Name: _____

1 Objective: Two Way ANOVA

In this lab we'll follow Littell/Freund/Stroup, SAS System for Linear Models, in their treatment of a 2 way anova example.

2 Key Words

`proc univariate, proc glm`, two way ANOVA, multiple comparison, factorial experiment.

3 Brief Description

Suppose we have:
3 Methods of promoting seed growth : `M`
5 Varieties of turf grass seed : `V`
6 pots planted with each `M, V` Combination.

Suppose the resulting 90 pots randomly placed in growing chamber and `Y = yield` measured at the end of four weeks.

This gives a 3×5 completely randomized factorial experiment. The data are in the file `grass.dat`, with each of the 6 yields given in each row. In addition the yields were multiplied by 10 to save entering decimal.

4 Lab Steps

[A] Read in the data and put it in ANOVA format.

[1] The following code reads the data structured as given:

```
DATA grass1;
        infile 'C:...\grass.dat';
        input m $ v y1-y6;
run;

PROC print data=grass1;
run;
```

[2] We need this in ANOVA format one long list of yields with corresponding treatment levels. Use this code to transform:

```
DATA grass;
        set grass1;
        drop y1-y6;
        y=y1/10; output;
        y=y2/10; output;
        y=y3/10; output;
        y=y4/10; output;
        y=y5/10; output;
        y=y6/10; output;
run;

PROC print data = grass;
run;
```

[B] Plot cell means.

[1] Find cell means.

```
PROC means data = grass noprint;
        by m v;
        output out = cellmn mean = ymean;
run;

PROC print data = cellmn;
run;
```

[2] Plot cell means vs `v` using `m` as symbol.

```
PROC plot data = cellmn;
        plot ymean*v = m;
run;
```

[3] SKETCH AND TURN IN the plot, connecting the same methods A, B, C.

[4] Which method produces most yield?

[5] Is there interaction, i.e. does the magnitude of the difference in cell means depend on the variety? Or are the plots parallel? Explain in a short paragraph.

[C] Perform an ANOVA test for interaction.

[1] Use the following `proc glm`:

```
PROC glm data = grass;
        class m v;
        model y=m v m*v;
run;
```

[2] What is the F-stt p-value for testing H_0: no interaction? (That is given by type III ss for `m*v`).

[D] Now, perform Duncan test of equality, of method means by variety. To do this, we must sort by variety.

[1] Sorting by variety:

```
PROC sort data = grass out = grass2;
        by v;
run;
```

[2] Perform two way anova (`proc glm`)

```
PROC glm data = grass2;
        by v;
        class m;
        model y = m;
        means m/duncan;
run;
```

[3] Look back at your plot. Because of randomness in data, you might think method A & C are the same and B less for variety 2. Look at Duncan's output for that variety. What does Duncan conclude? TURN IN: explain in a short paragraph.

[4] Look at plot for variety 3, What does Duncan say? TURN IN: explain in a short paragraph.

[E] Now try contrasts and estimation. NOTE: CAN ONLY ESTIMATE LINEAR CCMBINATIONS OF CELL MEANS.

[1] Recall relation between cell means and overparameterized model (We have notational difficulties here. So write MU for `mu`, `m` for model, `v` for variety).

$$MU_{ij} = MU + m_i + v_j + (mv)_{ij}$$

[2] Recall the three steps need to code estimates and contrasts into SAS:

[a] Write linear combination to be estimated in terms of cell means.

[b] Convert to model parameters using ****.

[c] Gather like terms.

[3] Suppose you want to estimate `.4*MU_A1 + .6*MU_B1`.
Perform the steps listed and then complete the code to get estimate:

```
PROC glm data = grass;
        class m v;
        model y = m v m*v;
        estimate '.4A+.6B in V1' intercept ??? m ??? v ??? m*v ???;
run;
```

TURN IN CODE, ESTIMATE AND STD ERR:
Check your answer using the appropriate weighted average of cell means from [B1].

[4] Suppose you want to estimate `.4*MU_A1 + .6*MU_B1 - MU_C1`.
Perform the steps listed and then complete the code to get estimate:
... estimate `'.4 A + .6 B - C in V1' intercept ??? m ??? v ??? m*v ???;`
TURN IN CODE, ESTIMATE AND STD ERR:

[5] Suppose you want to estimate `.4*MU_Abar + .6*MU_Bbar - MU_Cbar`, the simple average of method means across the variety all 5 varieties (This makes sense if NO INTERACTION, which is not the case, but do it anyway).
Perform the steps listed and then complete the code to get estimate:
... estimate `'.4 A + .6 B in V1' intercept ??? m ??? v ??? m*v ???;`
TURN IN CODE, ESTIMATE AND STD ERR: