Biostat 656: Lab 2 supplement

1. Introduction to SAS PROC MIXED

The MIXED procedure provides you with flexibility of modeling not only the means of yours data (as in the standard linear model) but also their variances and covariance as well (the mixed linear model).

- **THE MIXED LINEAR MODEL**
  - **The standardized linear model**
    - $\mathbf{Y=X\beta + \varepsilon}$
    - $\boldsymbol{\beta}$ is an unknown vector of fixed-effects parameters with known design matrix X.
    - $\boldsymbol{\varepsilon}$ is assumed to be independent and identically distributed Gaussian random variables with mean 0 and variance σ^2.
  - **The mixed linear model**
    - a generalized version of the standardized linear model as follows:

      $\mathbf{Y=X\beta + Zb + \varepsilon}$
    - $\mathbf{b}$ is an unknown vector of random-effects parameters with known design matrix Z
    - $\boldsymbol{\varepsilon}$ is the residuals vector whose elements are no longer required to be independent and homogeneous, and its variance is **R**.
    - The variance of **Y** is $\mathbf{V = ZGZ' + R}$
    - For **G** and **R**, you must select some covariance structure.
- **SYNTAX: (details refer to SAS help)**

PROC MIXED < options > **;**
BY variables **;**
CLASS variables **;**
ID variables **;**
MODEL dependent = < fixed-effects > < / options > **;**
RANDOM random-effects < / options > **;**
REPEATED < repeated-effect > < / options > **;**
PARMS (value-list) ... < / options > **;**
PRIOR < distribution > < / options > **;**
CONTRAST 'label' < fixed-effect values ... >
< | random-effect values ... > , ... < / options > **;**
ESTIMATE 'label' < fixed-effect values ... >
< | random-effect values ... >< / options > **;**
LSMEANS fixed-effects < / options > **;**
MAKE 'table' **OUT=**SAS-data-set **;**
WEIGHT variable **;**
RUN;

Let's look at **covtest method=options for PROC**, **CLASS**, **MODEL, RANDOM**, and **REPEATED**.

- **COVTEST**
  The option of COVTEST produces asymptotic standard errors and Wald *Z*-tests for the covariance parameter estimates G and R.
- **Method=option**
  The METHOD= option specifies the estimation method for the covariance parameters. The REML specification performs restricted maximum likelihood, and it is the default method. The ML specification performs maximum likelihood.
- **CLASS**
  - The CLASS statement names the classification variables to be used in the analysis.
  - If the CLASS statement is used, it must appear before the MODEL statement.
  - Classification variables can be either character or numeric.
- **MODEL**
  - MODEL dependent = < fixed-effects >< / options >;
  - The MODEL statement names a single dependent variable and the fixed effects, which determine the **X** matrix of the mixed model;
  - Solution option requests estimates for the fixed-effects parameters.
- **RANDOM**
  - RANDOM random-effects < / options >;
  - Define Z, **b,** G
- **REPEATED**
  - REPEATED < repeated-effect > < / options >;
  - Specify the R matrix in the mixed model.

2. Introduction to SAS PROC NLMIXED

The NLMIXED procedure fits nonlinear mixed models, which permit the outcome to be binary, count data etc. The models fit by PROC NLMIXED can be viewed as generalizations of the random coefficient models fit by the MIXED procedure.

- **SYNTAX: (details refer to SAS help)**

PROC NLMIXED options **;**
ARRAY array specification **;**
BOUNDS boundary constraints **;**
BY variables **;**
CONTRAST 'label' expression <,expression> **;**
ESTIMATE 'label' expression **;**
ID expressions **;**
MODEL model specification **;**
PARMS parameters and starting values **;**
PREDICT expression **;**
RANDOM random effects specification **;**

REPLICATE variable **;**
Program statements ;

Let's look **MODEL, PARMS, RANDO**M.

- **MODEL**
  - The MODEL statement specifies the conditional distribution of the data given the random effects;
  - You must specify a single dependent variable from the input data set, a tilde(~), and then a distribution with its parameters;
  - The distributions could be binary(p), binomial(n,p) and poisson(m) etc.
- **PARMS**
  - The PARMS statement lists names of parameters and specifies initial values.
- **RANDO**M
  - **RANDOM** *random-effects ~ distribution SUBJECT=variable <options>;*
  - The RANDOM statement defines the random effects and their distribution.
  - The only distribution currently available for the random effects is normal(*m*,*v*) with mean *m* and variance *v*.