

Review of Two-factor Completely Randomized Designs

Respond to the following questions individually then discuss your answers as a group. You should hand in your individual response. We will discuss your group responses and then I will lecture on other topics.

A breast cancer experiment studied expression level of an enzyme (TS) as a function of tumor stage (I, II, III) and Ethnicity (African-American and European-American). Analyze the data in SAS and comment on residuals. Typical SAS code would be:

```
data a; set cancer;
proc glm;
class Ethnicity Stage;
/* The symbol between Ethnicity and Stage is Shift-\, it expands the */
/* expression to include the interaction term AND main effects      */
model TS_Expression=Ethnicity|Stage;
output out=outa r=resid p=yhat;
proc plot;
plot resid*yhat;
run;
```

Look at interaction plots in Minitab (these are found under **Stat>ANOVA**, rather than **Graph**); be sure to vary the input order of the two factors and comment on what you learned from the two different sets of plots. In reality, this data set was badly balanced with one empty cell and reaching definitive conclusions was quite difficult.