Name: _____

1 Objective: Repeated Measures Designs and Accuracy

We'll consider two things today: [A] Repeated measures as given in the blood glucose data, and [B] Accuracy of calculations as checked using the `StRD` = statistical reference data sets.

2 Key Words

`proc univariate, proc glm`, repeated measures.

3 Lab Steps

We will consider the blood glucose data stored in `glucose.dat` for this lab. Note also that you have a printed version of it as a handout. Notice the missing values noted as -1.

[A] We'll read `glucose.dat` into SAS and make a few calculations.

[1] Open the file `glucose.dat` data in a word processor to see how it differs from the printed version:
- [a] no column headers
- [b] no column "420 min"
- [c] no meal time column

[2] Read the data into SAS; along the way create a new column, `tm`, which tells something about the time of the meal. For example, let `tm = 0` for the first 6 rows (the 10am real time), `tm = 1` for the next 6 (the 2pm meal time), etc. HINT: use the `int` function applied to `_n_/6`, or something close. You might also want to use `proc format` [CS 77-87]

[3] Find the means and standard deviations for `g1`, the glucose levels just before the meals, using class `tm`.

[4] Use anova or `glm` to test whether the means for class `tm` are the same for `g1`. For `g3`.

[5] If you use these data for a project you should
- [a] remove the two outliers coded as -1.00.
- [b] make plots like those shown on the handout.
- [c] make one plot of the separate `tm` class means across `gk`, `k` = 1:10.
- [d] run appropriate anova, using various statistics based on the glucose readings (for example `max(g1...g10), mean(g1...g10)`, area under curve,...)

[B] PROJECT - Comparing Accuracy:

There are data sets for which statistics calculations have been carried out to 500 decimal places, so that all answers are essentially exact, called CERTIFIED VALUES, These answers can be compared

with those produced on a day to day basis by statistics software like SAS and R, and other software like XL.

These data sets are available at "http://www.itl.nist.gov/div898/strd/general/dataarchive.html". We'll use the anova data set `AtmWtAgt.dat` which is copied from there. If you do a project in this area, you'll need to copy other interesting sets, from the UNIVARIATE, ANOVA, and REGRESSION data sets. And you'll have to make some tables to show how SAS accuracy compares with the certified values.

[1] Look at the data set `AtmWtAgt.dat` to see that it has two variables, `ins` = instrument used to measure = 1 or 2 , and `wt` = the measured weight. Note also that there are 60 lines of preamble before the actual data. Actual data lines are between 61st and 84th lines.

[2] Read the data into SAS (HINT: `firstobs` = 61). Print it. Is it all right, or all the same?

[3] Read it again, this time using the line `"format wt 20.15;"` following input. Print the data. Better?

[4] Apply `proc mean` to the data, by `ins`. What do you get?

[5] Repeat [4] with `fw` = 25 or `maxdec` = 12. What?

[6] Now try printing the output from

```
PROC means data = atwt noprint;
        by inst;
        output out = mnout mean = mn std = sd;
        format mn sd 25.20;
run;
```