Mixed Procedure and Covariances

**Submitted by:**

Binhong Cai

**Submitted to:**

Submitted to:

**Dr. Jack Mardekian**

The mixed procedure is used in providing normally distributed data with means, expected values, variances, and covariances of the data.

**Submitted on:**

April 20, 2020

Computing and Graphics in Applied Statistics

01:960:486

Rutgers University

**Problem**: Covariances and Mixed procedure analysis for self-modified corn data

**Program output**: 

**Results:**

An iteration gives a bunch of data suggesting and counting how many covariances parameters are there. There are 8 iterations in total, which decrease from 318.67817035 to 316.37065477, indicating that existing covariance data structure is linear and its slope tends to go down.

SAS Code:

data selfModifiedCornTrials; /\*trying to test different trils of corn\*/

input trial month @@;

monthc = month; /\*a new variable created in later use\*/

do i = 1 to 2; /\*two lists of data so it is 1 to 2 as listed below coln from 3 to 4\*/

input y @@;

output;

end;

datalines;

1 0 32.3 97.2

1 4 34.9 103.3

1 1 30.1 52.0

1 5 36.9 65.3

1 4 26.8 20.3

1 5 30.5 64.6

2 9 33.3 73.6

2 12 29.7 12.5

2 3 35.0 93.5

2 8 29.9 24.2

2 8 35.2 35.2

2 7 38.3 64.8

3 0 35.2 26.5

3 2 35.5 35.6

3 2 36.7 30.6

3 1 26.8 26.9

3 6 38.0 39.5

3 3 31.7 86.1

;

proc mixed data = selfModifiedCornTrials; /\*general procedure of mixed proc, model information in Random Coefficient Analysis\*/

class trial;

model Y = month / s;

random Int month / type = un sub = trial s;

run;

proc mixed data=selfModifiedCornTrials; /\*random coefficient with nested Error Analysis\*/

class trial;

model Y = month / s;

random Int month monthc / sub = trial s;

run;

**Data Modified from**:

Corn Yields and Rainfall DATA DESCRIPTION:

There are three variables:

year: from 1909 to 1927

yield: yearly corn yield in bushels per acre, in six Corn Belt

states(Iowa, Illinois, Nebraska, Missouri, Indiana, and

Ohio).

rain: rainfall measurements in inches, in the six states.

SOURCE:

M.Ezekiel and K.A.Fox, Methods of Correlation and

Regression Analysis, p.212.Copyright 1959, John Wiley

and Sons, Inc., New York.

Data originally from E. G.

Misner, "Studies of the Relationship of Weather to the

Production and Price of Farm Products, I. Corn",

mimeographed publication, Cornell University, March 1928.

**Approaching Analysis and Information**:           There are some explanations of variances and covariances working in Mixed procedure. Any two sets of parameters of a mixed linear model stipulate a complete dispersion of data due to Gaussian data being modeled utterly in terms of variances or covariances. Moreover, Fixed-effects parameters are parameters of the mean model, and the parameters of the variance-covariance model define covariance parameters. Accordingly, the fixed-effect parameters can be equated as explanatory variables used in the linear model. Those can either be qualitative or quantitative. However, what distinguishes the mixed linear model from the standard linear model is the covariance parameters. Two typical cases when we need to use covariances: The data from a standard cluster are correlated when the experimental units on which the data are measured can be grouped into those clusters; Repeated measurements are recorded in the same experimental unit, and those are correlated or show changing variability.

           Mixed Procedure in SAS offers a variety of options to create statistical inferences with different sorts of information. Analysis progress is similar to the GLM procedure learned in class, but the Mixed procedure fits a broader class of linear models with statistical tables than that of GLM.

Both of the coding structure include “LSMEANS,” “ESTIMATES,” “CONTRAST,” “MODEL,” and “CLASS,” but “REPEATED” and “RANDOM” functions are quite different. Also note that it is similar to REG, which has a CLASS statement. It allows us to analyze both categorical and numeric variables.

           Many types of output tables that explain various aspects of statistics, including “REPEATED” and “RANDOM” functions, as stated above. Mixed proc combines random effects with the vector by using a random statement. In tune, GLM proc deems it as a fixed model and a tool of presenting mean squares. Mixed proc is prone to analyze “ML” predictions of variances parameters. Additionally, it also generates “MIVQUE0” in the ANOVA procedure.

           Individually, by using Mixed procedure, we can obtain covariance structure, compound symmetry, unstructured, general linear and “Toeplitz,” appropriate standard errors of fixed and random effects corresponding F or T-tests for many specified linear combinations, some cluster or group effects that enable heterogeneity and blocking, “REML” and “ML” estimation methods implemented with a Newton’s algorithm, ability to generate unbalanced data with corresponding tables.