Stat 705, Spring 2015: Homework 3

RCB design, Fat in diets: Consider the data of problem 21.7 (p. 913). Note that the outcome Y_{ij} is the *reduction* in lipids after being on the diet.

1. Obtain the interaction plot for these data by typing `plots=all` in your `proc glm` statement, and including the interaction, e.g. `model lipids=age|dietfat;` Considering the sampling variability of estimating each μ_{ij} by one observation Y_{ij}, are the curves reasonably parallel? If so, that would indicate the additive model IV

$$Y_{ij} = \mu + \rho_i + \tau_j + \epsilon_{ij},$$

   is appropriate.

2. Still just using the plot in part 1, what generally happens to the ability to reduce lipids as age increases? Why might this be? Which diet is best at reducing lipids, i.e. has the highest *lipid reduction*?

3. As a double check, perform Tukey's 1 *df* test for additivity for these data and report the p-value and conclusions of your test. That is, test $H_0 : \delta = 0$ in the model

$$Y_{ij} = \mu + \rho_i + \tau_j + \delta\rho_i\tau_j + \epsilon_{ij}.$$

4. Fit the additive model IV to the data and write down the fitted model. Test whether there's differences in diet, e.g. $H_0 : \tau_j = 0$ at the 5% level using the Type III p-value. Was blocking on age effective?

5. Look at all three pairwise differences for diet using Tukey's procedure and make a "lines plot" for the three diet types with an overall FER of 5%.

6. Examine the standard SAS diagnostic plot and comment on the e_{ij}'s vs. the $\hat{Y}_{ij}$'s, the normal probability plot of the $\{e_{ij}\}$, and the t_{ij}'s vs. the $\hat{Y}_{ij}$. Is normality reasonable? Does variance seem roughly constant with the mean? Are there any outliers (e.g. $|t_{ij}| > 3$)?

7. Finally, prepare plots of the e_{ij} vs. i, and e_{ij} vs. j. Does constant variance seem reasonable across blocks and treatments?

ANCOVA, Eye contact: Consider the data of problems 19.12 and 22.17.

1. Define a (categorical) group variable that takes on the values 11, 12, 21, 22 for all possible pairings of eye contact and gender. Plot the success rating versus age using the grouping variable you just defined with four LOESS curves superimposed. Do you see four *approximately* parallel lines? If so, the ANCOVA model with two factors is appropriate
$$Y_{ijk} = \mu + \gamma x_{ijk} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk},$$
where $i = 1, 2$ denotes eye contact with the camera or no eye contact, and $j = 1, 2$ denotes a male or female personnel officer.

2. Fit the model in part 1 and test $H_0 : (\alpha\beta)_{ij} = 0$ at the 5% level using the Type III test. Can we drop the interaction?

3. Refit the ANCOVA model with additive treatments
$$Y_{ijk} = \mu + \gamma x_{ijk} + \alpha_i + \beta_j + \epsilon_{ijk}.$$
State the fitted model. Using the Type III tests, does having eye contact with the camera significantly affect the success rating at the 5% level? Which is better, eye contact or no eye contact? Does the gender of the personnel officer significantly affect the success rating at the 5% level? Who gives higher success ratings (i.e. who is nicer) in general, men or women?

4. Roughly, how does success rating change with the age of the personnel officer? Is the effect significant? Quantitify it using $\hat{\gamma}$.

5. Report SAS's standard diagnostic panel from `proc glm`. Comment on whether modeling assumptions are met.

6. Finally, prepare plots of the residuals e_{ijk} versus i, e_{ijk} versus j, and e_{ijk} versus x_{ijk} and comment on constant variance in each.