Notes for Thursday February 12, 1998

Today we consider two very common, somewhat complicated, designs:

- Latin Square designs let you account for two blocking factors, instead of the one factor accounted for by randomized block designs;
- Split-plot designs are one way to handle "nesting" in the responses (for example, pupils nested in a classroom).

1 A Latin Square Design

We have seen, with randomized block designs, that including the block in the model can eliminate its effect on the outcome, and make tests of the treatments more powerful. Carefully balanced designs can eliminate more than one blocking variable in an experiment or study. When the number of treatments equals the number of blocks, the most common design is *Latin Squares.*

Suppose that 4 cars and four drivers are employed in a study of possible differences between four gasoline additives. Even though the cars are identical models, slight systematic differences are likely to occur in their performance, and even though each driver may do her best to drive the car using the experimental protocol, slight systematic differences can occur from driver to driver. It would be desirable to design (and analyze) an experiment that eliminates both the driver-to-drive and automobile-to-automobile effects. The Latin Square arrangement below allows this to be done.

		cars			
		1	2	3	4
	I	A	B	D	C
drivers	II	D	C	A	B
	III	B	D	C	A
	IV	C	A	B	D

with cars (I,II,III, and IV), drivers (1, 2, 3, and 4) and additives [treatments] (A,B,C,and D).

Data will be assigned to each combination, so that data will be indexed by three things (car, driver, and additive), each with 4 levels, but there will only be 16 data points. We do **not** have a complete replication of the experiment.

The model might be written

$$Y_{ijk} = \mu + d_i + c_j + a_k + \epsilon_{ijk}$$

There is not much hope of estimating interactions.

Here is some data from this experiment (the data was actually artificially generated for an example in Box, Hunter and Hunter):

		cars			
		1	2	3	4
	I	A 21	B 26	D 20	C 25
drivers	II	D 23	C 26	A 20	B 27
	III	B 15	D 13	C 16	A 16
	IV	C 17	A 15	B 20	D 20

Lets do a quick analysis of these data, first entering them, and then an ANOVA and plots. Here is `gas.dat`:

```
A 21
B 26
D 20
C 25
D 23
C 26
A 20
B 27
B 15
D 13
C 16
A 16
C 17
A 15
B 20
D 20
```

and here is the data being entered.

```
402 > delme <-read.table("gas.dat")
402 > delme
   V1 V2
 1  A 21
 2  B 26
 3  D 20
 4  C 25
 5  D 23
 6  C 26
 7  A 20
 8  B 27
 9  B 15
10  D 13
11  C 16
12  A 16
13  C 17
14  A 15
15  B 20
16  D 20
402 > gas <- delme$V1
402 > resp <- delme$V2
402 > car <- rep(1:4, len=16)
402 > car
 [1] 1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4
402 > driver <- rep(1:4,rep(4,4))
```

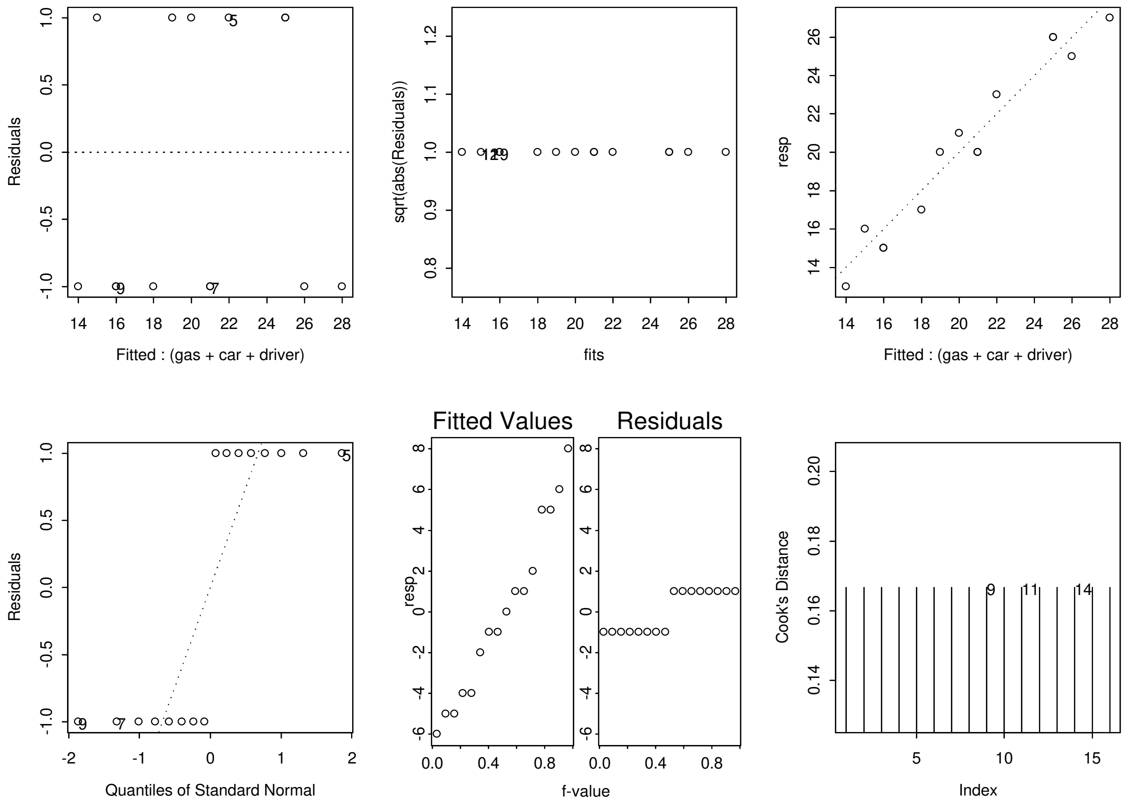
```
402 > driver
 [1] 1 1 1 1 2 2 2 2 3 3 3 3 4 4 4 4
402 > gas.frame <- data.frame(resp, gas, car=as.factor(car), driver=as.factor(driver) )
```

So here is my constructed data set

```
402 > gas.frame
   resp gas car driver
 1   21   A   1      1
 2   26   B   2      1
 3   20   D   3      1
 4   25   C   4      1
 5   23   D   1      2
 6   26   C   2      2
 7   20   A   3      2
 8   27   B   4      2
 9   15   B   1      3
10   13   D   2      3
11   16   C   3      3
12   16   A   4      3
13   17   C   1      4
14   15   A   2      4
15   20   B   3      4
16   20   D   4      4
```

Now onto an analysis.

```
402 > anova(mymod <- aov(resp ~ gas + car + driver,data=gas.frame) )
Analysis of Variance Table

Response: resp

Terms added sequentially (first to last)
          Df Sum of Sq  Mean Sq F Value     Pr(F)
gas        3        40 13.33333       5 0.0451975
car        3        24  8.00000       3 0.1169598
driver     3       216 72.00000      27 0.0006987
Residuals  6        16  2.66667
```



Wow, look at those residuals!!! What's going on???

2 Split-Plot Designs

A split-plot design contains more than one source of error. This can arise because factors are applied at different scales, as in the `guayule` example below. Split-plots are also encountered because of restrictions on the randomization. For example, an experiment involving oven temperature and baking time will probably not randomize the oven temperature totally, but rather only change the temperature after all of the runs for that temperature have been made. This type of design is often mistakenly analyzed as if there were no restrictions on the randomization (an indication of this can be p-values that are close to 1). See Hicks (1973) and Daniel (1976); Neter, Wasserman and Kutner (1990, Chapters 26 and 28) have a nice introduction to split plot designs.

2.1 A Basic Split-Plot Design

S-PLUS includes the `guayule` data frame from Chambers and Hastie (1992), which is also analyzed in Chapter 14 of the SPLUS Guide to Statistics. The guayule is a Mexican plant from which rubber is manufactured. The `guayule` data frame has 96 rows and five columns:

y_{ijk}	`plants`:	numeric vector giving the number seeds out of a batch of 100 that germinated.
α_i	`variety`:	factor with levels V1 through V8 labeling the variety of guayule.
β_j	`treatment`:	factor with levels T1 through T4 labeling the treatment given to the seeds.
k	`reps`:	factor = 1, 2, or 3. All eight varieties of plant are represented once in each rep.
l	`flats`:	factor that labels which of 24 flats (3 of each variety) a batch was planted in.

Batches of 100 seeds of eight varieties (variety) of guayule were given one of four treatments (treatment), and planted; the number of plants that came up in each batch (plants) was recorded. Since a flat (a shallow box for starting seedlings) was not large enough to contain all 32 combinations of variety and treatment, the design was to use only a single variety in each flat and to apply each treatment within each flat. Thus the flats each consist of four sub-plots. This is a split-plot design since flats are the experimental unit for varieties, but the sub-plots are the experimental unit for the treatments. The response is the number of plants that germinated in each sub-plot.

If we ignored the flats, the model would be the two-way ANOVA model that we have considered many times before:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

where $i = 1, \ldots, 8$ indicates the variety, $j = 1, \ldots, 4$ indicates treatment, $k = 1, \ldots, 3$ indicates replications, and

$$\epsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$$

The data for this model would be simply be a two-way layout with balanced replications

		Treatment			
		T1	T2	T3	T4
V	V1	66 63 70	12 10 13	13 13 11	6 12 7
a	V2	77 47 66	26 11 18	27 5 11	15 4 15
r	V3	51 81 63	8 16 14	20 30 29	10 14 10
i	V4	52 40 59	4 15 11	19 16 7	13 9 7
e	V5	45 51 52	20 13 16	9 10 12	12 12 11
t	V6	59 66 49	8 8 8	28 32 29	14 21 16
y	V7	56 38 45	12 16 16	26 16 24	15 8 7
	V8	49 41 54	14 20 25	30 28 36	9 15 12

which says (since flats are not present in the layout) that the flats do not contribute anything to the responses; the same error term ϵ_{ijk} is associated with every observation in this table.

The essential idea of the split-plot approach is that the flats exert a common, random influence on outcomes y for each of the four subplots in the flat. The split-plot ANOVA model is thus

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} + \delta_l$$

where now

$$\begin{aligned} \epsilon_{ijk} &\overset{iid}{\sim} N(0, \sigma^2) \text{ and} \\ \delta_l &\overset{iid}{\sim} N(0, \tau^2), \text{ independent of the } \epsilon_{ijk}\text{'s.} \end{aligned}$$

The data table for this model is indicated below (what is the relationship between this and the table above for the additive model?). The error term ϵ_{ijk} is associated with each observation, as with the additive model. In addition, the error term ϵ_l is associated with every four observations below; thus,

- The error term for the first observation below is $\epsilon_{111} + \delta_1$;
- The error term for the second observation is $\epsilon_{121} + \delta_1$;

  $\vdots$

- The error term for the fifth observation (first one in the second flat) is $\epsilon_{112} + \delta_2$;

  $\vdots$

Note that *every* observation has variance $\sigma^2 + \tau^2$. Also built into this model is the fact that, if we don't condition on the flats, observations from the same flat will be dependent (since they share the same stochastic term ϵ_l), and observations from different flats will be independent.

```
402 > split(guayule,guayule$flats) # shows the four subplots in each flat

   variety treatment reps plants flats
 1      V1        T1    1     66  1.V1
 9      V1        T2    1     12  1.V1
17      V1        T3    1     13  1.V1
25      V1        T4    1      6  1.V1

   variety treatment reps plants flats
33      V1        T1    2     63  2.V1
41      V1        T2    2     10  2.V1
49      V1        T3    2     13  2.V1
57      V1        T4    2     12  2.V1

   variety treatment reps plants flats
65      V1        T1    3     70  3.V1
73      V1        T2    3     13  3.V1
81      V1        T3    3     11  3.V1
89      V1        T4    3      7  3.V1

   variety treatment reps plants flats
 2      V2        T1    1     77  1.V2
10      V2        T2    1     26  1.V2
18      V2        T3    1     27  1.V2
26      V2        T4    1     15  1.V2

   variety treatment reps plants flats
34      V2        T1    2     47  2.V2
42      V2        T2    2     11  2.V2
50      V2        T3    2      5  2.V2
58      V2        T4    2      4  2.V2

   variety treatment reps plants flats
66      V2        T1    3     66  3.V2
74      V2        T2    3     18  3.V2
82      V2        T3    3     11  3.V2
90      V2        T4    3     15  3.V2

   variety treatment reps plants flats
 3      V3        T1    1     51  1.V3
11      V3        T2    1      8  1.V3
19      V3        T3    1     20  1.V3
27      V3        T4    1     10  1.V3

   variety treatment reps plants flats
35      V3        T1    2     81  2.V3
43      V3        T2    2     16  2.V3
51      V3        T3    2     30  2.V3
59      V3        T4    2     14  2.V3

   variety treatment reps plants flats
67      V3        T1    3     63  3.V3
75      V3        T2    3     14  3.V3
83      V3        T3    3     29  3.V3
91      V3        T4    3     10  3.V3

   variety treatment reps plants flats
 4      V4        T1    1     52  1.V4
12      V4        T2    1      4  1.V4
20      V4        T3    1     19  1.V4
28      V4        T4    1     13  1.V4
```

```
   variety treatment reps plants flats
36      V4        T1    2     40  2.V4
44      V4        T2    2     15  2.V4
52      V4        T3    2     16  2.V4
60      V4        T4    2      9  2.V4

   variety treatment reps plants flats
68      V4        T1    3     59  3.V4
76      V4        T2    3     11  3.V4
84      V4        T3    3      7  3.V4
92      V4        T4    3      7  3.V4

   variety treatment reps plants flats
 5      V5        T1    1     45  1.V5
13      V5        T2    1     20  1.V5
21      V5        T3    1      9  1.V5
29      V5        T4    1     12  1.V5

   variety treatment reps plants flats
37      V5        T1    2     51  2.V5
45      V5        T2    2     13  2.V5
53      V5        T3    2     10  2.V5
61      V5        T4    2     12  2.V5

   variety treatment reps plants flats
69      V5        T1    3     52  3.V5
77      V5        T2    3     16  3.V5
85      V5        T3    3     12  3.V5
93      V5        T4    3     11  3.V5

   variety treatment reps plants flats
 6      V6        T1    1     59  1.V6
14      V6        T2    1      8  1.V6
22      V6        T3    1     28  1.V6
30      V6        T4    1     14  1.V6

   variety treatment reps plants flats
38      V6        T1    2     66  2.V6
46      V6        T2    2      8  2.V6
54      V6        T3    2     32  2.V6
62      V6        T4    2     21  2.V6

   variety treatment reps plants flats
70      V6        T1    3     49  3.V6
78      V6        T2    3      8  3.V6
86      V6        T3    3     29  3.V6
94      V6        T4    3     16  3.V6

   variety treatment reps plants flats
 7      V7        T1    1     56  1.V7
15      V7        T2    1     12  1.V7
23      V7        T3    1     26  1.V7
31      V7        T4    1     15  1.V7

   variety treatment reps plants flats
39      V7        T1    2     38  2.V7
47      V7        T2    2     16  2.V7
55      V7        T3    2     16  2.V7
63      V7        T4    2      8  2.V7

   variety treatment reps plants flats
71      V7        T1    3     45  3.V7
79      V7        T2    3     16  3.V7
87      V7        T3    3     24  3.V7
95      V7        T4    3      7  3.V7

   variety treatment reps plants flats
 8      V8        T1    1     49  1.V8
16      V8        T2    1     14  1.V8
24      V8        T3    1     30  1.V8
32      V8        T4    1      9  1.V8

   variety treatment reps plants flats
40      V8        T1    2     41  2.V8
48      V8        T2    2     20  2.V8
56      V8        T3    2     28  2.V8
64      V8        T4    2     15  2.V8

   variety treatment reps plants flats
72      V8        T1    3     54  3.V8
80      V8        T2    3     25  3.V8
88      V8        T3    3     36  3.V8
96      V8        T4    3     12  3.V8
```

To analyze a split-plot design like this, put the variable that corresponds to the whole plot in an *Error* term in the formula of the aov call:

```
402 > gua.aov1 <- aov(plants ~ variety*treatment + Error(flats), guayule)
```

The term `Error(flats)` says to insert the extra error terms δ_l associated with the flats.

As usual, you can get an ANOVA table with `summary()` (though `anova()` no longer works:

```
402 > summary(gua.aov1)
```

```
Error: flats
          Df Sum of Sq  Mean Sq  F Value     Pr(F)
  variety  7   763.156 109.0223 1.232036 0.3420697
Residuals 16  1415.833  88.4896

Error: Within
                  Df Sum of Sq  Mean Sq  F Value        Pr(F)
        treatment  3  30774.28 10258.09 423.4386 0.00000e+000
variety:treatment 21   2620.14   124.77   5.1502 1.32674e-006
        Residuals 48   1162.83    24.23
```

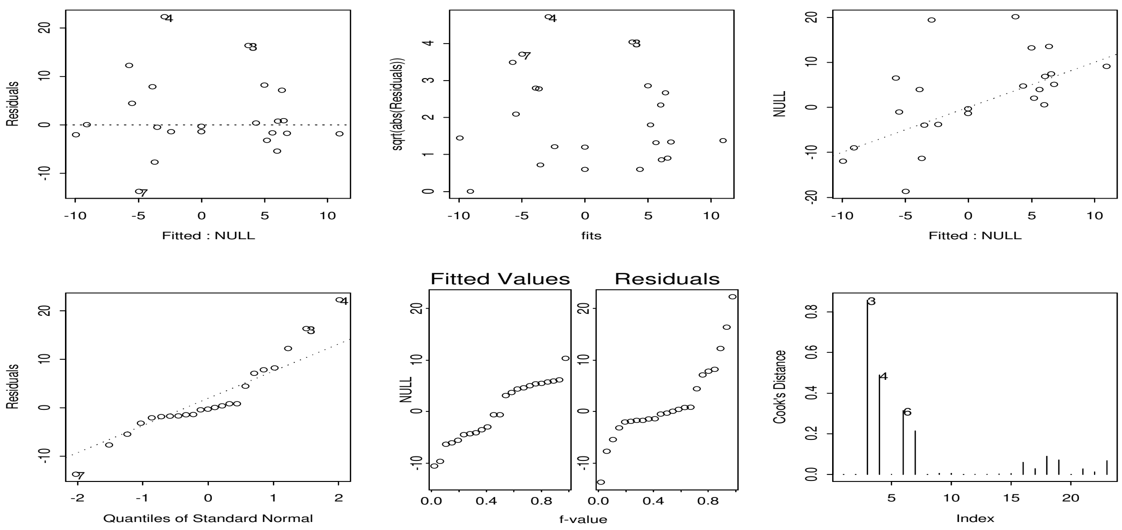
This shows varieties tested with the error from flats, while treatment and its interaction with variety are tested with the within-flat error (which is substantially smaller), analogous to a randomized-blocks design.

We can think of this as two ANOVA problems stacked on top of one another: a one-way ANOVA for the mean response of `variety` in each flat, and a two-way ANOVA for `variety`×`treatment` using the variability in the data that is unexplained by variety.

SPLUS will generate diagnostic plots of these two ANOVA problems, as follows:

```
402 > par(mfrow=c(2,3))
402 > gua.fm1 _ update(gua.aov1,qr=T)
402 > names(gua.fm1)
[1] "(Intercept)" "flats"       "Within"
402 > plot(gua.fm1$flats,title="Guayule Error(flats) flats")
```

The `update()` command above is not strictly necessary; it just makes the `plot()` command (and the `model.tables()` command below) go faster, by including some extra information in the fitted model object. See Venables and Ripley, p. 302).

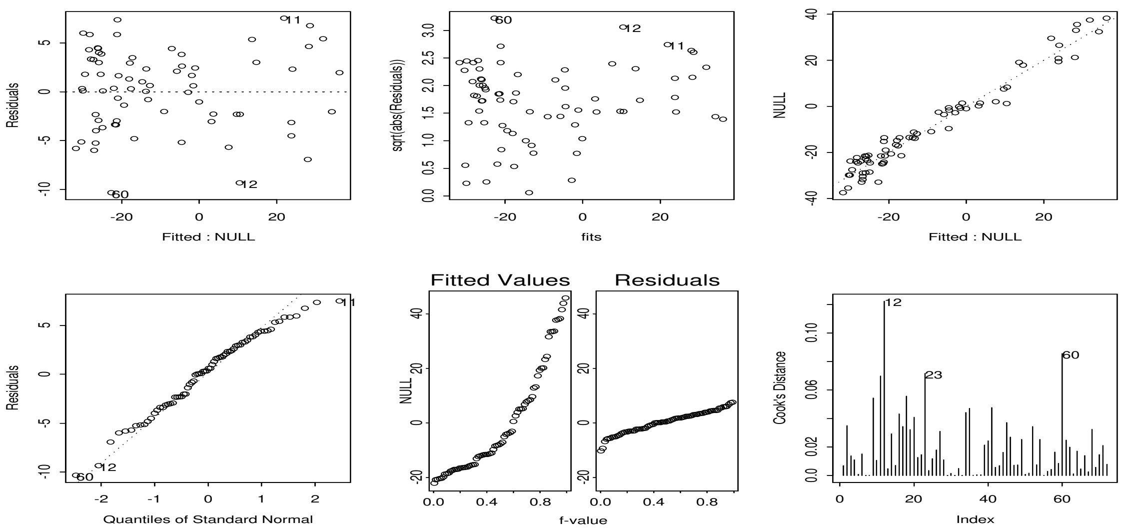


Note that there are 24 data points in these plots, corresponding to the 24 flats. Although residual diagnostic of this "second-level" model can be useful for suggesting general problems, they are much harder to interpret in terms of the original data: the "residuals" for example are something like $\bar{y}_{il} - \hat{y}_{il}$ where $\bar{y}_{il}$ is the mean number of germinating plants in each flat. I.e., we would draw much the same conclusions from the ANOVA of mean flat responses:

```
402 > tx.flat _ data.frame(tx=factor(rep(1:8,rep(3,8))),
+ meanresp=sapply(split(guayule[,c(1,4)],guayule$flats),apply,2,mean)[2,])
There were 24 warnings (use warnings() to see them)
402 > summary(aov(meanresp ~ tx,data=tx.flat))
          Df Sum of Sq  Mean Sq  F Value     Pr(F)
       tx  7  190.7891 27.25558 1.232036 0.3420697
Residuals 16  353.9583 22.12240
```

Here are the residual diagnostic plots for the "first-level" model.

```
402 > plot(gua.fm1$Within,title="Guayule Error(flats) within")
```



These plots are also helpful for suggesting problems with the model; however they are again difficult to interpret: although they are based on raw data (the "residual" really is $y_{ijkl} - \hat{y}_{ijkl}$), since `flat` information is missing, residuals within flats are dependent, so that visually assessing patterns and variability in the residuals is more difficult.

We can also obtain cell means for the two stacked ANOVA models

```
402 > model.tables(gua.fm1,type="means",se=T)

      ...

 variety
    V1    V2    V3 V4    V5    V6    V7    V8
 24.67 26.83 28.83 21 21.92 28.17 23.25 27.75

 treatment
```

```
    T1    T2    T3    T4
 55.83 13.92 20.04 11.42

 variety:treatment
Dim 1 : variety
Dim 2 : treatment
      T1    T2    T3    T4
V1 66.33 11.67 12.33  8.33
V2 63.33 18.33 14.33 11.33
V3 65.00 12.67 26.33 11.33
V4 50.33 10.00 14.00  9.67
V5 49.33 16.33 10.33 11.67
V6 58.00  8.00 29.67 17.00
V7 46.33 14.67 22.00 10.00
V8 48.00 19.67 31.33 12.00

Standard errors for differences of means

variety  3.84
replic. 12.00

treatment  1.421
  replic. 24.000

 variety:treatment
When comparing means with same levels of:
 variety otherwise
   4.019     5.183
replic. 3
```

We can now "eyeball" the significant contrasts, using two-sample t's—or just `(difference)`$\pm$`(two or three SE's)`, if we are in a hurry.

2.2 More Sources of Error: Split-Split-Plots

The guayule data actually represent an experiment in which the flats were grouped into replicates—i.e. all the flats for replication 1 were done at once; then all the flats for replication 2, then all the flats for replication 3.

This makes a total of three sources of error—"Within", "flats in reps", and "reps"—or a split-split-plot design. In math, the model is

$$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} + \delta_{lk} + \gamma_k$$

where

$$\begin{aligned} \epsilon_{ijk} &\overset{iid}{\sim} N(0,\sigma^2) \\ \delta_{lk} &\overset{iid}{\sim} N(0,\tau^2) \text{ and} \\ \gamma_k &\overset{iid}{\sim} N(0,\upsilon^2) \end{aligned}$$

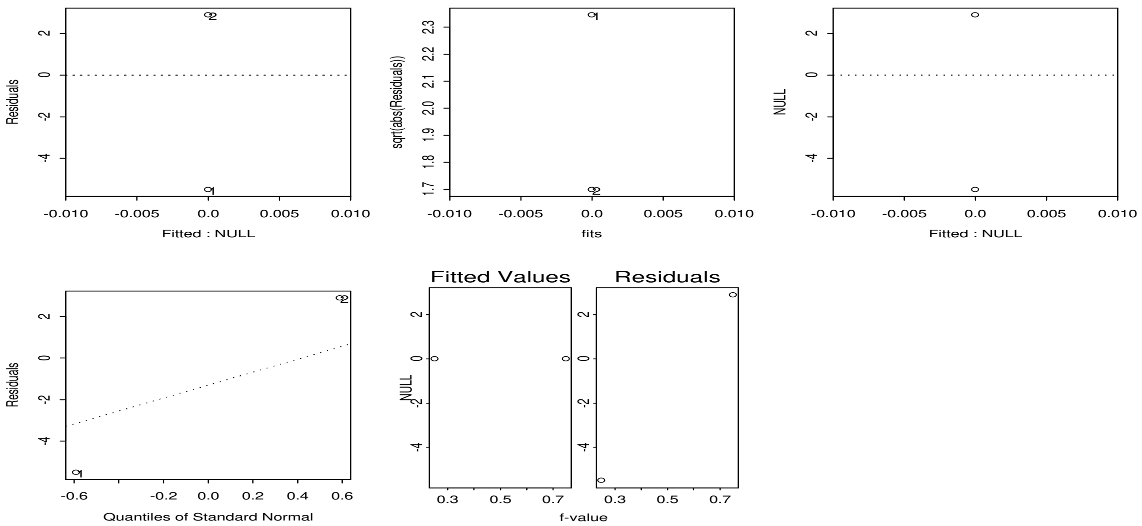
all independent of one another.

To model this we put more than one term inside the `Error` term—note the use of "`/`" to get the indexing of `flats` "within" `reps` right. See `help(aov)` for other uses of this "nesting" symbol.

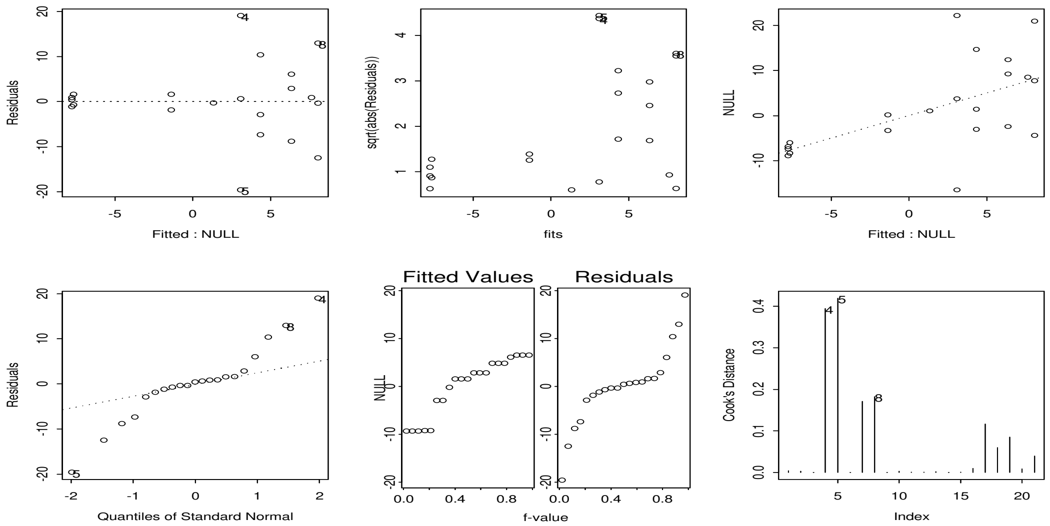
```
> gua.aov2 <- aov(plants ~ variety*treatment + Error(reps/flats), guayule)
> summary(gua.aov2)
Error: reps
          Df Sum of Sq  Mean Sq F Value Pr(F)
Residuals  2  38.58333 19.29167

Error: flats %in% reps
          Df Sum of Sq  Mean Sq  F Value     Pr(F)
  variety  7   763.156 109.0223 1.108232 0.4099625
Residuals 14  1377.250  98.3750

Error: Within
                  Df Sum of Sq  Mean Sq  F Value        Pr(F)
        treatment  3  30774.28 10258.09 423.4386 0.00000e+000
variety:treatment 21   2620.14   124.77   5.1502 1.32674e-006
        Residuals 48   1162.83    24.23
```

This can be interpreted as three ANOVA models stacked on top of each other. We can obtain diagnostic plots for each, with the same caveats on interpretability as before.

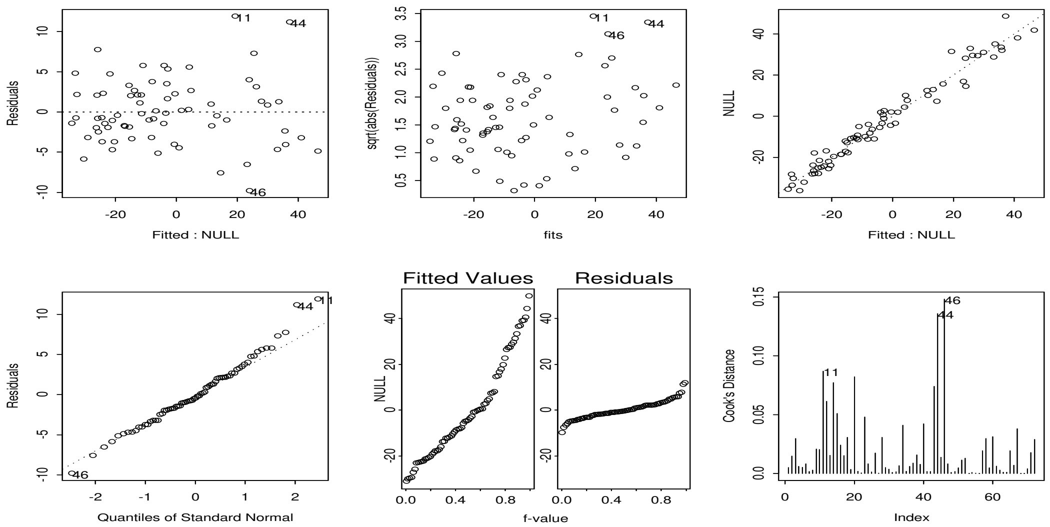
```
402 > gua.fm2 _ update(gua.aov2,qr=T)
402 > names(gua.fm2)
[1] "(Intercept)"     "reps"            "flats %in% reps" "Within"
402 > plot(gua.fm2$reps,title="Guayule Error(reps/flats) reps")
```




```
402 > plot(gua.fm2$"flats %in% reps",title="Guayule Error(reps/flats) flats")
```



`402 > plot(gua.fm2$Within,title="Guayule Error(reps/flats) Within")`



The plot for the `reps` level doesn't make much sense, since there are no predictors at that level. The plots for `"flats and for Within can be interpreted more or less as for the split plot design, and with the same caveats.`