Theory of Linear Models
Exercises 3 Solutions

3 February 2020

1. Consider the half-replicate fractional factorial design for three factors, each at two levels, plus two centre points:

X_1	X_2	X_3
-1	-1	-1
-1	1	1
1	-1	1
1	1	-1
0	0	0
0	0	0

Find a 95% confidence interval for $\beta_1 + \beta_{23}$ if $\widehat{\beta_1 + \beta_{23}} = 10.03$ and $s^2 = 5.34$. How would the result be interpreted?

The question should have made clear that the full second order model should be assumed, so that there is only one residual degree of freedom. We use the result at the top of page 2 in the notes for Lecture 4, with $\mathbf{c}' = [0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1]$, $n = 6$, $r = 5$, $\alpha = 0.05$, $s^2 = 5.34$ and

$$\mathbf{X} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

This gives $10.03 \pm 12.7062\sqrt{5.34 \times 0.25}$, i.e. $(-4.65, 24.71)$. This can be interpreted as an interval estimate for the effect of factor X_1, *under the assumption of no interaction between X_2 and X_3.*

2. Consider the following data:

Y	X_1	X_2	X_3
10	-1	-1	-1.0001
12	-1	1	-0.9999
20	1	-1	1.0000
22	1	1	1.0000
16	0	0	0.0000
18	0	0	0.0000

(a) Test the hypothesis $\beta_1 = \beta_2 = 0$ against a two-sided alternative.

This time the question should have specified the first order model and, in this part, that a normal distribution should be assumed. This, then, is a regular F test. The test statistic is

$$\frac{(7.331 - 3.333)/2}{3.333/2} = 1.200,$$

which gives a p-value of 0.455, so there is little evidence against H_0.

(b) Test the hypothesis $\beta_1 = \beta_3 = 0$ against a two-sided alternative.

The test statistic is

$$\frac{(103.333 - 3.333)/2}{3.333/2} = 30.03,$$

which gives a p-value of 0.032, so there is some evidence against H_0. Note that these results are despite the fact that, in the full model $\hat{\beta}_1 > \hat{\beta}_2 > \hat{\beta}_3$.

(c) Calculate the leverage of each observation.

The leverages are 0.9167, 0.9167 0.9167 0.9167 0.1667 and 0.1667.

(d) Assuming that the experiment was completely randomized, carry out a permutation test of $\beta_1 = 0$ against a two-sided alternative.

We have to consider $6! = 720$ permutations of the data and calculate $|\hat{\beta}_1|$ for each. I found that 644 of these have $|\hat{\beta}_1| \geq 5$, which is the observed value, so the p-value is $644/720 = 0.894$.

(e) Comment on the results.

It is difficult to believe any of them, given the high colinearity between X_1 and X_3.

3. For the general second order polynomial regression model, with q explanatory variables, find the location of the stationary point as a function of the parameters. (Hint: rewrite the model using the vector $\mathbf{b}' = [\beta_1 \ \cdots \ \beta_q]$ and the matrix

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \frac{1}{2}\beta_{12} & \cdots & \frac{1}{2}\beta_{1q} \\ \frac{1}{2}\beta_{12} & \beta_{22} & & \vdots \\ \vdots & & \ddots & \frac{1}{2}\beta_{(q-1)q} \\ \frac{1}{2}\beta_{1q} & \cdots & \frac{1}{2}\beta_{(q-1)q} & \beta_{qq} \end{bmatrix}$$

and then use vector differentiation.)

The model can be written

$$\begin{aligned} \mu &= \beta_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} \\ \Rightarrow \frac{\partial \mu}{\partial \mathbf{x}} &= \mathbf{b} + 2\mathbf{B}\mathbf{x}. \end{aligned}$$

Equating this to zero, we find that the stationary point is at

$$\mathbf{x} = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b},$$

as long as $\mathbf{B}$ is non-singular (which it will be for any real data).