Solution to Exercise 6.3 (Version 1, 26/10/14)

from Statistical Methods in Biology: Design & Analysis of Experiments and Regression (2014) S.J. Welham, S.A. Gezan, S.J. Clark & A. Mead. Chapman & Hall/CRC Press, Boca Raton, Florida. ISBN: 978-1-4398-0878-8

Exercise 6.3

A field experiment was done to investigate the effects of amount and timing of sulphur application on the level of scab disease in potatoes (Cochran & Cox, 1957, Table 4.1). Three doses of sulphur were used (300, 600 and 1200 lb per acre) and these were applied in either spring or autumn. Plots with no sulphur application were included as controls, giving seven treatments in total. The control treatment was replicated eight times and the six sulphur treatments each four times in a CRD with 32 plots in a four-row × eight-column layout. The average percentage surface area with scab for 100 potatoes per plot is the response to be analysed. The unit numbers (*Plot*), treatments applied (factor Treatment) and responses (variate *Scab*) can be found in file SCAB.DAT. Analyse these data on an appropriate scale using one-way ANOVA to compare the seven treatments. Is there any evidence that the application of sulphur affects the incidence of scab? (We re-visit these data in Exercises 8.5 and 11.4.)

Data 6.3 (SCAB.DAT)

Average % surface area with scab for 100 potatoes per plot (Scab) from a field experiment to investigate the effects of amount and timing of sulphur application on the level of scab disease. Plots were arranged in a 4-row × 8-column layout. Treatments are defined in Table S6.3.1.

Plot	Row	Col	Treatment	Scab	Plot	Row	Col	Treatment	Scab
1	1	1	2	9	17	3	1	2	9
2	1	2	1	12	18	3	2	7	7
3	1	3	6	18	19	3	3	3	18
4	1	4	4	10	20	3	4	1	30
5	1	5	6	24	21	3	5	3	18
6	1	6	7	17	22	3	6	7	16
7	1	7	5	30	23	3	7	2	16
8	1	8	3	16	24	3	8	4	4
9	2	1	1	10	25	4	1	5	9
10	2	2	5	7	26	4	2	1	18
11	2	3	4	4	27	4	3	7	17
12	2	4	3	10	28	4	4	6	19
13	2	5	5	21	29	4	5	1	32
14	2	6	1	24	30	4	6	4	5
15	2	7	1	29	31	4	7	1	26
16	2	8	6	12	32	4	8	2	4

Source: Cochran, W.G. & Cox, G.M. (1957) Experimental designs (2nd Edition). J. Wiley & Sons, New York.

Table S6.3.1 Definition of treatments.

Treatment number	Amount of sulphur, lb/acre	Time of application
1	0	–
2	300	autumn
3	600	autumn
4	1200	autumn
5	300	spring
6	600	spring
7	1200	spring

Solution 6.3

We first write a single factor model for the raw data, with symbolic form

Response variable: *Scab*
Explanatory component: [1] + Treatment

Analysis of this model gives the residual plots in Figure S6.3.1. Whilst the histogram and normal plot are acceptable, the fitted values and absolute residuals plots clearly indicate variance heterogeneity, with more variation among residuals corresponding to larger fitted values. This might have been expected, as the index values are percentages in the range 4-32% and percentage data often show variance heterogeneity, with maximum variance at 50% and minimum variance at 0 and 100%. We therefore transform the data to the logit scale (which is often used for percentage data) in an attempt to stabilise the variance, as *LogitScab* = log(Scab/(100-Scab)). The residual plots from analysis of the logit-transformed data, using the same model, are shown in Figure S6.3.2. The fitted values and absolute residuals plots are much improved, although the histogram and normal plots are slightly worse. However, the histogram and normal plots are acceptable, so we continue with the analysis on the logit scale.

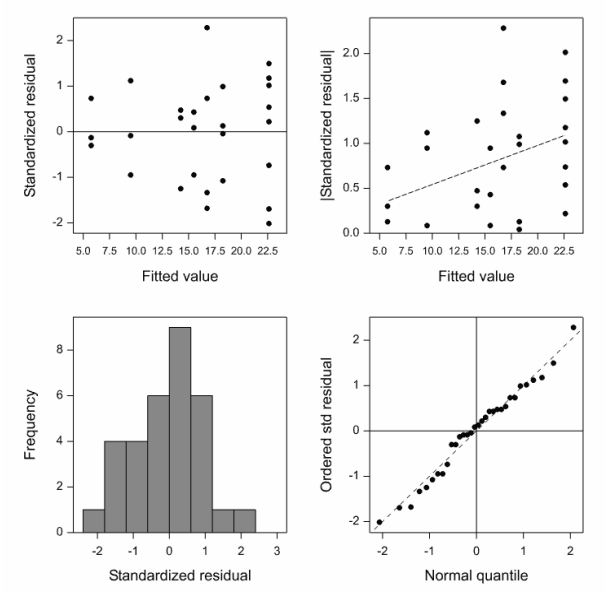

Figure S6.3.1. Composite set of residual plots based on standardized (std) residuals obtained from analysis of the observed index values (percentages).

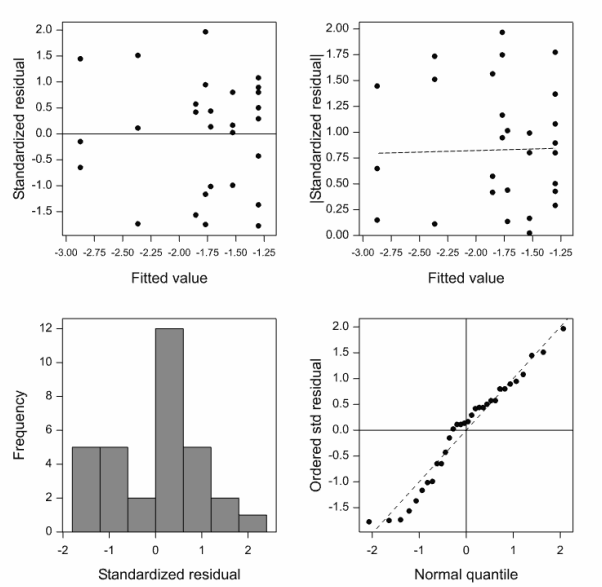

Figure S6.3.2. Composite set of residual plots based on standardized (std) residuals obtained from analysis of the logit transformed index values.

The ANOVA table for the logit-transformed scab index is shown in Table S6.3.2. The observed variance ratio ($F_{6,25} = 4.667$) is larger than the 1% critical value of the F-distribution ($F_{6,25}^{[0.01]} = 3.627$), with observed significance level $P = 0.003$. We therefore reject the null hypothesis and conclude that there is evidence that the application of sulphur affects the population mean incidence of scab (logit-transformed). The predicted treatment population means are given in Table S6.3.3. Least control is achieved when no sulphur is applied. The best control is achieved with the highest dose applied in the autumn. Further investigation of the specific effects of time of application and dose are best achieved by exploiting the factorial nature of the treatments, and this is the topic of Exercise 8.5.

Table S6.3.2 ANOVA table for one-way ANOVA of logit-transformed scab index (variate *Scab*).

Source of variation	df	Sum of squares	Mean square	Variance ratio	*P*-value
Treatment	6	8.1849	1.3642	4.667	0.003
Residual	25	7.3067	0.2923		
Total	31	15.4916			

Table S6.3.3 Predicted mean logit-transformed scab index with back-transformed values. Logit scale: SED for control vs other treatments = 0.3311, otherwise 0.3823 (df = 25); LSD for control vs other treatments = 0.6818, otherwise 0.7873 ($\alpha_s = 0.05$, df = 25). A = autumn, S = spring.

	Treatment						
	1 (control)	2	3	4	5	6	7
Amount of sulphur, lb/acre	0	300	600	1200	300	600	1200
Time of application	–	A	A	A	S	S	S
Mean (logit scale)	−1.300	−2.366	−1.722	−2.874	−1.768	−1.528	−1.854
Back-transformed mean (%)	21.41	8.58	15.16	5.34	14.58	17.83	13.54