Homework #7, ST518

1. Consider the unbalanced version of the cholesterol dataset available on moodle as "chol-hw7.dat".

   (a) Specify a statistical model for the cholesterol measurements in which the effects of age and gender are additive.

   (b) Report the four marginal means (mean for men, mean for women, mean for youngsters, mean for older folks). What function of factorial effects are estimated by each of these four marginal means. (Derive the expected value of each under the model.)

   (c) Using statistical software, estimate the four cell means for younger and older men and younger and older women. Report standard errors and 95% confidence intervals.

2. (Not graded) A food scientist runs a crossed, two-factor experiment in a completely randomized design. The levels of the "additive" factor are acetic, nothing, sorbate and the levels of the uv factor are 0,1. For each contrast specified in quotes in the ESTIMATE statements below, provide the correct SAS code to estimate it. (The given coefficients have many errors.) Alternatively, express the contrast estimator as a linear combination of treatment means, $\sum \sum c_{ij} \overline{Y}_{ij}$. and give the expectation of each contrast.

```
proc glm;
   class additive uv;
   model  ycount=additive uv uv*additive;
estimate 'acetic uv=0 vs acetic uv=1' uv  1 -1 uv*additive 1 -1 0 0 0 0;
estimate 'acetic uv=0 vs nothing uv=0' uv 1 -1 uv*additive 1 0 -1 0 0 0;
estimate 'acetic uv=0 vs nothing uv=1' uv 1 -1 uv*additive 1 0 0 -1 0 0;
estimate 'acetic uv=0 vs sorbate uv=0' uv 1 -1 uv*additive 1 0 0 0 -1 0;
estimate 'acetic uv=0 vs sorbate uv=1' uv 1 -1 uv*additive 1 0 0 0 0 -1;
estimate 'acetic uv=1 vs nothing uv=0' uv 1 -1 uv*additive 0 1 -1 0 0 0;
estimate 'acetic uv=1 vs nothing uv=1' uv 1 -1 uv*additive 0 1 0 -1 0 0;
estimate 'acetic uv=1 vs sorbate uv=0' uv 1 -1 uv*additive 0 1 0 0 -1 0;
estimate 'acetic uv=1 vs sorbate uv=1' uv 1 -1 uv*additive 0 1 0 0 0 -1;
estimate 'nothing uv=0 vs nothing uv=1' uv 1 -1 uv*additive 0 0 1 -1 0 0;
estimate 'nothing uv=0 vs sorbate uv=0' uv 1 -1 uv*additive 0 0 1 0 -1 0;
estimate 'nothing uv=0 vs sorbate uv=1' uv 1 -1 uv*additive 0 0 1 0 0 -1;
estimate 'nothing uv=1 vs sorbate uv=0' uv 1 -1 uv*additive 0 0 0 1 -1 0;
estimate 'nothing uv=1 vs sorbate uv=1' uv 1 -1 uv*additive 0 0 0 1 0 -1;
estimate 'sorbate uv=0 vs sorbate uv=1'uv 1 -1 uv*additive 0 0 0 0 1 -1;
estimate 'uv=0 vs uv=1' uv 1 -1;
estimate 'acetic vs nothing' additive 1 -1;
estimate 'acetic vs sorbate' additive 1 0 -1;
estimate 'nothing vs sorbate' additive 0 1 -1;
```

3. Find the data file "tilley.csv", which contains yields of corn from what can be regarded as a 4×2 completely randomized design, with four levels of planting density grown at two locations. (This problem derived from experiments run by Scott Tilley and is presented with permission.)

   (a) Propose a factorial effects model which will enable you to investigate the effects of density and location on yield and the possibility that density effects vary across yield.

   (b) Obtain an ANOVA table from your model.

   (c) Report the p-value for a test for interaction between density and location.

   (d) Test for a main effect of density at level $\alpha = .05$. Does this mean there is no effect of density on yield?

   (e) Construct an interaction plot.

   (f) Characterize the effects of the two factors, using p-values and possibly sums of squares to support your characterization. Is the effect of density positive or negative? Linear? Non-linear?