

# Statistical Enhancements in Release 6.12 of the SAS® System

Maura Stokes  
SAS Institute Inc., Cary, NC

## Abstract

Release 6.12 of the SAS System brings several key enhancements to SAS/STAT® software. Generalized Estimating Equations (GEE) methodology, a popular technique for the modeling of correlated response data, has been incorporated into the GENMOD procedure. Another key update is the computation of exact  $p$ -values for a number of statistical tests provided by the FREQ procedure. In addition, this paper provides an overview of enhancements to the MULTTEST, MIXED, GLM, and NPAR1WAY procedures.

## GEE Methods in the GENMOD Procedure

GEEs were introduced by Liang and Zeger (1986) as a method of dealing with correlated data when, except for the correlation among responses, the data can be modeled as a generalized linear model.

Correlated data can arise from situations such as

- longitudinal studies, in which multiple measurements are taken on the same subject at different points in time. A clinical trial for a new treatment of psoriasis may involve repeated measurements of symptom severity.
- clustering, where measurements are taken on subjects that share a common category or characteristic that leads to correlation. Incidence of pulmonary disease among family members may be correlated because of hereditary factors.

The correlation must be accounted for by analysis methods appropriate to the data. Possible consequences of analyzing correlated data as if they were independent are incorrect inferences due to underestimated standard errors and inefficient estimators. You model the correlated data by using the same link function and linear predictor setup as you use in a generalized linear model in the independent case; you describe the random component by the same variance function. However, in the GEE approach, you also model the covariance structure of the correlated measures.

You model the covariance matrix of  $Y_i$  as

$$V_i = \phi A_i^{1/2} R(\alpha) A_i^{1/2}$$

where  $A_i$  is an  $n_i \times n_i$  diagonal matrix with  $v(\mu_{ij})$  as the  $j$ th diagonal element.

Usually, you estimate the working correlation matrix

$$e_{ij} = \frac{y_{ij} - \mu_{ij}}{\sqrt{v(\mu_{ij})}}$$

using the current value of the parameter vector  $\beta$  to compute appropriate functions of the Pearson residual. There are several choices for the form of the working correlation matrix, including a fixed correlation matrix, the identity matrix, exchangeable correlation structure, unstructured correlation structure, and AR(1) structure. Both a model-based estimator of  $Cov(\hat{\beta})$  and an empirical estimator are available; both estimators are consistent, and the empirical estimator is consistent even if the working correlation matrix is not specified correctly. See the paper *Repeated Measurements Analysis Using GEE Methodology* in these proceedings for additional detail about the GEE approach.

The GEE strategy is implemented in the GENMOD procedure with the new REPEATED statement. In the following example, this strategy is applied to the Six Cities study of the effects of air pollution (Ware et al. 1984) for the cases analyzed in Lipsitz et al. (1994). The binary response, the wheezing status of children at ages 9, 10, 11, and 12, is modeled with a logistic regression for the explanatory variables city, age, and maternal smoking status at that age. Correlations among the binary responses are modeled as exchangeable correlations.

```
data six;
  input case city$ @@;
  do i=1 to 4;
    input age smoke wheeze @@;
    output;
  end;
datalines;
```

```

1 p 9 0 1 10 0 1 11 0 1 12 0 0
2 k 9 1 1 10 2 1 11 2 0 12 2 0
3 k 9 0 1 10 0 0 11 1 0 12 1 0
4 p 9 0 0 10 0 1 11 0 1 12 1 0
5 k 9 0 0 10 1 0 11 1 0 12 1 0
6 p 9 0 0 10 1 0 11 1 0 12 1 0
7 k 9 1 0 10 1 0 11 0 0 12 0 0
8 p 9 1 0 10 1 0 11 1 0 12 2 0
9 p 9 2 1 10 2 0 11 1 0 12 1 0
10 k 9 0 0 10 0 0 11 0 0 12 1 0
11 k 9 1 1 10 0 0 11 0 1 12 0 1
12 p 9 1 0 10 0 0 11 0 0 12 0 0
13 k 9 1 0 10 0 1 11 1 1 12 1 1
14 p 9 1 0 10 2 0 11 1 0 12 2 1
15 k 9 1 0 10 1 0 11 1 0 12 2 1
16 p 9 1 1 10 1 1 11 2 0 12 1 0
;
proc genmod data=six;
  class case city;
  model wheeze = city age smoke / dist=bin;
  repeated subject=case / type=exch covb corrw;
run;

```

You request the GEE strategy with the REPEATED statement. Specifying SUBJECT=CASE identifies the clustering variable, which must also be listed in the CLASS statement. The TYPE=EXCH option specifies an exchangeable working correlation structure. The COVB and CORRW options request that the parameter estimate covariance matrix and the final working correlation be printed, respectively.

The model information table lists the total number of clusters as well as the minimum and maximum cluster sizes.

GEE Model Information	
Description	Value
Correlation Structure	Exchangeable
Subject Effect	CASE (16 levels)
Number of Clusters	16
Correlation Matrix Dimension	4
Maximum Cluster Size	4
Minimum Cluster Size	4

NOTE: GEE parameter estimates converged after 4 iterations.

Figure 1. GEE Model Information

Both the model-based and empirical covariance matrices are printed.

Covariance Matrix (Model-Based)				
Covariances are Above the Diagonal and Correlations are Below				
Parameter Number	PRM1	PRM2	PRM4	PRM5
PRM1	5.71511	-0.22386	-0.53133	0.01658
PRM2	-0.13847	0.45733	-0.002411	0.01877
PRM4	-0.96838	-0.01553	0.05268	-0.01658
PRM5	0.01587	0.06353	-0.16530	0.19088

  

Covariance Matrix (Empirical)				
Covariances are Above the Diagonal and Correlations are Below				
Parameter Number	PRM1	PRM2	PRM4	PRM5
PRM1	9.33891	-0.85121	-0.83232	-0.16667
PRM2	-0.40467	0.47378	0.05737	0.04007
PRM4	-0.97676	0.29893	0.07775	-0.002201
PRM5	-0.15108	0.16125	-0.02187	0.13032

Figure 2. GEE Covariance Matrices

The working correlation matrix is displayed.

Working Correlation Matrix				
	COL1	COL2	COL3	COL4
ROW1	1.0000	0.1686	0.1686	0.1686
ROW2	0.1686	1.0000	0.1686	0.1686
ROW3	0.1686	0.1686	1.0000	0.1686
ROW4	0.1686	0.1686	0.1686	1.0000

Figure 3. GEE Working Correlation Matrix

Finally, the parameter estimates table contains parameter estimates, standard errors, confidence intervals, Z-scores, and *p*-values for the parameter estimates.

Analysis Of GEE Parameter Estimates					
Empirical Standard Error Estimates					
Parameter	Estimate	Empirical Std Err	95% Confidence Limits		Z
			Lower	Upper	
INTERCEPT	1.2754	3.0560	-4.7141	7.2650	0.4174
CITY k	0.1219	0.6883	-1.2272	1.4709	0.1771
CITY p	0.0000	0.0000	0.0000	0.0000	0.0000
AGE	-0.2036	0.2788	-0.7501	0.3429	-7.302
SMOKE	-0.0928	0.3610	-0.8003	0.6147	-2.571
Scale	0.9991	.	.	.	.

Figure 4. GEE Parameter Estimates Table

### More Exact *p*-Values

Exact *p*-values provide an alternative strategy when data are sparse, skewed, or unbalanced so that the assumptions required for standard asymptotic tests are violated. Advances in computer performance and developments in network algorithms over the last decade have made exact *p*-values accessible for a number of statistical tests. In Release 6.11, exact *p*-values were added for the simple linear rank statistics produced by the NPAR1WAY procedure. In Release 6.12, exact *p*-values are produced for

many of the statistics computed by the FREQ procedure. You are now able to request exact  $p$ -values for the following chi-square statistics: Pearson's chi-square, likelihood-ratio chi-square, Mantel-Haenszel chi-square, Fisher's exact test and  $r$  by  $c$  exact test, Jonckheere-Terpstra test, and McNemar's test. In addition, you can also obtain exact  $p$ -values for hypothesis tests that the following statistics are equal to 0: Pearson correlation coefficient, Spearman correlation coefficient, simple kappa statistic, and weighted kappa statistic. Exact confidence bounds are also available for the odds ratios produced for 2 by 2 tables.

The following example illustrates the use of the new EXACT statement to produce an exact  $p$ -value for the simple kappa statistic. Researchers studied two scoring systems for evaluating fitness in fifth grade students. Forty-three students were classified into one of four fitness categories. Interest lies in determining whether there is agreement between the two scoring systems, which can be assessed by testing whether the kappa coefficient is equal to 0.

```
data fitness;
  input score_1 $ score_2 $ count;
  datalines;
  poor poor 5
  average average 4
  good good 4
  superior superior 3
  poor average 3
  average poor 1
  average good 6
  good average 5
  good superior 1
  superior average 10
  superior good 1
  ;
```

To request the exact  $p$ -value for the kappa statistic, you specify the keyword KAPPA in the EXACT statement. The AGREE option in the MODEL statement requests the measures of agreement.

```
proc freq;
  weight count;
  tables score_1 * score_2 / agree;
  exact kappa;
run;
```

The following figure displays the contingency table form of the data. Note the number of zero cells, which makes the use of the asymptotic test questionable.

SCORE_1	SCORE_2				
	average	good	poor	superior	Total
average	4 9.30 36.36 18.18	6 13.95 54.55 54.55	1 2.33 9.09 16.67	0 0.00 0.00 0.00	11 25.58
good	5 11.63 50.00 22.73	4 9.30 40.00 36.36	0 0.00 0.00 0.00	1 2.33 10.00 25.00	10 23.26
poor	3 6.98 37.50 13.64	0 0.00 0.00 0.00	5 11.63 62.50 83.33	0 0.00 0.00 0.00	8 18.60
superior	10 23.26 71.43 45.45	1 2.33 7.14 9.09	0 0.00 0.00 0.00	3 6.98 21.43 75.00	14 32.56
Total	22 51.16	11 25.58	6 13.95	4 9.30	43 100.00

Figure 5. Exact Test for Simple Kappa

The resulting exact  $p$ -value for the hypothesis that the simple kappa statistic is equal to 0 is  $p=0.055$ , which may be considered to have marginal significance at best. Note the value  $p=0.038$  for the asymptotic test. Using exact  $p$ -values for this analysis leads to a very different conclusion than using the asymptotic test.

Test of Symmetry			
-----			
Statistic = 11.091	DF = 6	Prob = 0.086	
Simple Kappa Coefficient			
-----			
95% Confidence Bounds			
Kappa = 0.167	ASE = 0.102	-0.032	0.366
Asymptotic P-Values		Exact P-Values	
(Right-sided) = 0.019	(Right-sided) = 0.034		
(Two-sided) = 0.038	(Two-sided) = 0.055		
Weighted Kappa Coefficient			
-----			
95% Confidence Bounds			
Kappa = 0.100	ASE = 0.100	-0.096	0.297
Sample Size = 43			

Figure 6. Exact Test for Simple Kappa

There have been other additions to the FREQ procedure. These include

- goodness-of-fit test for one-way tables
- test for specified proportions or frequencies for one-way tables
- confidence bounds for MEASURES statistics

PROC FREQ users will also be pleased to know that the procedure no longer truncates values of character variables to 16 characters.

### Tests for Homogeneity of Variance

One of the assumptions of one-way ANOVA is *homogeneity of variance* (HoV)—that is, that the groups have equal variance. The GLM procedure in Release 6.12 of SAS/STAT software includes several methods for testing HoV in a one-way model and also provides Welch's HoV-robust test for group mean differences.

The traditional test for HoV is Bartlett's (Bartlett 1937). While this is the optimal test under a normal model, it can be very inaccurate when the data deviate from normality (Box 1953). Levene's test, O'Brien's test, and the Brown and Forsythe test are all based on an ANOVA for a dispersion variable derived from the dependent variable values. These tests are fairly robust to underlying distributions (Conover, Johnson, and Johnson 1981).

These tests are now available for groups defined in the MEANS statement and are requested with the HOVTEST= option and the following keywords: BARTLETT, LEVENE, OBRIEN, and BF. Specifying HOVTEST alone produces Levene's test. The option is ignored unless the MODEL statement specifies a simple one-way model; investigating homogeneity of variance for more complicated models is an area of ongoing research.

The following example displays the use of homogeneity variance testing for a study on olfactory capability (O'Brien and Heft 1995). Each observation contains the score of an olfactory index for a study participant and the age group (1–5), where age group ranges from younger to older. There is evidence that as people age, their sense of smell decreases. However, some people experience no decline, so there may be greater variability within the older age groups than the younger age groups.

```
data upsit;
  input agegroup smell @@;
  datalines;
1 1.381 1 1.322 1 1.162 1 1.275 1 1.381 1 1.275 1 1.322
1 1.492 1 1.322 1 1.381 1 1.162 1 1.013 1 1.322 1 1.322
1 1.275 1 1.492 1 1.322 1 1.322 1 1.492 1 1.322 1 1.381
1 1.234 1 1.162 1 1.381 1 1.381 1 1.381 1 1.322 1 1.381
1 1.322 1 1.381 1 1.275 1 1.492 1 1.275 1 1.322 1 1.275
1 1.381 1 1.234 1 1.105
2 1.234 2 1.234 2 1.381 2 1.322 2 1.492 2 1.234 2 1.381
2 1.381 2 1.492 2 1.492 2 1.275 2 1.492 2 1.381 2 1.492
2 1.322 2 1.275 2 1.275 2 1.275 2 1.322 2 1.492 2 1.381
2 1.322 2 1.492 2 1.196 2 1.322 2 1.275 2 1.234 2 1.322
2 1.098 2 1.322 2 1.381 2 1.275 2 1.492 2 1.492 2 1.381
2 1.196
3 1.381 3 1.381 3 1.492 3 1.492 3 1.492 3 1.098 3 1.492
3 1.381 3 1.234 3 1.234 3 1.129 3 1.069 3 1.234 3 1.322
3 1.275 3 1.230 3 1.234 3 1.234 3 1.322 3 1.322 3 1.381
4 1.322 4 1.381 4 1.381 4 1.322 4 1.234 4 1.234 4 1.234
4 1.381 4 1.322 4 1.275 4 1.275 4 1.492 4 1.234 4 1.098
4 1.322 4 1.129 4 0.687 4 1.322 4 1.322 4 1.234 4 1.129
4 1.492 4 0.810 4 1.234 4 1.381 4 1.040 4 1.381 4 1.381
4 1.129 4 1.492 4 1.129 4 1.098 4 1.275 4 1.322 4 1.234
4 1.196 4 1.234 4 0.585 4 0.785 4 1.275 4 1.322 4 0.712
```

```
4 0.810
5 1.322 5 1.234 5 1.381 5 1.275 5 1.275 5 1.322 5 1.162
5 0.909 5 0.502 5 1.234 5 1.322 5 1.196 5 0.859 5 1.196
5 1.381 5 1.322 5 1.234 5 1.275 5 1.162 5 1.162 5 0.585
5 1.013 5 0.960 5 0.662 5 1.129 5 0.531 5 1.162 5 0.737
5 1.098 5 1.162 5 1.040 5 0.558 5 0.960 5 1.098 5 0.884
5 1.162 5 1.098 5 0.859 5 1.275 5 1.162 5 0.785 5 0.859
;
```

The following PROC GLM invocation requests Levene's test and Welch's test for the groups defined by the AGEGROUP variable.

```
proc glm data=upsit;
  class agegroup;
  model smell = agegroup;
  means agegroup / hovtest welch;
run;
```

The usual ANOVA indicates that the olfactory index mean differs for the various age groups.

Analysis of Variance Procedure					
Dependent Variable: SMELL					
Source	DF	Anova SS	Mean Square	F Value	Pr > F
AGEGROUP	4	2.13878141	0.53469535	16.65	0.0001

Figure 7. ANOVA Results

However, Levene's test indicates that the assumption of homogeneity of variance is violated, with an  $F$  value of 6.3494 for 4 df ( $p=0.001$ ).

Analysis of Variance Procedure					
Levene's Test for Equality of SMELL Variance					
ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
AGEGROUP	4	0.0799	0.0200	6.3494	0.0001
Error	175	0.5503	0.00314		

Figure 8. Levene's Test

Using Welch's ANOVA, you would still conclude that there are differences in mean olfactory index for the various groups. The  $F$  statistic has the value 13.7208 ( $p=0.001$ ).

Analysis of Variance Procedure			
Welch's ANOVA for SMELL			
Source	DF	F Value	Pr > F
AGEGROUP	4.0000	13.7208	0.0001
Error	78.7489		

Figure 9. Welch's ANOVA

Note that unless group variances are extremely different or unless you have a very large number of groups,

the usual ANOVA is quite robust when groups are roughly the same size.

## Upgrading the MULTTEST Procedure

The MULTTEST procedure computes regular and multiplicity-adjusted  $p$ -values for a variety of statistical tests. The input comes from a multivariate ANOVA model, and the response variables can be either discrete or continuous. In the case of a continuous variable,  $p$ -values for the  $t$ -test are computed; in the case of discrete response variables,  $p$ -values for the Cochran-Armitage trend test, the permutation exact test, and the Peto test can be produced, among others.

With Release 6.12, the MULTTEST procedure has undergone substantial improvement. A very useful enhancement is the capability of accepting unadjusted  $p$ -values as input and adjusting them by using any method that doesn't require raw data. Options for the Peto test have expanded: it can now be computed via a permutation test, all relevant tables are now printed, and a new permutation-based multiplicity adjustment is available.

Several new multiplicity adjustments have been added to PROC MULTTEST, including the false discovery rate controlling method of Benjamini and Hochberg, Hochberg's step-up Bonferroni adjustment, Bonferroni approximations using exact permutation distributions, and independence-assuming approximations using exact permutation distributions.

Other new enhancements include:

- Freeman-Tukey and  $t$ -tests are weighted by the STRATA variable.
- Simulation standard errors and intermediate calculations are now included in the OUT= data set.
- The hypergeometric variance has become the default for the Cochran-Armitage test.
- Handling of missing data has improved.

As an example of the new features, the following analysis illustrates the use of PROC MULTTEST to multiplicity-adjust a set of raw  $p$ -values. In the DATA step, the unadjusted  $p$ -values from eight tests are assigned to the required variable RAW\_P in the data set RAW. In the PROC MULTTEST invocation, the option PDATA identifies the data set with the raw  $p$ -values, and the options HOLM, HOC, and FDR request the Holm, Hochberg, and Benjamini and Hochberg methods, respectively.

```
data raw;
  input test$ raw_p;
datalines;
test1 .08201
test2 .67132
test3 .00112
test4 .54333
test5 .23112
test6 .02454
test7 .01423
test8 .04222
;
proc multtest pdata=raw
  holm hoc fdr out=newvalue;
run;
proc print;
run;
```

No output from PROC MULTTEST is produced. The following listing from PROC PRINT displays the multiplicity-adjusted  $p$ -values.

OBS	TEST	RAW_P	STPBON_P	HOC_P	FDR_P
1	test1	0.08201	0.32804	0.32804	0.13122
2	test2	0.67132	1.00000	0.67132	0.67132
3	test3	0.00112	0.00896	0.00896	0.00896
4	test4	0.54333	1.00000	0.67132	0.62095
5	test5	0.23112	0.69336	0.67132	0.30816
6	test6	0.02454	0.14724	0.14724	0.06544
7	test7	0.01423	0.09961	0.09961	0.05692
8	test8	0.04222	0.21110	0.21110	0.08444

Figure 10. Adjusted  $p$ -Values

The Hochberg method controls the familywise error rate under independence, using step-up Bonferroni, while the Holm method uses step-down Bonferroni. The Benjamini and Hochberg method controls the false discovery rate and not the familywise error rate. Note that the adjusted  $p$ -values for the Hochberg method (HOC\_P) are less than or equal to those for the Holm method (STPBON\_P), and those from the Benjamini and Hochberg method (FDR\_P) are less than or equal to those for the Hochberg method.

## Enhancements to the MIXED Procedure

The MIXED procedure continues to be improved with Release 6.12. New covariance structures have been added to facilitate repeated measurements analysis, including the spatial anisotropic exponential structure and direct product structures designed for multivariate repeated measures. The former structure has the  $(i, j)$ th element equal to

$$\sigma^2 \prod_{k=1}^c \exp[-\theta_k d(i, j, k)^{p_k}]$$

where  $c$  is the number of coordinates and  $d(i, j, k)$  is the absolute distance between the  $k$ th coordinate ( $k = 1, \dots, c$ ) of the  $i$ th and  $j$ th observations in the

input data set. There are  $2c + 1$  parameters to be estimated:  $\theta_k, p_k$  ( $k = 1, \dots, c$ ), and  $\sigma^2$ .

The direct product structures are constructed by taking the Kronecker product of an unstructured matrix (modeling covariance across the multivariate observations) with an additional covariance matrix (modeling covariance across time or another factor). The upper left value in the second matrix is constrained to be equal to 1 to identify the model. Refer to Galecki (1994).

The computations for predicted values requested with the P options in the MODEL statement have changed for specifications with a REPEATED statement with the SUBJECT= option and missing dependent variables. Now, the predicted values are computed using empirical best linear unbiased prediction (EBLUP) instead of using the estimated means. This means that you can perform kriging predictions when you specify the spatial correlation structures.

In addition, many of the tables in the PROC MIXED output have been altered to improve readability and interpretation.

## LSMEANS Enhancements

Previously, the GLM procedure computed multiple comparisons of group averages only. With highly unbalanced data, this didn't necessarily produce useful estimates. With Release 6.12, the LSMEANS facility of both the GLM and MIXED procedures have been updated to provide for multiple comparisons based on a general methodology described by Hsu (1996). The implementation of this approach provides multiple comparisons for the complete set of linear models fitted with the GLM procedure as well as providing multiple comparisons for mixed models fit in the MIXED procedure with non-trivial error structures.

The LSMEANS strategy now includes a more general approach that allows the user to define alternate weighting schemes. Enhancements include options that allow you to specify the population over which the expected marginal means are to be computed. Covariates can be handled by setting them to their mean value in the analysis data set, and you can also set their values directly.

When an interaction A\*B is significant, you may want to analyze the marginal differences between the A\*B means for different values of A. You can test these so-called *simple effects* of B by "slicing" the LSMEANS of the A\*B interaction by A:

```
lsmeans A*B / slice=A;
```

This produces a familiar ANOVA table for the marginal

differences.

## Miscellaneous Enhancements

The 6.12 Release of SAS/STAT contains other enhancements. These include:

- The NPAR1WAY procedure now includes an OUTPUT statement. You can put the statistics produced into an output data set for further processing and reporting.
- The NLIN procedure now computes analytic first and second order derivatives. This feature was previously available in an experimental version of PROC NLIN called PROC TNLIN.
- The REG procedure now includes new options that allow you to output to the OUTEST= data set the standard errors of the parameter estimates, the corresponding  $t$  statistics and  $p$ -values, error degrees of freedom, and the model  $R^2$ .

## Conclusion

With Release 6.12, SAS/STAT software continues to provide current statistical methodology in a variety of application areas. For documentation of all the new features, you have two choices. *SAS/STAT Software: Changes and Enhancements in Release 6.12* describes the features new in the 6.12 Release. *SAS/STAT Software: Changes and Enhancements through Release 6.12* contains all of the features added since the two-volume *SAS/STAT User's Guide* was produced. For up-to-date information on the statistical products and development, please visit us at <http://www.sas.com/rnd/>.

## References

- Bartlett, M.S. (1937), "Properties of Sufficiency and Statistical Tests," *Proceedings of the Royal Society of London, Series A* 160, 268–282.
- Box, G.E.P. (1953), "Non-normality and Tests on Variance," *Biometrika*, 40, 318–335.
- Brown, M.B. and Forsythe, A.B. (1974), "Robust Tests for Equality of Variances," *Journal of the American Statistical Association*, 69, 364–367.
- Carey, V., Zeger, S.L., and Diggle, P. (1993), "Modelling Multivariate Binary Data with Alternating Logistic Regressions," *Biometrika*, 517–526
- Conover, W.J., Johnson, M.E., and Johnson, M.M. (1981), "A Comparative Study of Tests for Homogeneity of Variances, with Applications to the Outer

Continental Shelf Bidding Data,” *Technometrics*, 23, 351–361.

Diggle, P.J., Liang, K.-Y. and Zeger, S.L. (1994), *Analysis of Longitudinal Data*, Oxford: Oxford Science

Galecki, A.T. (1994), “General Class of Covariance Structures for Two or More Repeated Factors in Longitudinal Data Analysis,” *Communications in Statistics-Theory and Methods*. 23(11), 3105–3119.

Hsu, J.C. (1992), “The factor analytic approach to simultaneous confidence interval for multiple comparisons with the best,” *Journal of Computational Statistics and Graphics*, 1, 151–168.

Levene, H. (1960), “Robust Tests for the Equality of Variance,” in *Contributions to Probability and Statistics*, ed. I. Olkin, Palo Alto, CA: Stanford University Press, 278–292.

Liang, K.-Y. and Zeger, S.L. (1986), “Longitudinal Data Analysis Using Generalized Linear Models” *Biometrika*, 13–22

Lipsitz, S.H., Fitzmaurice, G.M., Orav, E.J., and Laird, N.M. (1994), “Performance of Generalized Estimating Equations in Practical Situations,” *Biometrics* 50, 270–278

O’Brien, R.G. (1979), “A General ANOVA Method for Robust Tests of Additive Models for Variances,” *Journal of the American Statistical Association*, 74, 877–880.

O’Brien, R.G. and Heft, M.W. (1995), “New Discrimination Indexes and Models for Studying Sensory Functioning in Aging,” *Journal of Applied Statistics*, 22, 9–27.

Thall, P.F. and Vail, S.C. (1990), “Some Covariance Models for Longitudinal Count Data with Overdispersion,” *Biometrics*, 657–671

Ware, J.H., Dockery, Spiro A. III, Speizer, F.E., and Ferris, B.G., Jr. (1984), “Passive Smoking, Gas Cooking, and Respiratory Health of Children Living in Six Cities,” *American Review of Respiratory Diseases*, 129, 366–374.

Welch, B.L. (1951), “On the Comparison of Several Mean Values: An Alternative Approach,” *Biometrika*, 38, 330–336.

Zeger, S.L. and Liang, K.-Y. (1986), “Longitudinal Data Analysis for Discrete and Continuous Outcomes,” *Biometrics*, 121–130

SAS and SAS/STAT are registered trademarks of SAS Institute Inc. in the USA and in other countries. ® indicates USA registration.