Fall, 2007

Applied Statistics Comprehensive Examination

Regression and Linear Models

Answers to all questions require complete explanations to receive full credit.

(20) 1. Consider the multiple linear regression model, $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$ where $\epsilon \sim iid$ Normal$(0, \sigma^2)$. Using the procedure for testing a general linear hypothesis, show how to test the following hypotheses:

(a) $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4$

(b) $H_0 : \beta_1 = 2\beta_2$ and $\beta_3 = -\beta_4$

(20) 2. Discuss what is meant by the following regression terms: (a) outlier, (b) leverage and (c) influence. Your discussion should include a description of how they are measured and why they are important in regression analysis.

(10) 3. Assuming a simple linear regression model, construct a sample $x - y$ scatterplot to illustrate each of the following two situations: (1) a dataset containing an observation with a low residual, high leverage and high influence and (2) a dataset containing an observation with a high residual, low leverage and low influence. You should make one plot for each situation and in each plot, be sure to clearly identify the observation with the desired characteristics.

(30) 4. Consider an experiment with a two-way treatment structure having three rows, two columns and four observations per treatment combination which is conducted in a completely randomized manner. (a) Specify the population marginal means for row 1 assuming a cell means model. (b) Specify the population marginal means for column 2 assuming a fixed effects model with interaction (use α_i for row effects, β_j for column effects and γ_{ij} for interaction effects). (c) Specify two orthogonal contrasts for interaction effects assuming an effects model with interaction.

(20) 5. An experiment was conducted to investigate the warping of copper plates. The two factors studied were temperature (50° and 100°) and copper content of the plates (40% and 80%) and the response variable was the amount of warping. Ten observations were obtained at each of the four combinations of the two factors and the resulting averages were recorded:

	Content (%)	
Temp (°)	40	80
50	2	4
100	10	8

Construct both interaction plots for these data and interpret your results. Describe the impact of these results on the testing of main effect.