Generalised Linear Models

Proc GENMOD

proc genmod data= ;

     class ;

     model y = x / dist= link= type3 ;

     repeated subject= / corr= ;

     lsmeans x / diff=control ('ref') cl adjust=;

run;

Parameters estimated with maximum likelihood methods. This is an iterative process, i.e., the computer adjusts the parameter until the log likelihood function is maximised.

Distribution	Link function	Note
normal	identity	continuous dependent variable, mean difference
binomial	log	risk ratio from exponentiation of the parameter estimate
binomial	logit	odds ratio
poisson	log	risk ratio from exponentiation of the parameter estimate; robust error estimates can be requested with the REPEAT statement
multinomial		categorical response variable with more than two levels

Sums of squares	Effect type	Note
Type I	Sequential	The SS for each factor is the incremental improvement in the error SS as each factor effect is added to the model. In other words it is the effect as the factor were considered one at a time into the model, in the order they are entered in the model selection.
Type II	Hierarchical or partially sequential	The SS for each factor is the reduction in residual error obtained by adding that term to a model consisting of all other terms that do not contain the term in question.
Type III	Marginal	The effect of each variable is evaluated after all other factors have been accounted for.

Adjustment for pairwise comparison

- Tukey: all pairs
- Dunnett: paired with control

Generalized Estimating Equations

GEEs are used for analysis of correlated data, e.g., subjects are measured at different points in time, or subjects are clustered, i.e., share a common characteristic. GEE analysis can be performed in GENMOD by specifying a REPEATED statement which provides clustering information and a working correlation matrix. The REPEATED statement requests a GEE analysis.

- Subject: Responses from different subjects are assumed to be statistically independent, and responses within subjects are assumed to be correlated. Variables used in defining the subject-effect must be listed in the CLASS statement. The input data set does not need to be sorted by subject.
- Corr= specifies the correlation structure: Un unstructured; IND independent.