Chapter 8

Experimental design

1 Studies in crop variation

R. A. Fisher created a lot of statistical theory (which is still heavily used) while working at Rothamsted agricultural research station. In particular, Fisher developed a method of designing and analyzing complex experiemnts. (For more about the history see a presentation by Roger Payne.)

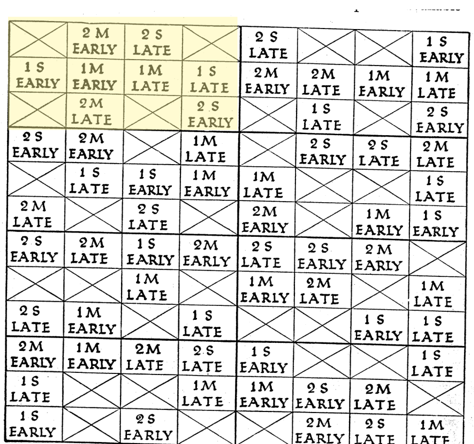
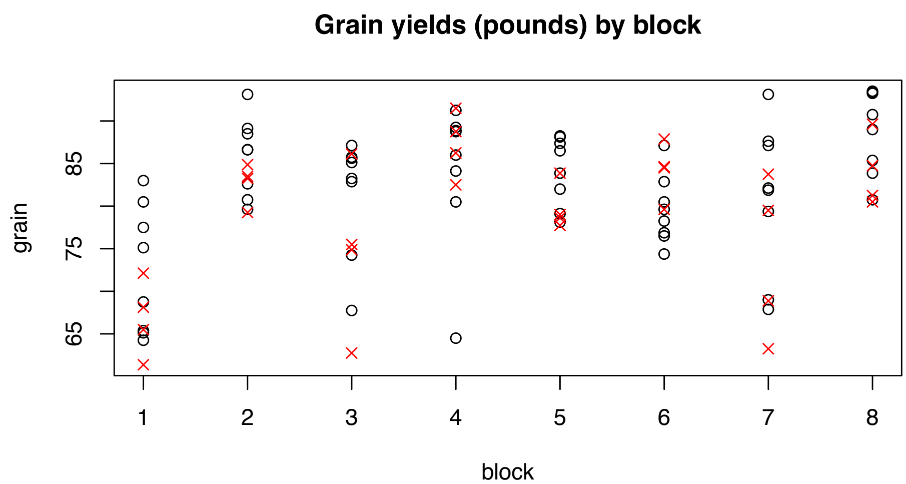


Fig. 1. A complex experiment with winter oats. (Reproduced from the *Journal of the Ministry of Agriculture* by permission of the Controller of H.M. Stationery Office.)

In a important early paper, Eden and Fisher (1927) described a way to compare the effects of various fertilizer treatments on the yield of grain (Grey Winter oats). They had two different nitrogen fertilizers (M = muriate of ammonia, S = sulphate of ammonia), applied in three different amounts (0, 1, or 2 cwt/acre), at two different stages of crop growth (E= early, L= late). They assigned the "treatments" to 96 plots of size 1/40 acre, arranged in 8 blocks of 12 plots each. Within each block, they assigned "treatments" to plots in a random order: 4 plots with no treatment (that is, amount = 0), and each of the eight possible combinations of $\{M, S\}$, $\{1, 2\}$, and $\{E, L\}$ appearing once.

The random allocation was intended to offset differences in fertility between different plots within each block, which were known to exist. (You can see these differences by looking at just the yields for the no-treatment plots, the red X's in the next plot.)




```
##            I      II    III      IV       V      VI     VII    VIII   total
## none   61.38   79.25  75.50   91.50   78.62   84.62   68.88   81.25  621.00
## none   65.50   83.50  74.88   86.25   79.00   84.50   79.50   80.50  633.62
## none   68.12   83.25  62.75   88.75   83.88   87.88   63.25   89.62  627.50
## none   72.12   84.88  86.12   82.50   77.75   79.62   83.75   84.75  651.50
## 1ME    77.50   80.75  85.12   80.50   88.25   76.88   69.00   90.75  648.75
## 1ML    80.50   93.12  67.75   88.88   88.12   79.62   67.88   80.75  646.62
## 1SE    65.38   89.12  85.75   86.00   86.50   76.50   79.38   93.50  662.12
## 1SL    75.12   86.62  85.62   89.25   87.38   87.12   87.62   93.25  692.00
## 2ME    83.00   86.62  83.25   64.50   82.00   82.88   82.12   85.38  649.75
## 2ML    64.25   79.62  87.12   88.75   79.12   74.38   87.12   89.00  649.38
## 2SE    68.75   88.50  82.88   84.12   83.88   78.25   81.88   83.88  652.12
## 2SL    65.12   82.62  74.25   91.25   78.12   80.50   93.12   93.38  658.38
## total 846.75 1017.88 951.00 1022.25 992.62 972.75 943.50 1046.00 7792.75
```

Remark. E&F Table I gave the grain yields in eighths of a pound:

```
##         I   II  III   IV    V   VI  VII VIII total
## none  491  634  604  732  629  677  551  650  4968
## none  524  668  599  690  632  676  636  644  5069
## none  545  666  502  710  671  703  506  717  5020
## none  577  679  689  660  622  637  670  678  5212
```

```
## 1ME    620  646  681  644  706  615  552  726  5190
## 1ML    644  745  542  711  705  637  543  646  5173
## 1SE    523  713  686  688  692  612  635  748  5297
## 1SL    601  693  685  714  699  697  701  746  5536
## 2ME    664  693  666  516  656  663  657  683  5198
## 2ML    514  637  697  710  633  595  697  712  5195
## 2SE    550  708  663  673  671  626  655  671  5217
## 2SL    521  661  594  730  625  644  745  747  5267
## total 6774 8143 7608 8178 7941 7782 7548 8368 62342
```

If you look at the paper, be aware that some tabulations are for pounds and some are for eighths of a pound.

The four untreated plots within each block give a way of estimating the variability within blocks:

```
## lm(formula = grain ~ -1 + block, data = EFdata, subset = notreat)
## Rsquared:  0.996
##            Estimate Std. Error t value Pr(>|t|)
## blockI       66.781      2.838  23.531        0
## blockII      82.719      2.838  29.147        0
## blockIII     74.812      2.838  26.361        0
## blockIV      87.250      2.838  30.743        0
## blockV       79.812      2.838  28.123        0
## blockVI      84.156      2.838  29.653        0
## blockVII     73.844      2.838  26.020        0
## blockVIII    84.031      2.838  29.609        0
## Estimate of sigma =  5.68 from  24 degrees of freedom
```

E&F also estimated σ using the residuals from an additive fit (not the way they put it):

```
out.bt <- lm(grain ~ -1 + block + treat,EFdata)
sighat <- sqrt(sum(out.bt$res^2)/out.treat$df)
round(sighat,3) # on 80 degrees of freedom

## [1] 6.405
```

Using an F-test, E&F decided that the two estimates of σ were not significantly different. They then declared that "the value derived from the whole 80 degrees of freedom may be used with confidence".

The analysis of variance table suggests that overall effect of the treatments is only at the noise level:

```
anova(out.treat)

## Analysis of Variance Table
##
## Response: grain
##           Df Sum Sq Mean Sq F value    Pr(>F)
## block      7 2286.4  326.63  7.9620 2.617e-07 ***
## treat      8  387.0   48.38  1.1792     0.322
## Residuals 80 3281.9   41.02
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

However, E&F also pointed out that the design of the experiment allows for more detailed comparisons. For example, there were 32 plots for each of the fertilizer treatments. The average yields provide a broad comparison of the three levels:

```
mean.fert <- tapply(EFdata$grain,EFdata$fert,mean)
```

The difference 2.19 between the means for the M and S fertilizers is down at the level of the (estimated) standard error for a difference of two such averages: $\widehat{\sigma}\sqrt{2/32} = 1.6$.

E&F commented that the only significant differences appeared to be in the amount of fertilizer used:

```
round(tapply(EFdata$grain,EFdata$amount,mean),2)

##     1     2     0
## 82.80 81.55 79.18
```

> **Remark.** These numbers are different from those in the first row of Table VI (E&F$_{560}$). My numbers are about 1.04 times bigger. Maybe that is the conversion factor for pounds to bushels, although I have my doubts. Probably I have made a silly mistake somewhere. The effects differences don't look very significant to me.

It is possible to carry out formal t-tests without so much manual labor. I'll return to this idea in a later section.

2 Decomposition of treatment effects

The E&F design involves a few complications that I'll avoid by first discussing a simpler data set from Box et al. (1978, Section 10.1). The data involve three factors: temperature (at 160 or 180 degrees Celsius), concentration (at 20% or 40%), and catalyst (A or B), with the yield (in grams) as the response in a pilot study. (BHH devote quite a few pages to the example.)

```
##   temp conc catal yield
## 1  180   40     B    60
## 2  160   40     B    72
## 3  180   20     B    54
## 4  160   20     B    68
## 5  180   40     A    52
## 6  160   40     A    83
## 7  180   20     A    45
## 8  160   20     A    80
```

I will explain what is going on with things like:

```
##
## Call:
## lm(formula = yield ~ temp + conc + catal, data = bhh)
##
## Residuals:
##    1    2    3    4    5    6    7    8
##  5.5 -5.5  4.5 -4.5 -4.0  4.0 -6.0  6.0
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   64.250      2.531  25.385 1.43e-05 ***
## temp1        -11.500      2.531  -4.544   0.0105 *
## conc1          2.500      2.531   0.988   0.3792
## catal1        -0.750      2.531  -0.296   0.7817
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.159 on 4 degrees of freedom
## Multiple R-squared:  0.8444,Adjusted R-squared:  0.7277
## F-statistic: 7.236 on 3 and 4 DF,  p-value: 0.04297
```

```
## Analysis of Variance Table
##
## Response: yield
##           Df Sum Sq Mean Sq F value  Pr(>F)
## temp       1 1058.0 1058.00 20.6439 0.01047 *
## conc       1   50.0   50.00  0.9756 0.37920
## catal      1    4.5    4.50  0.0878 0.78173
## Residuals  4  205.0   51.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Call:
## lm(formula = yield ~ temp * conc * catal, data = bhh)
##
```

```
## Residuals:
## ALL 8 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         6.425e+01         NA      NA       NA
## temp1              -1.150e+01         NA      NA       NA
## conc1               2.500e+00         NA      NA       NA
## catal1             -7.500e-01         NA      NA       NA
## temp1:conc1         7.500e-01         NA      NA       NA
## temp1:catal1        5.000e+00         NA      NA       NA
## conc1:catal1       -1.963e-15         NA      NA       NA
## temp1:conc1:catal1 -2.500e-01         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,Adjusted R-squared:    NaN
## F-statistic:   NaN on 7 and 0 DF,  p-value: NA
```

```
## Warning in anova.lm(lm(yield ~ temp * conc * catal, bhh)):  ANOVA F-tests on an essentially perfect fit are unreliable
```

```
## Analysis of Variance Table
##
## Response: yield
##                 Df Sum Sq Mean Sq F value Pr(>F)
## temp             1 1058.0  1058.0
## conc             1   50.0    50.0
## catal            1    4.5     4.5
## temp:conc        1    4.5     4.5
## temp:catal       1  200.0   200.0
## conc:catal       1    0.0     0.0
## temp:conc:catal  1    0.5     0.5
## Residuals        0    0.0
```

To be continued.

References

Box, G. E. P., W. G. Hunter, and J. S. Hunter (1978). *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.* New York: Wiley.

Eden, T. and R. A. Fisher (1927, 10). Studies in crop variation IV: The experimental determination of the value of top dressings with cereals. *The Journal of Agricultural Science 17*(4), 548–562.