**Analysis of Variance in SAS**

### The Preliminaries

Dale Carnegie stated that smiling helps win friends and influence people. Research on the effects of smiling has backed this up and shown that a smiling person is judged to be more pleasant, attractive, sincere, sociable, and competent than a non-smiling person.

There is evidence that smiling can attenuate judgments of possible wrongdoing. This phenomenon termed the "smile-leniency effect" was the focus of a study by Marianne LaFrance & Marvin Hecht in 1995.

The questions that LaFrance and Hecht sought to answer were “Does smiling increase leniency (in terms of a judgment in a criminal trial)?” and “Are different smile types differentially effective?” They presented “judges” with four pictures of a single person. In one picture the person is smiling a false smile. In one picture, the person is smiling a genuine smile. In one picture, the person is smiling a miserable smile, and in the last picture the person has a neutral expression. Each judge saw only one of the conditions; therefore, the observations are independent. The response variable is a rating of how lenient the judgments were for each judge.

### The Procedures

SAS Procedures Used: PROC Univariate, PROC Means, PROC ANOVA, PROC GLM, PROC GPLOT, Output statement

### The Problems

1. Checking Assumptions

**/\* Code to enter the data \*/**

**filename smile 'C:\correct\pathname\leniency.csv';**

**data smile;**

**infile smile firstobs = 2, dlm = ‘,’;**

**input smile rating;**

**run;**

**Proc Print data=smile; run;**

**/\* Always print the data to check that it has been \*/**

**/\* entered correctly! \*/**

Check the ANOVA assumptions using PROC UNIVARIATE. Specifically, do the following:

1. First of all, name the factors, levels, response, and subjects.
2. Use PROC Boxplot to create parallel boxplots for the four conditions. Explain what you see in the boxplots.
3. Find the mean, median, standard deviation, and interquartile range for each group. Explain what each of these numbers tells you about the distribution of the observations for each group.
4. Analysis of Variance

There are two main ways to run an analysis of variance in SAS. The first procedure is PROC ANOVA and the second is PROC GLM (GLM = General Linear Model). PROC ANOVA is computationally faster of the two, and is for balanced, well-behaved data. It does not allow for post hoc tests or contrasts. PROC GLM is the more powerful procedure (in the sense of what it can do, not in the sense of statistical power). It can run post hoc tests and contrasts. So, we will use PROC GLM most of the time.

* 1. Run an ANOVA in SAS. Here is the code:

**Proc GLM data=smile;**

**class smile;**

**model rating = smile;**

**run;**

* 1. Is there evidence that leniency varies with smile type? Justify your answer.

1. Post-hoc Tests of Significance.

Below is an excerpt from a SAS help document that shows the types of post hoc tests available in SAS.

MEANS effects / options ;

The following MEANS statement options are used to select a multiple comparison procedure:

BON DUNCAN DUNNETT DUNNETTL DUNNETTU GABRIEL GT2 LSD REGWF REGWQ SCHEFFE SIDAK SMM SNK T TUKEY WALLER.

The following MEANS statement options specify details for the multiple comparison procedure: ALPHA= p CLDIFF CLM

To perform post-hoc analyses in SAS, you need to add a “means” statement after the model statement (see code below).

**Proc GLM data=smile;**

**class smile;**

**model rating = smile;**

**means smile / bon Tukey SNK REGWQ;**

**run;**

* 1. Use the Bonferroni, Tukey, SNK, and REGWQ procedures for post-hoc analysis. State the method, the test statistic for that method, the p-value for each of the pairwise comparisons and the conclusion for each pair.
  2. Which method has the fewest significant pairs? Which method has the most?

1. Comparisons to a control

The Dunnett’s test is used for comparisons of each group to a control group. The syntax is a little different, since the control group must be specified. Is there evidence that the mean rating for the control group is statistically significantly from the mean ratings of other groups?

**Proc GLM data=smile;**

**class smile;**

**model rating = smile;**

**means smile / dunnett(‘CONTROL’); /\* You have to supply the control group! \*/**

**run;**