EXST 7015
Fall 2014
Lab 13: Factorial Design and EMS

OBJECTIVES:

The objective of an experimental design is to provide the maximum amount of reliable information at the minimum cost. In statistical terms, the reliability of information is measured by the standard error of estimates (that is directly related with the population variance, inversely related to sample size). Properly applied experimental design may effectively reduce the population variance, and/or could structure data collection to reduce the magnitude of the experimental error. Usually data resulting from the implementation of experimental designs are described by linear model and analyzed by the analysis of variance as was introduced in previous week.

A factorial design is an experimental design in which two or more series of treatments are tried in all combinations. Such an experiment allows studying the effect of each factor on the response variable, as well as the effects of interactions between factors on the response variable. A full **factorial design** may also be called a **fully crossed design**. The following linear model describes a factorial design with three treatment factors:

$$y_{ijkl} = \mu + \tau 1_i + \tau 2_j + \tau 3_k + \tau 4_{ij} + \tau 5_{ik} + \tau 6_{jk} \ \ \tau 7_{ijk} + \varepsilon_{ijkl}$$

$$(i = 1, 2, \ldots, t1;\ j = 1, 2, \ldots, t2;\ k = 1, 2, \ldots, t3;\ l = 1, 2, \ldots, n)$$

Where y_{ijkl} is the observed individual of the response variable; μ is the overall mean; $\tau 1_i$ is the fixed effect of treatment i; $\tau 2_j$ is the fixed effect of treatment j; $\tau 3_k$ is the fixed effect of treatment k. $\tau 1_i$, $\tau 2_j$ and $\tau 3_k$ are called main effects. $\tau 4_{ij}$, $\tau 5_{ik}$ and $\tau 6_{jk}$ are two-way interaction effects, and $\tau 7_{ijk}$ is the three-way interaction effects. Keep in mind that the effects of interactions are fixed effects too. ε_{ijkl} is the sampling error, a random variable with mean zero and variance σ^2.

Factorial design has several important features. First, it has great flexibility for exploring the treatment effects. Second, factorial designs are efficient. Instead of conducting a series of independent studies one can effectively combine several studies into one. Finally, factorial designs are the only effective way to examine interaction effects between the treatments.

LABORATORY INSTRUCTIONS

Housekeeping Statements

```
dm 'log; clear; output; clear';
options nodate nocenter pageno = 1 ls=78 ps=53;
title1 'EXST7015 lab 13, Name, Section#';
ods rtf  file = 'c:/temp/lab13.rtf';
ods html file = 'c:/temp/lab13.html';
```

The data for today's assignment is taken from Kuehl(2000). The experiment was designed to test the effects of salinity **(dS/m)**, water content (%), and soil type on the electro-conductivity of soil. It is a 2X2X2 factorial design. Below is the SAS code that you will use for this lab. Feel free to edit the code however you like (adding options and titles etc.).

```
data salinity;
input y salinity water soil $;
datalines;
0.60  2  0  sand
0.48  2  0  sand
3.47  2  15  sand
3.30  2  15  sand
1.37  2  0  clay
1.50  2  0  clay
5.74  2  15  clay
5.38  2  15  clay
0.07  16  0  sand
0.06  16  0  sand
0.22  16  15  sand
0.17  16  15  sand
0.40  16  0  clay
0.57  16  0  clay
1.95  16  15  clay
2.87  16  15  clay
;

proc mixed data=salinity method=type3;
  class salinity water soil;
  model y= salinity | water | soil / outp=outdata;
run;
proc univariate data=outdata normal plot;
  var resid;
run;
proc plot data=outdata;
  plot resid*pred;
run;

proc mixed data=salinity method=type3;
  class salinity water soil;
  model y= salinity | water / outp=outdata;
  random soil salinity*soil water*soil salinity*water*soil;
run;
proc univariate data=outdata normal plot;
  var resid;
run;
proc plot data=outdata;
  plot resid*pred;
run;
```

Method=type3: Forces mixed procedure produce the Expected Mean Squares.

The vertical bar "|": Tells SAS to model salinity, water, soil and all possible interactions. This works in both **proc glm** and **proc mixed**, and is very useful when dealing with large models and multiple interaction terms.

LAB ASSIGNMENT:

1. Write the linear model for this factorial experiment. Clearly describe each term in the model, and the range of the subscripts. Write the null hypotheses that you are testing (there are seven of them!).
2. Consider all three treatments fixed and write the ANOVA source table, including degrees of freedom and expected mean squares.
3. Use proc mixed to test the hypotheses you stated in Question 1. Report your results, including your F-values and degrees of freedom, p-values, and conclusions for each test. List the assumptions necessary for your analysis and determine whether they have been violated. Provide evidence to support your conclusions.
4. Consider the treatment soil random and write the ANOVA source table, including degrees of freedom and expected mean squares.
5. Use proc mixed to test the hypotheses you stated in Question 1. Report your results, including your F-values and degrees of freedom, p-values, and conclusions for each test. List the assumptions necessary for your analysis and determine whether they have been violated. Provide evidence to support your conclusions.