

Data Analysis Exercises for Chapter 17 *Applied Regression Analysis, Generalized Linear Models, and Related Methods, Third Edition* (Sage, 2016)

John Fox

Last modified: 2015-01-29

Exercise D17.1 The data in `ginzberg.txt` (collected by Ginzberg) were analyzed by Monette (1990). The data are for a group of 82 psychiatric patients hospitalized for depression. The response variable in the data set is the patient’s score on the Beck scale, a widely used measure of depression. The explanatory variables are “simplicity” (measuring the degree to which the patient “sees the world in black and white”) and “fatalism.” (These three variables have been adjusted for other explanatory variables that can influence depression.) Using the full quadratic regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2 + \beta_4 X_2^2 + \beta_5 X_1 X_2 + \varepsilon$$

regress the Beck-scale scores on simplicity and fatalism. Are the quadratic and product terms needed here? If you have access to suitable software, graph the data and the fitted regression surface in three dimensions. Do you see any problems with the data? What do standard regression diagnostics for influential observations show?

Exercise D17.2 The following table gives the population of Canada for decennial censuses between 1851 and 2001.

| <i>Year</i> | <i>Population</i> | <i>Year</i> | <i>Population</i> |
|-------------|-------------------|-------------|-------------------|
| 1851 | 2.436 | 1931 | 10.377 |
| 1861 | 3.230 | 1941 | 11.507 |
| 1871 | 3.689 | 1951 | 13.648 |
| 1881 | 4.325 | 1961 | 17.780 |
| 1891 | 4.833 | 1971 | 21.046 |
| 1901 | 5.371 | 1981 | 23.774 |
| 1911 | 7.207 | 1991 | 26.429 |
| 1921 | 8.788 | 2001 | 30.007 |

Fit the logistic growth model to these data, assuming (a) additive errors (Equation 17.8 in the text, reproduced here),

$$Y_i = \frac{\beta_1}{1 + e^{\beta_2 + \beta_3 X_i}} + \varepsilon_i$$

and (b) multiplicative errors (as in Exercise 17.9 in the text, also reproduced here)

$$Y_i = \frac{\beta_1}{1 + e^{\beta_2 + \beta_3 X_i}} \varepsilon_i$$

Compare the results of the two fits. Be sure to graph the data and the fit of the models; also graph the residuals from the models and examine their autocorrelation.