EXST 7015
Fall 2014
Lab 12: Latin Square Design and Expected Mean Square

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.

In last week's lab, a special RBD example was exercised, which help you understand why it is important to consider the block effect in the experimental design. You are also getting familiar with how to construct source ANOVA table with variance sources, their degree of freedom and expected mean squares. Today, **Latin Square Design (LSD)** with row and column blocks will be introduced. Compared with factorial design, LSD is more efficient which allows the researchers to control variations in two directions with minimum experimental unit. In LSD, there are equal numbers of treatment, row block and column block. Treatments are assigned at random within rows and columns, with each treatment only once per row and only once per column. Two blocks of column and row are in a factorial structure. The linear model for data from such an experiment is

$$\mathbf{y_{ijk} = \mu + a_i + b_j + \tau_k + e_{ijk}\ (\mathit{i} = 1, 2, \ldots, r;\ \mathit{j} = 1, 2, \ldots, r;\ \mathit{k} = 1, 2, \ldots, r)}$$

Where $\mathbf{a_i}$ is the effect of row block $\boldsymbol{i}$; $\boldsymbol{\beta_j}$ is the effect of column block $\boldsymbol{j}$; $\mathbf{y_{ijk}}$ is the observed value of the response variable in the $\boldsymbol{i}^{\mathbf{th}}$ row of the column $\boldsymbol{j}$; $\boldsymbol{\mu}$ is the overall mean; $\boldsymbol{\tau_k}$ is the fixed effect of treatment $\boldsymbol{k}$; $\boldsymbol{\varepsilon_{ijk}}$ is the random error with mean zero and variance σ^2.

The Latin Square Design is appropriate only if effects of all three factors (row block, column block and treatment) are additive, i.e., all interactions are zero. It is a very important assumption of Latin Square Design. The ANOVA table of LSD is as the following:

Source	DF	EMS
Treatment	r – 1	$\sigma^2 + r\,\sigma^2_\tau$
Row	r – 1	$\sigma^2 + r\,\sigma^2_\alpha$
Column	r – 1	$\sigma^2 + r\sigma^2_\beta$
Error	(r-1)(r-2)	σ^2
Total	r^2 – 1	

PROC MIXED will be used to analyze a set of data with 4x4 Latin Square Design. **Expected Mean Squares** will be required in PROC MIXED to further help you understand the ANOVA table.

LABORATORY INSTRUCTIONS

Housekeeping Statements

```
dm 'log; clear; output; clear';
options nodate nocenter pageno = 1 ls=78 ps=53;
title1 'EXST7015 lab 12, Name, Section#';
ods rtf  file = 'c:/temp/lab12.rtf';
ods html file = 'c:/temp/lab12.html';
```

Latin Square Design:

The data set is from **Statistics for Experimenters** by Box, Hunter and Hunter (Chapter 8). An experiment was carried out to compare four gasoline additives (A, B, C and D) for reducing NOX emissions from cars. Four drivers (I, II, III and IV) and four cars (1, 2, 3 and 4) will be used in the experiment. Because there may be significant driver effects and car effects, we would like to block on both variables. A Latin square design accomplishes this. In this design, each additive is used exactly once by each driver, and exactly once in each car. However, not every combination of driver and car is used with each additive. The design can be visualized as a square with four rows corresponding to the drivers and four columns corresponding to the cars. Each additive occurs exactly once in each row and exactly once in each column.

```
data autoemission;
  input car driver $ additive $ reduct;
  cards;
1   I     A   21
2   I     B   26
3   I     D   20
4   I     C   25
1   II    D   23
2   II    C   26
3   II    A   20
4   II    B   27
1   III   B   15
2   III   D   13
3   III   C   16
4   III   A   16
1   IV    C   17
2   IV    A   15
3   IV    B   20
4   IV    D   20
run;
Proc mixed data=autoemission cl covtest method = REML;
  class car driver additive;
  model reduct = additive/ddfm=satterth outp=outdatalsd1;
  random car driver;
Run;
proc univariate data=outdatalsd1 normal plot;
```

```
  var resid;
run;
proc plot data=outdatalsd1;
  plot resid*pred;
run;

Proc mixed data=autoemission cl covtest method = type3;
  class car driver additive;
  model reduct = additive/ddfm=satterth outp=outdatalsd3;
  random car driver;
Run;

proc univariate data=outdatalsd3 normal plot;
  var resid;
run;
proc plot data=outdatalsd3;
  plot resid*pred;
run;
```

CL: requests confidence limits for the covariance parameter estimates.

COVTEST: produces asymptotic standard errors and Wald Z-tests for the covariance parameter estimates. The method is based on computing the covariance matrix and returning the diagonal elements.

METHOD = TYPE3: forces PROC MIXED procedure to produce the Expected Mean Square.

LAB ASSIGNMENT:

1. Write the linear model and source ANOVA table for the Latin Square Design in symbolic notation. The ANOVA table should include the variance sources, degree of freedom and expected mean squares.

2. According to the model, do different additives have significant effect on reducing NOX emission? Can the effects of Cars or Drivers be ignored in this model?

3. There is an important assumption for Latin Square Design besides the usual assumptions for the linear model (normality, homogenous variance et al). What is that?