    Name         Variance Std.Dev.
 Grp       (Intercept) 0.005619 0.07496
 Residual              0.034579 0.18595
Number of obs: 108, groups: Grp, 27

Fixed effects:
               Estimate Std. Error t value
(Intercept)     2.35923    0.11575   20.381
LocationA       0.64930    0.16370    3.966
LocationB       0.06643    0.16370    0.406
LocationC       0.54533    0.16370    3.331
LocationD       0.37413    0.16370    2.285
LocationE       0.55000    0.16370    3.360
LocationF       0.99810    0.16370    6.097
LocationG       0.36057    0.16370    2.203
LocationH       1.01403    0.16370    6.194
Trt1            0.22720    0.15183    1.496
Trt2           -0.00140    0.15183   -0.009
Trt3            0.42323    0.15183    2.788
LocationA:Trt1 -0.18853    0.21472   -0.878
LocationB:Trt1 -0.27523    0.21472   -1.282
LocationC:Trt1 -0.04000    0.21472   -0.186
LocationD:Trt1 -0.53513    0.21472   -2.492
LocationE:Trt1 -0.26297    0.21472   -1.225
LocationF:Trt1 -0.27153    0.21472   -1.265
LocationG:Trt1  0.20323    0.21472    0.946
LocationH:Trt1 -0.14953    0.21472   -0.696
LocationA:Trt2 -0.09347    0.21472   -0.435
LocationB:Trt2 -0.32273    0.21472   -1.503
LocationC:Trt2  0.08960    0.21472    0.417
LocationD:Trt2 -0.29693    0.21472   -1.383
LocationE:Trt2 -0.30693    0.21472   -1.429
LocationF:Trt2 -0.30993    0.21472   -1.443
LocationG:Trt2 -0.10860    0.21472   -0.506
LocationH:Trt2 -0.33060    0.21472   -1.540
```

```
LocationA:Trt3 -0.40247     0.21472   -1.874
LocationB:Trt3 -0.56550     0.21472   -2.634
LocationC:Trt3 -0.12247     0.21472   -0.570
LocationD:Trt3 -0.54840     0.21472   -2.554
LocationE:Trt3 -0.32863     0.21472   -1.531
LocationF:Trt3 -0.46257     0.21472   -2.154
LocationG:Trt3 -0.25297     0.21472   -1.178
LocationH:Trt3 -0.37203     0.21472   -1.733


Correlation matrix not shown by default, as p = 36 > 20.
Use print(object, correlation=TRUE)   or
    vcov(object)            if you need it
> anova(fm1Mult)
Analysis of Variance Table
            Df Sum Sq Mean Sq F value
Location     8 6.9476 0.86845 25.1150
Trt          3 1.2217 0.40725 11.7774
Location:Trt 24 0.9966 0.04152  1.2008
> (fm2Mult <- lmer(Adj ~ Location + Trt + (1|Grp), Multilocation))
Linear mixed model fit by REML ['lmerMod']
Formula: Adj ~ Location + Trt + (1 | Grp)
   Data: Multilocation

REML criterion at convergence: -6.0011

Random effects:
 Groups    Name        Variance Std.Dev.
 Grp       (Intercept) 0.005085 0.07131
 Residual              0.036715 0.19161
Number of obs: 108, groups: Grp, 27

Fixed effects:
            Estimate Std. Error t value
(Intercept)  2.53296    0.07599   33.33
LocationA    0.47818    0.09752    4.90
LocationB   -0.22443    0.09752   -2.30
LocationC    0.52712    0.09752    5.41
LocationD    0.02902    0.09752    0.30
LocationE    0.32537    0.09752    3.34
LocationF    0.73709    0.09752    7.56
```

```
LocationG     0.32098     0.09752     3.29
LocationH     0.80099     0.09752     8.21
Trt1          0.05834     0.05215     1.12
Trt2         -0.18802     0.05215    -3.61
Trt3          0.08379     0.05215     1.61


Correlation of Fixed Effects:
          (Intr) LoctnA LoctnB LoctnC LoctnD LoctnE LoctnF LoctnG LoctnH
LocationA -0.642
LocationB -0.642  0.500
LocationC -0.642  0.500  0.500
LocationD -0.642  0.500  0.500  0.500
LocationE -0.642  0.500  0.500  0.500  0.500
LocationF -0.642  0.500  0.500  0.500  0.500  0.500
LocationG -0.642  0.500  0.500  0.500  0.500  0.500  0.500
LocationH -0.642  0.500  0.500  0.500  0.500  0.500  0.500  0.500
Trt1      -0.343  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000
Trt2      -0.343  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000
Trt3      -0.343  0.000  0.000  0.000  0.000  0.000  0.000  0.000  0.000
          Trt1   Trt2
LocationA
LocationB
LocationC
LocationD
LocationE
LocationF
LocationG
LocationH
Trt1
Trt2       0.500
Trt3       0.500  0.500
> (fm3Mult <- lmer(Adj ~ Location + (1|Grp), Multilocation))
Linear mixed model fit by REML ['lmerMod']
Formula: Adj ~ Location + (1 | Grp)
   Data: Multilocation


REML criterion at convergence: 9.8205


Random effects:
 Groups   Name         Variance Std.Dev.
```

```
 Grp       (Intercept) 0.001654 0.04067
 Residual               0.050439 0.22459
Number of obs: 108, groups: Grp, 27

Fixed effects:
           Estimate Std. Error t value
(Intercept)  2.52149    0.06895   36.57
LocationA    0.47818    0.09752    4.90
LocationB   -0.22443    0.09752   -2.30
LocationC    0.52712    0.09752    5.41
LocationD    0.02902    0.09752    0.30
LocationE    0.32537    0.09752    3.34
LocationF    0.73709    0.09752    7.56
LocationG    0.32098    0.09752    3.29
LocationH    0.80099    0.09752    8.21

Correlation of Fixed Effects:
          (Intr) LoctnA LoctnB LoctnC LoctnD LoctnE LoctnF LoctnG
LocationA -0.707
LocationB -0.707  0.500
LocationC -0.707  0.500  0.500
LocationD -0.707  0.500  0.500  0.500
LocationE -0.707  0.500  0.500  0.500  0.500
LocationF -0.707  0.500  0.500  0.500  0.500  0.500
LocationG -0.707  0.500  0.500  0.500  0.500  0.500  0.500
LocationH -0.707  0.500  0.500  0.500  0.500  0.500  0.500  0.500
> (fm4Mult <- lmer(Adj ~ Trt + (1|Grp), Multilocation))
Linear mixed model fit by REML ['lmerMod']
Formula: Adj ~ Trt + (1 | Grp)
   Data: Multilocation

REML criterion at convergence: 31.5057

Random effects:
 Groups    Name        Variance Std.Dev.
 Grp       (Intercept) 0.11092  0.3330
 Residual              0.03672  0.1916
Number of obs: 108, groups: Grp, 27

Fixed effects:
```

```
            Estimate Std. Error t value
(Intercept)  2.86567     0.07395   38.75
Trt1         0.05834     0.05215    1.12
Trt2        -0.18802     0.05215   -3.61
Trt3         0.08379     0.05215    1.61


Correlation of Fixed Effects:
     (Intr) Trt1   Trt2
Trt1 -0.353
Trt2 -0.353  0.500
Trt3 -0.353  0.500   0.500
> (fm5Mult <- lmer(Adj ~ 1 + (1|Grp), Multilocation))
Linear mixed model fit by REML ['lmerMod']
Formula: Adj ~ 1 + (1 | Grp)
   Data: Multilocation


REML criterion at convergence: 47.3273


Random effects:
 Groups    Name         Variance Std.Dev.
 Grp       (Intercept) 0.10749  0.3279
 Residual              0.05044  0.2246
Number of obs: 108, groups: Grp, 27


Fixed effects:
            Estimate Std. Error t value
(Intercept)  2.85419    0.06669   42.79
> anova(fm2Mult)
Analysis of Variance Table
         Df Sum Sq Mean Sq F value
Location  8 7.3768 0.92209  25.115
Trt       3 1.2217 0.40725  11.092
> (fm2MultR <- lmer(Adj ~ Trt + (Trt - 1|Location) + (1|Block), Multilocation
+               verbose = TRUE))
Linear mixed model fit by REML ['lmerMod']
Formula: Adj ~ Trt + (Trt - 1 | Location) + (1 | Block)
   Data: Multilocation


REML criterion at convergence: 1.4072
```

```
Random effects:
 Groups    Name         Variance Std.Dev. Corr
 Location  Trt1         0.13594  0.3687
           Trt2         0.10704  0.3272   0.989
           Trt3         0.11911  0.3451   0.998 0.996
           Trt4         0.11415  0.3379   0.927 0.972 0.948
 Block     (Intercept) 0.00000  0.0000
 Residual               0.03777  0.1943
Number of obs: 108, groups: Location, 9; Block, 3

Fixed effects:
            Estimate Std. Error t value
(Intercept)  2.86567    0.11867  24.149
Trt1         0.05834    0.07012   0.832
Trt2        -0.18802    0.05922  -3.175
Trt3         0.08379    0.06448   1.299

Correlation of Fixed Effects:
     (Intr) Trt1   Trt2
Trt1 -0.150
Trt2 -0.306  0.620
Trt3 -0.236  0.681  0.620
```

I PBIB

```
> str(PBIB)
'data.frame':        60 obs. of  3 variables:
 $ response : num  2.4 2.5 2.6 2 2.7 2.8 2.4 2.7 2.6 2.8 ...
 $ Treatment: Factor w/ 15 levels "1","10","11",..: 7 15 1 5 11 13 14 1 2 1 .
 $ Block    : Factor w/ 15 levels "1","10","11",..: 1 1 1 1 8 8 8 8 9 9 ...
 - attr(*, "ginfo")=List of 7
  ..$ formula     :Class 'formula' length 3 response ~ Treatment | Block
  .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
  ..$ order.groups: logi TRUE
  ..$ FUN         :function (x)
  ..$ outer       : NULL
  ..$ inner       : NULL
  ..$ labels      : list()
  ..$ units       : list()
> ## compare with output 1.7  pp. 24-25
> (fm1PBIB <- lmer(response ~ Treatment + (1 | Block), PBIB))
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: response ~ Treatment + (1 | Block)
   Data: PBIB

REML criterion at convergence: 51.9849

Random effects:
 Groups    Name         Variance Std.Dev.
 Block     (Intercept)  0.04652  0.2157
 Residual               0.08556  0.2925
Number of obs: 60, groups: Block, 15

Fixed effects:
             Estimate Std. Error t value
(Intercept)  2.891311   0.166413  17.374
Treatment1  -0.073789   0.222061  -0.332
Treatment10 -0.400250   0.222061  -1.802
Treatment11  0.007387   0.222061   0.033
Treatment12  0.161510   0.222061   0.727
Treatment13 -0.273542   0.222061  -1.232
Treatment14 -0.400000   0.227200  -1.761
Treatment15 -0.032078   0.222061  -0.144
Treatment2  -0.485996   0.222061  -2.189
Treatment3  -0.436368   0.222061  -1.965
Treatment4  -0.107482   0.227200  -0.473
Treatment5  -0.086413   0.222061  -0.389
Treatment6   0.019382   0.222061   0.087
Treatment7  -0.102327   0.222061  -0.461
Treatment8  -0.109706   0.222061  -0.494

Correlation of Fixed Effects:
            (Intr) Trtmn1 Trtm10 Trtm11 Trtm12 Trtm13 Trtm14 Trtm15 Trtmn2
Treatment1  -0.667
Treatment10 -0.667  0.500
Treatment11 -0.667  0.477  0.500
Treatment12 -0.667  0.500  0.500  0.500
Treatment13 -0.667  0.500  0.500  0.500  0.500
Treatment14 -0.683  0.512  0.512  0.512  0.512  0.512
Treatment15 -0.667  0.500  0.477  0.500  0.500  0.500  0.512
Treatment2  -0.667  0.500  0.500  0.500  0.477  0.500  0.512  0.500
```

```
Treatment3   -0.667   0.500   0.500   0.500   0.500   0.477   0.512   0.500   0.500
Treatment4   -0.683   0.512   0.512   0.512   0.512   0.512   0.500   0.512   0.512
Treatment5   -0.667   0.500   0.477   0.500   0.500   0.500   0.512   0.477   0.500
Treatment6   -0.667   0.477   0.500   0.477   0.500   0.500   0.512   0.500   0.500
Treatment7   -0.667   0.500   0.500   0.500   0.477   0.500   0.512   0.500   0.477
Treatment8   -0.667   0.500   0.500   0.500   0.500   0.477   0.512   0.500   0.500
             Trtmn3  Trtmn4  Trtmn5  Trtmn6  Trtmn7
Treatment1
Treatment10
Treatment11
Treatment12
Treatment13
Treatment14
Treatment15
Treatment2
Treatment3
Treatment4    0.512
Treatment5    0.500   0.512
Treatment6    0.500   0.512   0.500
Treatment7    0.500   0.512   0.500   0.500
Treatment8    0.477   0.512   0.500   0.500   0.500
```

J   SIMS

```
> str(SIMS)
'data.frame':        3691 obs. of  3 variables:
 $ Pretot: num  29 38 31 31 29 23 23 33 30 32 ...
 $ Gain  : num  2 0 6 6 5 9 7 2 1 3 ...
 $ Class : Factor w/ 190 levels "1","10","100",..: 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "ginfo")=List of 7
  ..$ formula      :Class 'formula' length 3 Gain ~ Pretot | Class
  .. .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
  ..$ order.groups: logi TRUE
  ..$ FUN          :function (x)
  ..$ outer        : NULL
  ..$ inner        : NULL
  ..$ labels       :List of 2
  .. ..$ Pretot: chr "Sum of pre-test core item scores"
  .. ..$ Gain  : chr "Gain in mathematics achievement score"
  ..$ units        : list()
```

31

```
> ## compare to output 7.4, p. 262
> (fm1SIMS <- lmer(Gain ~ Pretot + (Pretot | Class), SIMS))
Linear mixed model fit by REML ['lmerMod']
Formula: Gain ~ Pretot + (Pretot | Class)
   Data: SIMS

REML criterion at convergence: 22380.57

Random effects:
 Groups    Name        Variance  Std.Dev. Corr
 Class     (Intercept) 14.489513 3.80651
           Pretot       0.009203 0.09593  -0.641
 Residual              22.235739 4.71548
Number of obs: 3691, groups: Class, 190

Fixed effects:
            Estimate Std. Error t value
(Intercept)   7.0596     0.3659   19.29
Pretot       -0.1860     0.0161  -11.56

Correlation of Fixed Effects:
       (Intr)
Pretot -0.760
```