Biostats VHM 802 Course
Winter 2012, Atlantic Veterinary College, PEI
Henrik Stryhn

Solution to Additional Exercise 9.2

The dataset consists of milk yields of 18 cows during 3 periods where the cows were subjected to different diets. The study was laid out in 6 Latin squares of size 3 × 3, such that for each set of 3 cows and their 3 periods the treatments formed a Latin square. The 6 Latin squares were of two types (determined by the order in which the treatments followed each other), with 3 replications of each type. In total, the data comprises 54 observations, and we use a single index notation where y_i denotes the (natural) logarithmic milk yield for observation i and $i = 1, \dots, 54$.

1. Additive statistical model

Our initial (basic) model is a simplification of the ANOVA model discussed in Example 13.12 because it omits the residuals effects. This model, with additive effects of cows, periods, treatments and squares, would usually be a reasonable starting point for analysis. The statistical model can be written

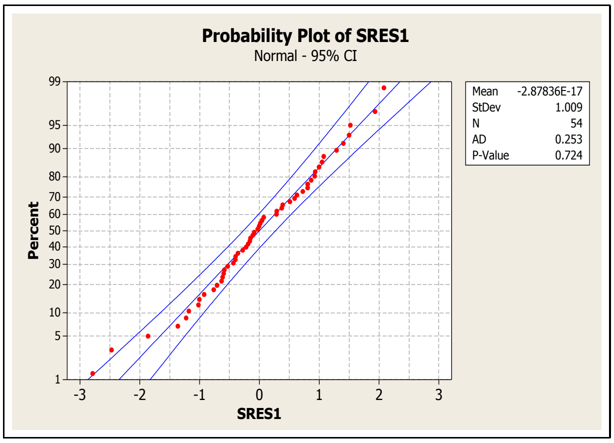
$$y_i = \mu + \alpha_{\mathrm{cow}(i)} + \beta_{\mathrm{per}(i)} + \gamma_{\mathrm{tx}(i)} + \varepsilon_i, \tag{1}$$

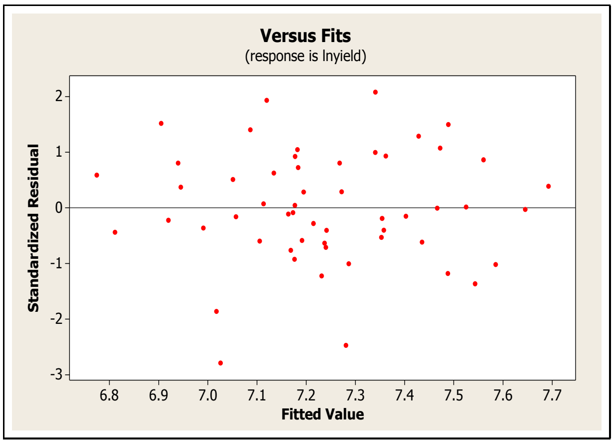
where the errors $\varepsilon_1, \dots, \varepsilon_{54}$ are assumed independent and identically, normally distributed $\mathrm{N}(0, \sigma^2)$. The ANOVA table is given below:

```
Source  DF   Seq SS   Adj SS   Adj MS       F      P
cow     17  0.90727  0.90727  0.05337   11.00  0.000
period   2  0.99807  0.99807  0.49903  102.86  0.000
tx       2  0.40999  0.40999  0.20500   42.25  0.000
Error   32  0.15525  0.15525  0.00485
Total   53  2.47058
```

The table shows all effects strongly significant, indicating that not only did the treatments have some effect (to be explored further), but the layout of the design in Latin squares seems to have been very effective in reducing the residual variation. The proportion of explained variation is very high ($R^2 = 94\%$).

Before proceeding with the analysis, we evaluate the model assumptions (plots on the next page). The normal plot of the residuals looks reasonably good despite the points off the line at the lower tail of the distribution. Normality tests are clearly non-significant (P-values above 0.5). The most extreme residual has a deletion residual of -3.156 which by the outlier test corresponds to $P = 0.19$, and thus offers no conclusive evidence that this observation is an outlier. The plot against fitted values looks fine, apart from the three points corresponding to the lowest residuals being separated from the rest of the points; there is no indication of heteroscedasticity problems.





Note that one would usually evaluate the "full" model and not a reduced model such as the present one. When adding terms to the model, the appearance of the residual plots can change appreciably, and therefore one should monitor the residuals for expanded models as well. The advantage of evaluating the full model is that one would rarely encounter major problems with reduced models if the full model was deemed acceptable.

2. Added period by square interactions

Even though the periods are the same in all squares, there is a choice between taking their effect as common to all squares or as separate effects for each square. The latter would be particularly relevant if the periods for the cows were only standardized within a square but not across squares. (For example, it may be difficult to find 18 homogeneous cows at the roughly same stage of lactation.) When we introduce square effects, the cows become "nested" within squares (because the cows are different in the different squares), and some software algorithms (e.g., Minitab) need to have cows explicitly specified as nested within squares. The expanded statistical model can be written

$$y_i = \mu + \alpha_{\text{cow}(i)} + \beta_{\text{per(sq)}(i)} + \gamma_{\text{tx}(i)} + \delta_{\text{sq}(i)} + \varepsilon_i. \tag{2}$$

The model does not have treatments nested within squares as well, because the squares are considered a blocking factor, and the experiment should have been carried out so that there was no reason to expect different treatment effects in the blocks. We get the ANOVA table:

```
Source          DF    Seq SS    Adj SS    Adj MS      F      P
square           5  0.623029  0.623029  0.124606  26.70  0.000
cow(square)     12  0.284241  0.284241  0.023687   5.08  0.001
period(square)  12  1.050659  1.050659  0.087555  18.76  0.000
tx               2  0.409995  0.409995  0.204997  43.93  0.000
Error           22  0.102658  0.102658  0.004666
Total           53  2.470583
```

The three somewhat extreme residuals from the previous model all have reduced substantially in size so that the model diagnostics for this model look very nice.

We can furthermore split the effect of periods within squares with 12 degrees of freedom (df) into an overall period effect with 2 df and the square by period interaction with 10 df, corresponding to the model equation below.

$$y_i = \mu + \alpha_{\text{cow}(i)} + \beta_{\text{per}(i)} + \gamma_{\text{tx}(i)} + \delta_{\text{sq}(i)} + (\beta\delta)_{\text{per}*\text{sq}(i)} + \varepsilon_i, \tag{3}$$

Note that this is merely a reparametrization, the model still has the same fit and residuals. The ANOVA table is:

```
Source          DF    Seq SS    Adj SS    Adj MS       F      P
square           5  0.623029  0.623029  0.124606   26.70  0.000
cow(square)     12  0.284241  0.284241  0.023687    5.08  0.001
period           2  0.998069  0.998069  0.499034  106.94  0.000
square*period   10  0.052591  0.052591  0.005259    1.13  0.387
tx               2  0.409995  0.409995  0.204997   43.93  0.000
Error           22  0.102658  0.102658  0.004666
Total           53  2.470583
```

The square by period interaction does not show any significance. With 10 df there may still be a potential for significant effects within the full interaction, so we continue our exploration of square type effects.

3. Added period by square type interaction

The ANOVA table from 2.) can be further expanded by splitting the variation between squares into the variation between the two square types and the variation between squares within square type. This is still the same model, but allows to test more effects.

$$y_i = \mu + \alpha_{\text{cow}(i)} + \beta_{\text{per}(i)} + \gamma_{\text{tx}(i)} + \delta_{\text{sqtype}(i)} + \eta_{\text{sq(i)}} + (\beta\delta)_{\text{per}*\text{sqtype}(i)} + (\beta\eta)_{\text{per}*\text{sq}(i)} + \varepsilon_i. \tag{4}$$

In the above formula, the term “square(sqtype)” has for simplicity been written as “square” (sq) . Note that this is also a reparametrization, so that the model still has the same fit and residuals. The corresponding ANOVA table is:

```
Source                  DF    Seq SS    Adj SS    Adj MS       F      P
sqtype                   1  0.028038  0.028038  0.028038    6.01  0.023
square(sqtype)           4  0.594992  0.594992  0.148748   31.88  0.000
cow(sqtype square)      12  0.284241  0.284241  0.023687    5.08  0.001
period                   2  0.998069  0.998069  0.499034  106.94  0.000
sqtype*period            2  0.007390  0.007390  0.003695    0.79  0.466
period*square(sqtype)    8  0.045201  0.045201  0.005650    1.21  0.338
tx                       2  0.409995  0.409995  0.204997   43.93  0.000
Error                   22  0.102658  0.102658  0.004666
Total                   53  2.470583
```

It is seen that also the interaction between square type and period is non-significant. The reason that this particular hypothesis might have been interesting is that the two types of squares differ in the order in which the treatments are applied and in the “pairs” of adjacent treatments (previous and current treatment). If there was a residual effect that depended on the pairs of adjacent treatments, there should be a difference between the square types. As the treatment pairs only can affect periods 2 and 3, the potential impact should appear in the interaction between square type and period. Apparently such an effect is just not present, and generally there does not seem to be any square-specific period effects.

4. Direct modelling of residual effects

The absence of a square type by period interaction does not mean that the data show no residual effects. It means that any residual effects do not seem to differ betweenn squares with different pairs. This would for example be the case if the residual effect was of a simple additive type, where each treatment exerts a certain residual effect in the next period independently of the next treatment. GO, Example 13.12, describes how such a residual effect may be modelled, using indicator (dummy) variables. These dummy variables are included in the dataset, labeled `r1` and `r2`. The results are shown in Listing 13.3. However, the residual plot in Figure 13.3 contains one large negative residual for observation no. 24, and the correponding deletion residual is -4.699 ($\sim P = 0.003$ by the outlier test). Analysis without this outlying observation gives similar results, as shown in the listing below.

```
Source  DF    Seq SS    Adj SS    Adj MS       F      P
cow     17  1.061368  0.922163  0.054245   22.81  0.000
period   2  0.820576  0.844883  0.422442  177.65  0.000
tx       2  0.337185  0.368317  0.184158   77.45  0.000
r1       1  0.042403  0.019671  0.019671    8.27  0.007
r2       1  0.006141  0.006141  0.006141    2.58  0.119
Error   29  0.068959  0.068959  0.002378
Total   52  2.336633

S = 0.0487637   R-Sq = 97.05%   R-Sq(adj) = 94.71%

Term          Coef   SE Coef        T      P
Constant   7.24469   0.00675  1073.08  0.000
tx
1         -0.11612   0.01086   -10.69  0.000
2          0.00342   0.01086     0.32  0.755
r1        -0.04058   0.01411    -2.88  0.007
r2        -0.02383   0.01483    -1.61  0.119
```

Taken together, the residual effects have increased in magnitude and significance, but the pattern shown by the estimates is still the same. There is a negative residual effect of in particular the first treatment but also to some extent the second treatment, whereas the third treatment has a positive residual effect (obtained by the zero-sum rule as 0.064). We don't have sufficient information about the dataset to assess whether the analysis with or without the extreme observation is preferable.