Assignment 2

Stat 623

Due on Oct 8, 2020

Turn in problems 1–2. Problem 3 is for extra credit.

1. The file corn.txt shows the effect of nitrogen fertilizer on the yield of corn. The experiment was arranged in three blocks or contiguous areas, most likely three parts of the same field, but possibly three different fields. The first column gives the amount of fertilizer used in lbs/acre, the second column gives the block number, and the third column gives the observed yield in lbs per plot.

   (a) Calculate the sum of squares associated with variation between blocks, the sum of squares associated with nitrogen, and the sum of squares associated with departures from additivity (or interactions)

   (b) The purpose of blocking is to eliminate recognizable variation between plots, due for example to different soil composition or differences of exposure to sun or rain. Has blocking been effective in this experiment?

   (c) The response to nitrogen appears to be smooth but non-linear. Illustrate this by a suitable plot. Fit the linear model **block + beta log (gamma + nitrogen)** for various values of γ in the range $1-30$. Plot the residual sum of squares against γ and find the least squares estimate.

   (d) Can the model in (c) be substantially improved by response transformation? Check this using the Box-Cox method.

   (e) Give a brief summary of your conclusions.

2. The data taken from Snedecor and Cochran (1967, p. 354) were obtained as part of an experiment to determine the effects of temperature and storage time on the loss of ascorbic acid in snap-beans. The beans were all harvested under uniform conditions at the Iowa Agricultural Station before eight o'clock morning. They were prepared and quick-frozen before noon the same day. Three packages were assigned at random to each temperature and storage-lime combination. The some of the three ascorbic acid determinations is shown in the table

Table 1: Ascorbic acid concentration of samples of snap-beans after a period of cold storage

	Weeks of storage				
Temp. °F	2	4	6	8	*Total*
0	45	47	46	46	184
10	45	43	41	37	166
20	34	28	21	16	99
Total	124	118	108	99	449

Suppose for the purpose of model construction that ascorbic acid concentration decays exponentially fast, with a decay rate that is temperature dependent. In order words, for a given storage temperature T, the expected concentration after time t (measured in weeks)

$$\mu = E(Y) = e^{-\alpha - \beta_T t}.$$

The initial concentration, e^{α} is assumed in this model to be independent of the storage temperature.

(a) Express the above theory as a regression model, treating temperature as a factor and storage time as a variate. Write down the details of the model and the score equations. Fit the model to the data using the glm function with family=gaussian and link="log" and obtain the estimates of the parameters and their standard errors. Provide summary from the fitted model.

(b) Estimate the times taken at each of the three temperature for the ascorbic acid combination to be reduced to 50% of the original value. Consider carefully how you might construct confidence intervals for this half-life.

3. The following provides excerpts of R-output from two linear regressions. In one case, sales performance (y) is regressed on education (x_1) and extroversion scores (x_2), while in the other sales performance (y) is regressed only on education scores (x_1).

```
lm(formula = y ~ x1 + x2)
```

```
Coefficients:
                 Estimate Std. Error    t value     Pr(>|t|)
(Intercept)       99.39          78.80      1.2      0.27
x1                8.00            7.00       1.1     0.20
x2                52.10          20.40       2.5     0.02 *


lm(formula = y ~ x1)

Coefficients:
                  Estimate  Std. Error.    t value    Pr(>|t|)
(Intercept)     170.00         80.45            2.11     0.04 *
x1                10.5          7.09            1.49     0.16
```

What can we say about the relationship between education (x_1) and extroversion (x_2) scores of these sales people? What assumptions will you make? Explain your answer with reasoning.