Statistics 850 **yandell** **Spring 1995**

Homework # 3—Due Mon 20 Feb

Gray cast iron is a common commercial iron-based casting alloy. The two principal alloying elements are carbon and silicon. This problem examines the relationship between carbon level (`C`) and a particular melting property (`resp`), specifically called the "primary austenite liquidus arrest temperature of the cooling curve". Three levels of added carbon were used, low (1), medium (2) and high (3). However, it is not possible to set these levels exactly, for a variety of practical reasons. Instead, the scientist can measure the actual carbon afterward (`Cactual`). The questions for this homework concern how to relate carbon level to the response. Is it enough to just consider the 3 discrete levels (coded as `Clevel=1,2,3`) or is there further information about the response in the actual carbon? Use plots, models and tables of means and/or analyses of variance to present your arguments.

1. Run a one-factor analysis of variance comparing the effect of the three levels of carbon on response. Include multiple comparison of means. In particular, consider whether the linear contrast of means $(1\mu_3 + 0\mu_2 - 1\mu_1)$ is zero.
2. Regress the discrete carbon levels on response. Why does the variation explained differ from the contrast considered in the previous problem?
3. Test lack of linear fit. That is, is there evidence that the relationship between carbon level and respose could be nonlinear?
4. Regress the actual carbon on response. Why is the fit worse than in problem 2, even though more "accurate" measure of carbon is used? [You may want to use plots to show this.]
5. Fit a model with both discrete level and actual carbon to examine the question of error in variables. How strong is the evidence?

The data can be found in file `hwk3.dat` (in the usual directory, `/p/stat/course/st850-yandell/data`), with some SAS and S suggestions for this problem in `hwk3.sas` and `hwk3.s`, respectively, for your information. [The last SAS `proc glm` (or S `ervar2`) checks for a nonlinear error in variables, finding no significant evidence. This is for your information.]