Research Report

Loglinear Smoothing: An Alternative Numerical Approach Using SAS

Tim Moses

Alina A. von Davier

Jodi Casabianca

Research & Development

July 2004
RR-04-27

Loglinear Smoothing: An Alternative Numerical Approach Using SAS

Tim Moses, Alina A. von Davier, and Jodi Casabianca

ETS, Princeton, NJ

July 2004

Abstract

The purpose of this report is to demonstrate loglinear smoothing using SAS PROC GENMOD. The results from four published examples, which include the smoothing of a) univariate distributions, b) bivariate distributions, c) distributions with teeth, and d) bivariate distributions with structural zeros, are reproduced to show the flexibility of the SAS procedure. Comparisons of graphical displays and likelihood ratio statistics show that the SAS procedure yields results identical to the published results. SAS PROC GENMOD provides an alternative approach to smoothing that is readily available for researchers and graduate students in educational measurement, as well as researchers from other scientific fields.

Key words: Loglinear smoothing, SAS PROC GENMOD, test equating

Acknowledgements

The authors thank Paul Holland, Skip Livingston, Dan Eignor, and Shelby Haberman for their valuable comments and suggestions on the previous version of this paper. The opinions expressed herein are those of the authors and not of ETS.

Table of Contents

Smoothing represents a set of statistical procedures that aim to replace a discrete empirical dataset with a discrete dataset that preserves some features of the observed data without the irregularities that are attributable to sampling. The type of smoothing described in this paper involves the fitting of loglinear models to discrete test score distributions (Holland & Thayer, 1987; 2000). These loglinear models can preserve a variety of different features in observed data with a relatively small number of parameters. Loglinear smoothing has numerous applications, one of which is in educational assessment as a preliminary step in the equating of scores on different forms of a test. More precisely, loglinear smoothing is first applied and then the smoothed results can be used with nonlinear equating procedures such as the traditional equipercentile procedure or the kernel procedure (von Davier, Holland, & Thayer, 2004; Hanson, 1996; Holland & Thayer, 1989; Rosenbaum & Thayer, 1987). Loglinear smoothing can also be used, however, prior to the actual equating in situations where available sample sizes are small (Livingston, 1993).

While there are software routines that carry out loglinear smoothing, they are usually implemented as part of larger software packages that perform test equating (e.g., the operational software used by ETS; the equating software from Iowa Testing Programs, http://www.uiowa.edu/~itp/pages/SWEQUATING.SHTML). Because they are embedded within a larger system of routines oriented towards very specific purposes, these software packages are either neither generally available nor flexible enough for general smoothing purposes. Therefore, it is important to have alternative tools for smoothing that are readily available to researchers and graduate students.

The purpose of this report is to illustrate how SAS PROC GENMOD can be used to implement the same loglinear smoothing procedures that are described in previously published descriptions of loglinear smoothing (von Davier et al., 2004; Holland & Thayer, 1987, 2000). The first section gives a brief overview of the relevant models and the model-fitting process that accomplishes loglinear smoothing. Then the model-fitting process is described in terms of how it can be implemented with PROC GENMOD. The results of four published examples from Holland and Thayer (2000) and von Davier et al. (2004) are reproduced with PROC GENMOD, including the smoothing of univariate distributions, bivariate distributions, distributions with "teeth" (a regular pattern of cells with frequencies that are much lower than those of neighboring cells, usually due to the use of rounded formula scores), and bivariate distributions with

structural zeros (impossible score combinations in the score probabilities that can arise in the joint distribution of a total test and an internal anchor, so that the total test score can never be less than the score on the internal anchor test). To emphasize the agreement of the results of the published examples with the results from SAS PROC GENMOD, references are made to results in the published examples.

Evaluations of the fit of the models themselves are discussed at length in the previously published papers and are, therefore, only briefly discussed in this report.

Loglinear Smoothing Models

Assume we have a random variable X that defines test form X (we use the same notation for a test form and a random variable) with possible values $x_0,\ldots,x_J$, or x_j, with $j = 0,\ldots,J$ (the possible score values), and a corresponding vector of observed score frequencies $n = (n_0,\ldots,n_J)^t$ that sum to the total sample size, N. Under some distributional assumptions about n, like multinomial or Poisson distributional assumptions, the vector of the population score probabilities $p = (p_0,\ldots,p_J)^t$ is said to satisfy a loglinear model if

$$\log_e(p_j) = \alpha + u_j + \mathbf{b_j}\boldsymbol{\beta}$$

where the $\{p_j\}$ are assumed to be positive and sum to one, $\mathbf{b_j}$ is a row vector of constants referred to as score functions throughout this text (e.g., x_j^1, x_j^2, x_j^3), $\boldsymbol{\beta}$ is a vector of free parameters, u_j is a known constant that specifies the distribution of the $\{p_j\}$ when the vector $\boldsymbol{\beta}$ is set to zero, and α is a normalizing constant that insures that the probabilities sum to one.

Under different choices of $\mathbf{u}$, $\mathbf{b}$, or $\boldsymbol{\beta}$, the loglinear model becomes equivalent to the discrete uniform distribution ($\mathbf{u} = 0$, $\boldsymbol{\beta} = 0$) or the binomial distribution (see Holland & Thayer, 1987, 2000, for details).

Loglinear models are a class of the exponential families of discrete distributions, which can be described in terms of their sample moments. As in Holland and Thayer (1987, 2000), we will make use of this property and the fact that u_j are known constants. Therefore, in this paper, the loglinear model used to fit a univariate distribution is

$$\log_e(p_j) = \alpha + \sum_{i=1}^{I} \beta_i (x_j)^i , \qquad (1)$$

where the u_j are set to zero. When the data are test score data, the terms in this model can be defined as follows: b_j is a vector of score functions; $(x_j)^i$ are the score functions; and β_i are the I free parameters to be estimated in the model-fitting process.

The value of I determines the number of moments of the actual test score distribution that are preserved in the fitted distribution. If $I = 1$, then the fitted distribution preserves the first moment (the mean) of the observed distribution. If $I = 4$, then the fitted distribution preserves the first, second, third, and fourth moments (mean, variance, skewness, and kurtosis) of the observed distribution.

The model in (1) can be extended to fit the bivariate distribution of the scores of two tests (call them X and Y):

$$\log_e(p_{jk}) = \alpha + \sum_{i=1}^{I} \beta_{xi}(x_j)^i + \sum_{h=1}^{H} \beta_{yh}(y_k)^h + \sum_{g=1}^{G} \sum_{f=1}^{F} \beta_{gf}(x_j)^g (y_k)^f , \qquad (2)$$

where p_{jk} is the joint score probability of the score (x_j, y_k) (score x_j on test X and score y_k on test Y). The fitting of Model (2) produces a fitted bivariate distribution that preserves I moments in the marginal (univariate) distribution of X, H moments in the marginal (univariate) distribution of Y, and a number of cross-moments ($G \le I$, $F \le H$) in the bivariate X-Y distribution. Model (2) is also appropriate for the smoothing of bivariate distributions with impossible X-Y score combinations, or "structural zeros." Distributions with structural zeros arise when X and Y represent the scores of a total test and an internal anchor, so that the total test score can never be less than the score on the internal anchor test. The fitting of distributions with structural zeros is described in the fourth example of this report and involves models of the form in (2) with appropriate specification of the data layout.

Another extension of (1) incorporates indicator functions. These indicator functions allow for the fitting of both the full univariate distribution and a subset of the distribution (e.g., "teeth" or lumps at different score points). One example of such a model is:

$$\log_e(p_j) = \alpha + \beta_1(x_j)^1 + \beta_2(x_j)^2 + \beta_3 I_s(j) + \beta_4(x_j)^1 I_s(j) , \qquad (3)$$

where the indicator function $I_S(j) = 1$ if j belongs to a defined subset, S, of all js and $I_S(j) = 0$ otherwise. S denotes the set of the score points where the frequencies are systematically lower or higher than most of the test frequencies. Model (3) will preserve the mean and variance of the

total distribution of X (β_1 and β_2), the total frequency in the cells denoted by S (β_3), and the mean of the cell values for the cells in S (β_4).

Fitting Loglinear Smoothing Models

Under the assumption that the vector of the frequencies is multinomial, the estimation of the free parameters (β_i) proceeds by maximizing the following log-likelihood function:

$$L = \sum_j n_j \log_e(p_j), \tag{4}$$

where n_j and p_j are the observed frequencies and the population score probabilities in the jth cell, respectively (Holland & Thayer, 1987, 2000).

The maximization can be accomplished through the use of the Newton-Raphson algorithm (Holland & Thayer, 1987, p. 11). Holland and Thayer specify two criteria for the convergence solution from the algorithm. The first criterion involves the maximization of the log-likelihood function, which happens when the relative change in the log-likelihood is less than some specified value. The second criterion involves the satisfaction of the likelihood equation for all of the estimated parameters (β), meaning that the relative error in each fitted moment must be less than some specified value. At convergence both criteria should be met.

To add stability to the algorithm, it has been suggested that the score functions be transformed so that they sum to zero and their squares sum to one (Holland & Thayer, 1987, 2000; Rosenbaum & Thayer, 1987). Holland and Thayer also suggest specific starting values. The suggested starting values for the parameter estimates are based on converting the observed frequencies into a smoother form with nonzero frequencies at all score points and then computing a function of these converted frequencies and the score functions.

Fitting Loglinear Smoothing Models With SAS PROC GENMOD

The examples that are presented next will demonstrate how SAS PROC GENMOD can be used to fit Models (1), (2), and (3). While modeling procedures based on Poisson and multinomial distributional assumptions produce the same maximum likelihood estimates (Bishop, Fienberg, & Holland, 1975; Fisher, 1922; Haberman, 1974), the Poisson-based modeling procedures in SAS are much more flexible than the multinomial-based modeling procedures. Therefore, for our purposes, the discrete observed frequencies $\{n_j\}$ are assumed to be

independently Poisson-distributed with parameters Np_j. In this case, the log-likelihood to be maximized is:

$$L = \sum_j \log_e \left[\frac{(Np_j)^{n_j} e^{(-Np_j)}}{n_j!} \right]. \quad (5)$$

(SAS Institute, 2002, pp. 1534–1535).

Maximization is accomplished through a ridge-stabilized Newton-Raphson algorithm (SAS Institute, 2002, p. 1536). This algorithm converges when the change in parameter estimates between iterations is less than some specified value. After convergence is determined for the parameters, the algorithm checks the convergence of the inverse matrix of second derivatives relative to the log-likelihood function, which is another indication of whether or not the likelihood function has been maximized. The user has the option of specifying the values for both of these convergence criteria. The default initial parameter values are weighted least squares estimates based on using the observed frequencies for the initial mean estimates. The user has the option of specifying these initial values. For the examples discussed in this report, the default convergence criteria, initial parameter values from SAS, and the unscaled score functions yield acceptable solutions. Appendix A describes the general form of SAS PROC GENMOD.

Evaluating Model Fit With SAS PROC GENMOD

The examples will focus on comparing likelihood ratio chi-square statistics, the statistics most often reported in the published sources, to deviances from SAS outputs. The likelihood ratio chi-square statistic is defined as:

$$G^2 = 2\sum_j n_j \log_e \left(\frac{n_j}{\hat{p}_j N}\right), \quad (6)$$

where $\hat{p}_j$ is the fitted value of p_j under the model. This measure is used in Holland and Thayer (1987, 2000). It should be equal to one of the measures of fit produced in the SAS results, the deviance:

$$\text{Deviance} = 2\sum_j [n_j \log_e \left(\frac{n_j}{\hat{p}_j N}\right) - (n_j - \hat{p}_j N)]. \quad (7)$$

The difference between the likelihood ratio chi-square and the deviance, $2\sum_j [n_j - \hat{p}_j N]$, is defined as twice the maximum achievable likelihood (SAS, 2002, p. 1537) and is zero for all cases where there is complete convergence.

Example 1: Fitting Univariate Distributions

An example from von Davier et al. (2004, pp. 100–104) is used to illustrate how SAS PROC GENMOD smooths univariate distributions. Univariate smoothing models are of the form in (1). von Davier et al. selected a 2-moment model for test *X* and a 3-moment model for test *Y*. The SAS code given in Appendix B illustrates how to enter their data (which are already in a frequency form) and create the relevant score functions (which are the scores in their original, squared, and cubed form for preserving the first, second, and third moments).

von Davier et al. (2004) report that the likelihood ratio chi-square statistic (computed as in [6]) is 18.35 on 18 degrees of freedom for the fit of *X* and 20.24 on 17 degrees of freedom for the fit of *Y*. Appendix C shows the commands and partial results from SAS PROC GENMOD—first for the fit of *X*, then for the fit of *Y*. These results show that the degrees of freedom (DF) match the published results. In addition, the deviances from SAS PROC GENMOD also match the likelihood ratio chi-square statistics reported by von Davier et al. (2004).

Appendix D shows the observed and fitted frequencies from SAS PROC GENMOD, along with the fitted frequencies reported in Table 7.2 of von Davier et al. (2004, p. 102). The observed-fitted plots that correspond to Figures 7.1 and 7.2 in von Davier et al. (2004, pp. 103–104) are shown in Figures 1 and 2.

The moments from the fitted and observed distributions can also be compared. The 2-moment fit of *X* should result in observed and fitted distributions with equal means and variances. The 3-moment fit of *Y* should result in observed and fitted distributions with equal means, variances, and skewness. The SAS commands of Appendices E and F demonstrate how a dataset of individual observations can be obtained from the observed or fitted score frequencies. The fitted and observed moments are compared for *X* (Appendix E) and for *Y* (Appendix F). The results agree with the published results, except that the statistic called "kurtosis" in the SAS output is actually the standardized fourth moment (i.e., a deviation from zero rather than a deviation from 3) rather than the actual kurtosis that is reported in von Davier et al. (2004).

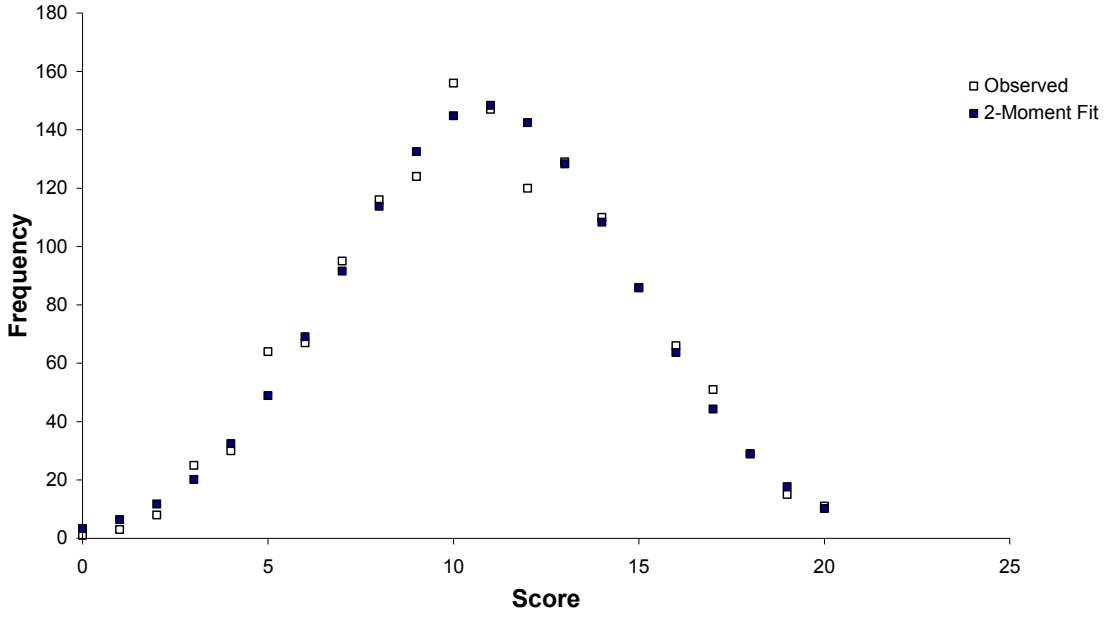

Figure 1. **Example 1: Fitting univariate distributions for *X* and *Y* (2-moment fit).**

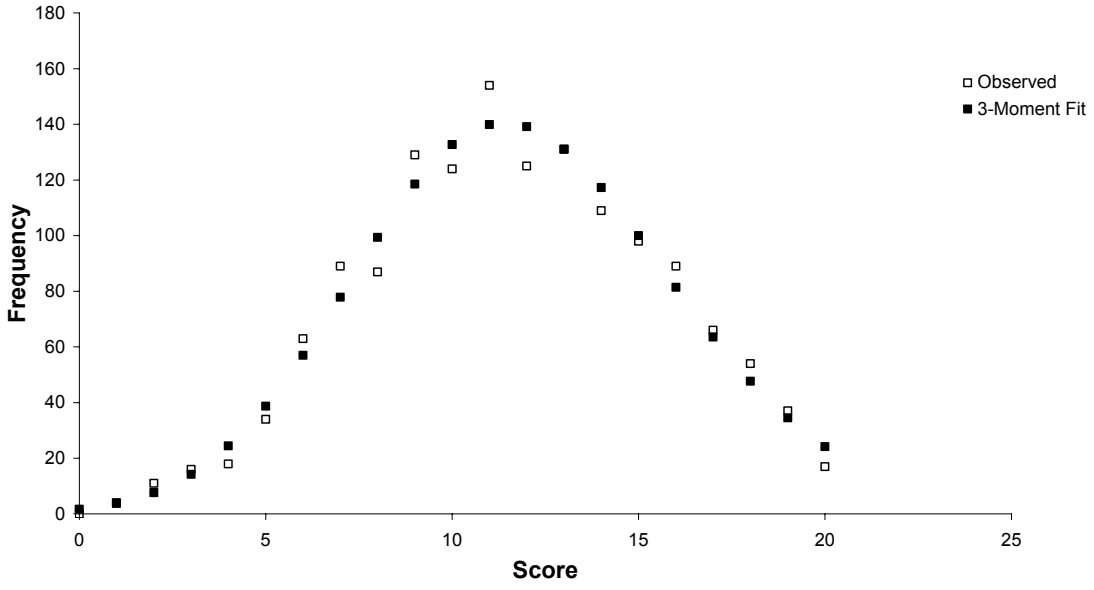

Figure 2. **Example 1: Fitting univariate distributions for *X* and *Y* (3-moment fit).**

Example 2: Fitting Bivariate Distributions

To illustrate the fitting of bivariate data of the form described in (2), another example from von Davier et al. (2004, pp. 114–120) is used. This modeling calls for a fitting of the bivariate frequencies based on two tests, *X* and *Y* (*Y* is different from the *Y* used in Example 1). von Davier et al. fit a model that preserved 3 moments on *X*, 3 moments on *Y*, and 1 cross-moment for *X* and *Y*. The reported likelihood ratio of the chi-square statistic is 242.73 on 433 degrees of freedom (note that this statistic is not chi-square distributed due to the sparseness of the data).

The SAS code in Appendix G indicates how to enter the bivariate data (where frequency is now the total number of cases with a particular score on *X* and a particular score on *Y*), and create the score functions used for the desired model. The code and results of SAS PROC GENMOD are presented in Appendix H.

von Davier et al. (2004) present the marginal fitted frequencies of *X* (summed over *Y*) and *Y* (summed over *X*) in Table 8.3 (p. 118). The corresponding marginal fitted frequencies from SAS PROC GENMOD are shown in Appendix I, along with von Davier et al.'s fitted frequencies. Table 1 shows the fitted bivariate frequencies from SAS PROC GENMOD. Table 2 shows the fitted bivariate frequencies from von Davier et al. (from Table 8.4, p. 120).

Table 1

Example 2: The Fitted X-Y Frequencies From SAS PROC GENMOD

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	0.3	0.4	0.4	0.4	0.4	0.2	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.4	0.6	0.8	1.0	0.9	0.7	0.4	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.4	0.8	1.3	1.7	1.9	1.7	1.2	0.7	0.4	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.4	0.9	1.7	2.6	3.3	3.4	2.8	2.0	1.1	0.5	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.3	0.9	1.9	3.3	4.8	5.6	5.5	4.4	2.9	1.6	0.7	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.2	0.7	1.7	3.5	5.7	7.8	8.8	8.1	6.2	3.9	2.0	0.9	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.1	0.5	1.3	3.0	5.8	9.1	11.7	12.5	11.0	8.0	4.8	2.4	1.0	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0
7	0.1	0.3	0.8	2.2	4.9	8.8	13.1	16.1	16.4	13.7	9.4	5.4	2.5	1.0	0.3	0.1	0.0	0.0	0.0	0.0	0.0
8	0.0	0.1	0.4	1.4	3.4	7.2	12.3	17.4	20.4	19.6	15.6	10.2	5.5	2.5	0.9	0.3	0.1	0.0	0.0	0.0	0.0

(Table continues)

Table 1 (continued)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
9	0.0	0.0	0.2	0.7	2.0	4.9	9.7	15.8	21.3	23.7	21.7	16.4	10.2	5.3	2.2	0.8	0.2	0.1	0.0	0.0	0.0
10	0.0	0.0	0.1	0.3	1.0	2.8	6.4	12.1	18.8	24.0	25.4	22.1	15.9	9.4	4.6	1.9	0.6	0.2	0.0	0.0	0.0
11	0.0	0.0	0.0	0.1	0.4	1.4	3.6	7.8	14.0	20.6	25.1	25.1	20.8	14.2	8.1	3.8	1.4	0.5	0.1	0.0	0.0
12	0.0	0.0	0.0	0.0	0.2	0.6	1.7	4.3	8.8	14.9	20.9	24.2	23.1	18.2	11.9	6.4	2.8	1.0	0.3	0.1	0.0
13	0.0	0.0	0.0	0.0	0.0	0.2	0.7	2.0	4.7	9.2	14.8	19.8	21.7	19.7	14.8	9.2	4.7	2.0	0.7	0.2	0.0
14	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.8	2.1	4.8	8.9	13.7	17.4	18.2	15.7	11.2	6.6	3.2	1.3	0.4	0.1
15	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.8	2.1	4.6	8.1	11.8	14.2	14.2	11.7	8.0	4.5	2.1	0.8	0.3
16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.8	2.0	4.1	6.9	9.5	10.9	10.4	8.1	5.3	2.8	1.3	0.5
17	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.7	1.8	3.4	5.4	7.2	7.9	7.1	5.3	3.3	1.7	0.7
18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.6	1.4	2.7	4.1	5.1	5.3	4.6	3.3	1.9	0.9
19	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.5	1.1	2.0	2.9	3.4	3.4	2.8	1.9	1.1
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.4	0.8	1.4	1.9	2.2	2.1	1.6	1.0

Table 2

Example 2: The Fitted X-Y Frequencies From von Davier, Holland, and Thayer

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	0.2	0.3	0.4	0.4	0.3	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.3	0.6	0.8	0.9	0.9	0.6	0.4	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.4	0.8	1.3	1.7	1.8	1.6	1.2	0.7	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.4	0.9	1.7	2.6	3.2	3.3	2.8	1.9	1.1	0.5	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.3	0.9	1.9	3.3	4.7	5.6	5.4	4.4	2.9	1.5	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.2	0.7	1.7	3.4	5.7	7.8	8.7	8.1	6.2	3.9	2.0	0.8	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.1	0.4	1.3	3.0	5.7	9.0	11.7	12.5	11.0	7.9	4.7	2.3	0.9	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0
7	0.0	0.2	0.8	2.2	4.8	8.8	13.1	16.1	16.3	13.6	9.4	5.3	2.5	0.9	0.3	0.1	0.0	0.0	0.0	0.0	0.0
8	0.0	0.1	0.4	1.3	3.4	7.1	12.3	17.4	20.3	19.6	15.5	10.2	5.5	2.4	0.9	0.2	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.2	0.7	2.0	4.9	9.7	15.8	21.3	23.6	21.6	16.3	10.2	5.2	2.2	0.7	0.2	0.0	0.0	0.0	0.0
10	0.0	0.0	0.0	0.3	1.0	2.8	6.4	12.1	18.9	24.0	25.3	22.0	15.8	9.4	4.6	1.8	0.6	0.1	0.0	0.0	0.0

(Table continues)

Table 2 (continued)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
11	0.0	0.0	0.0	0.1	0.4	1.3	3.6	7.8	13.9	20.6	25.0	25.1	20.8	14.2	8.0	3.7	1.4	0.4	0.1	0.0	0.0
12	0.0	0.0	0.0	0.0	0.1	0.5	1.7	4.2	8.7	14.9	20.9	24.2	23.1	18.2	11.8	6.3	2.8	1.0	0.3	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.2	0.6	1.9	4.6	9.1	14.8	19.7	21.7	19.7	14.8	9.1	4.7	1.9	0.7	0.2	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.7	2.1	4.7	8.9	13.6	17.3	18.1	15.7	11.2	6.6	3.2	1.3	0.4	0.1
15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.8	2.1	4.5	8.0	11.8	14.2	14.2	11.6	7.9	4.4	2.0	0.8	0.2
16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.8	1.9	4.0	6.8	9.5	10.9	10.3	8.1	5.2	2.8	1.2	0.4
17	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.7	1.7	3.4	5.4	7.2	7.8	7.1	5.3	3.2	1.6	0.7
18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.6	1.4	2.6	4.0	5.1	5.3	4.6	3.2	1.9	0.9
19	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.5	1.1	1.9	2.8	3.4	3.4	2.8	1.9	1.0
20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.4	0.8	1.3	1.9	2.1	2.0	1.6	1.0

Note. From von Davier et al. (2004, p. 120).

The final part of Example 2 compares the moments and cross-moment of the observed and fitted distributions. To accomplish this, two large datasets are constructed, one based on the observed bivariate frequencies and the other based on the fitted bivariate frequencies. Then the moments and cross-moment of X and Y in the datasets are compared. Appendix J shows that the means, variances, and skewness of X and Y are preserved in the fitted distribution. Appendix K shows that the X-Y covariance is also preserved. Since the variances of X and Y are matched, the X-Y correlation is also preserved in the fitted distribution.

Example 3: Fitting Distributions With Teeth

SAS PROC GENMOD can also fit separate parts of distributions, which can be useful when distributions have naturally occurring teeth or when negative formula scores get rounded to zero, creating a "lump at zero." The form of the model for this example is a combination of (2) because the data are bivariate, and (3) because of the teeth. In this example, parts of Example 3 in Holland and Thayer (2000, pp. 157–169) will be reproduced.

The suggested progression of the modeling is to work from the "outside" (marginal distributions) to the "inside" (bivariate distribution). The first model fits the marginal distributions for the two tests in the bivariate data. This model does not include any cross-product terms to match the cross-moments, and therefore allows the two tests (X and Y) to be independent.

Holland and Thayer's (2000) first model fits the first five moments of X's overall distribution, the frequency of the teeth in X, the first five moments of the teethed distribution in part of X, and the first five moments in Y. This model has 16 total parameters, and results in a likelihood ratio chi-square statistic of 24,187.26 on 534 degrees of freedom. This large statistic indicates that the model does not fit the data very well, and might be improved upon by accounting for the dependence of X and Y. The SAS code that produces the score functions is shown in Appendix L and then the code and relevant results from SAS PROC GENMOD are shown in Appendix M.

Holland and Thayer (2000) consider a second and third model, so that the first model considered is nested within the second and third models, which contain all of the terms as used in the first model plus additional terms for the cross-moments of X and Y. Because the third model contains all of the terms of the second model, the second model is also nested within the third model. For evaluating the second and third model, a large reduction in the likelihood ratio chi-square relative to the reduction in the degrees of freedom as compared to the first model indicates a superior fit. The second model adds one term to the first model to fit the cross-moment of X and Y. This results in a likelihood-ratio chi-square statistic of 577.72 on 533 degrees of freedom, a large decrease in the likelihood-ratio chi-square statistic (23,609.54) relative to the decrease in degrees of freedom (1). The third model adds three terms to the second model to fit the X^2–Y cross-moment, the X–Y^2 cross-moment, and the X^2–Y^2 cross-moment. This 20 parameter model has a likelihood-ratio chi-square statistic of 435.48 on 530 degrees of freedom, also a large decrease in the likelihood-ratio chi-square statistic (23,751.78 when compared to the first model, 142.24 when compared to the second model) relative to the decrease in degrees of freedom (4 when compared to the first model; 3 when compared to the second model). See Appendixes N and O.

The evaluations and comparisons of the three models would need to be more extensive if the purpose were to select one of them as the "best." Additional fit indexes would be useful for comparing the relative fits of the three models, including Akaike Information Criterion, Consistent Akaike Information Criterion, Freeman-Tukey residuals for the marginal distributions, and plots of the observed and the fitted conditional means and conditional variances (see von Davier et al., 2004). These additional fit indexes are mentioned but not

reported here because the purpose of this report is simply to show the agreement in the results of SAS PROC GENMOD with Holland and Thayer's (2000) results.

The SAS code and output that fits these two models are shown in Appendixes N and O. While the results of these two models agree with Holland and Thayer's (2000), it is interesting to note that SAS reports convergence problems for the second model. Messages about the Mean Parameter being out of range are supposed to alert the user that SAS PROC GENMOD was unable to get the fitted frequencies to conform to the restrictions of the Poisson model and sum to *N*. The agreement of the results of SAS PROC GENMOD and the results of Holland and Thayer for this second model and the absence of convergence problems for the more complex third model suggest that the problem in SAS PROC GENMOD's results is not that there is nonconvergence, but that the convergence criterion used does not always recognize that the likelihood function has been maximized and that the converged solution has been achieved.

One of the ways Holland and Thayer (2000) report the results of these three model fits is with figures of the actual and fitted frequencies of the marginal *X* and *Y* distributions. Figures 3 and 4 show the observed and fitted marginal frequencies from SAS PROC GENMOD. These plots correspond to Figures 7 and 8 of Holland and Thayer (p. 162).

Example 4: Fitting Bivariate Distributions With Structural Zeros

Holland and Thayer (2000, pp. 178–181) describe how to smooth data that make up half of an array. For their 5th smoothing example, they describe a data set of *X* and *Y* scores with a structure where *X* cannot be greater than *Y*. The frequencies of the dataset are shown in Table 3, where the shaded, zero frequencies are impossible combinations of *X* and *Y*.

Table 3

Example 4: The Dataset for the Structural Zeros Problem

Scores	*Y = 1*	*Y = 2*	*Y = 3*	*Y = 4*	*Y = 5*
X = 1	8	15	12	23	11
X = 2	0	1	4	10	9
X = 3	0	0	4	4	6
X = 4	0	0	0	5	4
X = 5	0	0	0	0	5

Note. The shaded area represents impossible *X*-*Y* combinations.

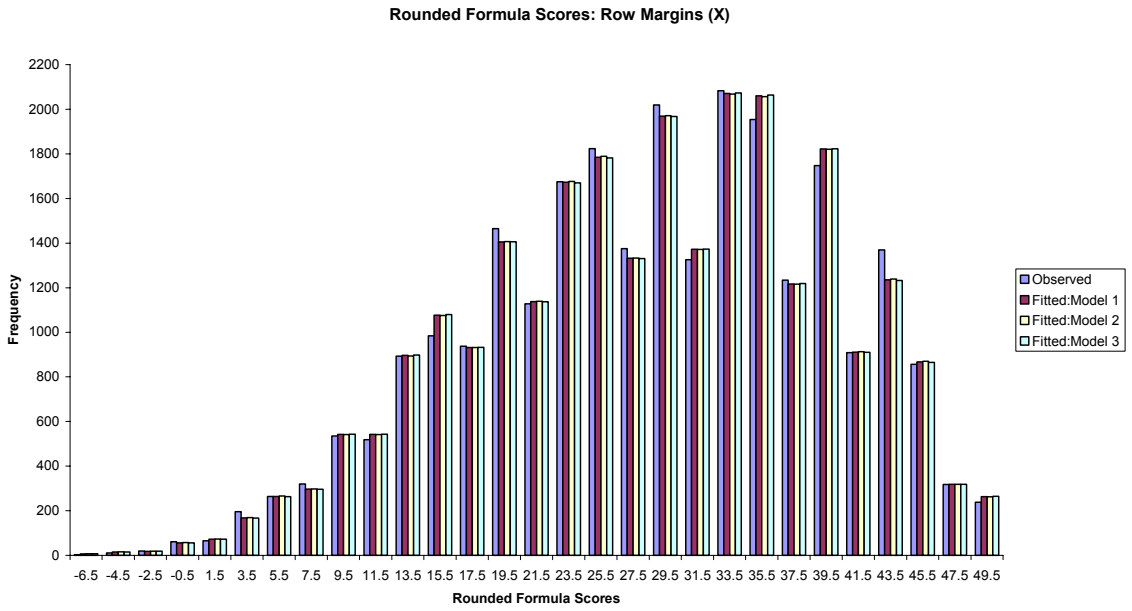

Figure 3. **Example 3: Observed and fitted marginal frequencies plots based on Models 1–3 for row margins (X).**

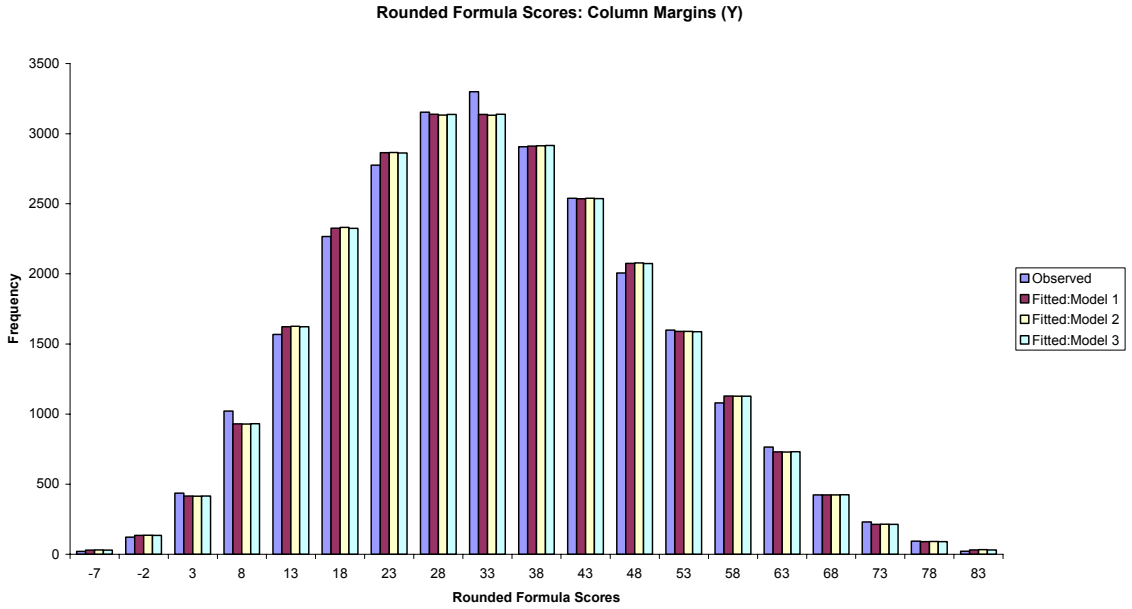

Figure 4. **Example 3: Observed and fitted marginal frequencies plots based on Models 1–3 for column margins (Y).**

Fitting bivariate distributions with structural zeros with SAS PROC GENMOD requires only one additional step in the definition of the dataset. The model-fitting process then proceeds with fitting bivariate frequencies with models of the form in (2). The SAS code in Appendix P shows how the entire dataset is entered and then reduced to the array of interest.

Holland and Thayer (2000) fit two models to these data. The first model fits the first two moments of *X* and *Y*. The likelihood ratio chi-square statistic for this model was 13.45 on 10 degrees of freedom. The corresponding code and output from SAS PROC GENMOD is shown in Appendix Q. The second model adds one *X*-*Y* cross-moment term. The result is a likelihood ratio chi-square statistic of 10.13 on 9 degrees of freedom. For the second model, Holland and Thayer note that the estimated parameter for the *X*-*Y* cross moment was 0.2614. Appendix R shows the code and output, including the parameter estimate that verifies the agreement with Holland and Thayer's results. Finally, the fitted frequencies for the two models are shown in Appendix S, along with Holland and Thayer's results (from Tables 10 and 11, p. 181).

Conclusions

This report illustrates the use of SAS PROC GENMOD for estimating loglinear models. This approach represents an alternative to the existent software routines that are embedded within larger software packages. In this form the smoothing methodology can be made more readily available for researchers and graduate students in educational measurement, as well as from other scientific fields. The results of this report suggest that SAS PROC GENMOD can be used for many applied smoothing problems.

An important issue that will be addressed in a follow-up report is the identification of situations where SAS PROC GENMOD will not provide satisfactory results. There are datasets with characteristics and models with specifications that SAS PROC GENMOD is not able to handle. For these modeling situations, it is important to understand not only why the models do not converge, but also the extent to which previously proposed strategies for improving the stability of the algorithm will help. Again, the strategies that have been proposed involve centering and/or rescaling of the score functions, adjustments to the convergence criteria, and specific choices of starting values for the parameter estimates. Also important in this discussion is the issue of convergence criteria, as the criterion used within SAS PROC GENMOD has been described as "indirect" relative to alternative convergence criteria that might be used for loglinear smoothing (Holland & Thayer, 2000, p. 147).

So far, the results only cover the estimation of the score probabilities. The demonstrations in this paper could be expanded so that SAS PROC GENMOD could provide other measures of model fit (e.g., AIC, CAIC, Freeman-Tukey residuals, plots of conditional means, variances, and skewness). It will also be important to demonstrate how to generate C-matrices, the factorization matrices of the covariance matrix of the estimated frequencies from the loglinear model (von Davier et al., 2004; Holland & Thayer, 1987, 2000). The C-matrices cannot be directly requested in SAS PROC GENMOD and will instead need to be produced from the SAS PROC GENMOD fitted frequency output by making use of the library of matrix operations in SAS IML.

References

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis.* Cambridge, MA: MIT Press.

von Davier, A. A., Holland, P. W., & Thayer, D. T. (2004). *The kernel method of test equating.* New York: Springer-Verlag.

Fisher, R. A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of *Journal of the Royal Statistical Society, 85,* 87–94.

Haberman, S. J. (1974). *The analysis of frequency data*. Chicago: University of Chicago Press.

Hanson, B. A. (1996). Testing for differences in test score distributions using log-linear models. *Applied Measurement in Education, 9,* 305–321.

Holland, P. W., & Thayer, D. T. (1987). *Notes on the use of log-linear models for fitting discrete probability distributions* (Technical Report 87-79). Princeton, NJ: ETS.

Holland, P. W., & Thayer, D. T. (1989). *The kernel method of equating score distributions* (ETS RR-89-07), Princeton, NJ: ETS.

Holland, P. W., & Thayer, D. T. (2000). Univariate and bivariate loglinear models for discrete test score distributions. *Journal of Educational and Behavioral Statistics, 25,* 133–183.

Livingston, S. (1993). Small-sample equatings with log-linear smoothing. *Journal of Educational Measurement, 30,* 23–39.

Rosenbaum, P. R., & Thayer, D. (1987). Smoothing the joint and marginal distributions of scored two-way contingency tables in test equating. *British Journal of Mathematical and Statistical Psychology, 40,* 43–49.

SAS Institute. (2002). The GENMOD procedure, Version 9 [Computer software manual]. Cary, NC: Author.

List of Appendixes

Appendix A

The Form of SAS PROC GENMOD That Implements Loglinear Smoothing

The form of the SAS code for PROC GENMOD is:

```
proc genmod data=DATA;
output out=RESULTS p=fitted;
model freq=score score2 score3 …. / link=log dist=p type3;
run;
```

The first line invokes the GENMOD procedure for a desired dataset called DATA. The second line asks for the fitted frequencies (along with observed frequencies and predictors) to be written to a dataset (called RESULTS) that can be used to evaluate the results. The third line specifies the model to be fit, where the observed frequencies are to be related to the score functions. An intercept that corresponds to α and a scale parameter that is constrained to 1 are included by default in the model. After the required backslash (/), the appropriate link function for linking the frequencies to the score functions (log) and distribution (p = Poisson) are requested. The user can request a sequential modeling process (type 1), where the score functions are entered in one at a time, or a nonsequential modeling process (type 3). If it is of interest, an α-only model can be fit by specifying the model in the third line as:

```
model freq= / link=log dist=p type3;
```

The fitted frequencies from an α-only model are equal to N divided by the total number of score levels so that the smoothed test score distribution will be uniform. The fourth line asks SAS to run the desired model.

Appendix B

Example 1—Entering the Dataset and Defining the Score Functions

```
data llin;
input score xfreq yfreq;
cards;
0 1 0
1 3 4
2 8 11
3 25 16
4 30 18
5 64 34
6 67 63
7 95 89
8 116 87
9 124 129
10 156 124
11 147 154
12 120 125
13 129 131
14 110 109
15 86 98
16 66 89
17 51 66
18 29 54
19 15 37
20 11 17
;



data llin;set llin;
/*Defining the score functions for preserving the 2nd and
3rd moments.*/
score2=score**2;
score3=score**3;
```

Appendix C

Example 1—Fitting the Two Univariate Models

```
/*Here is the 2-moment fit for X.*/
proc genmod data=llin;
output out=llinoutx p=xpp;
model xfreq=score score2 /link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.LLIN
Distribution	Poisson
Link Function	Log
Dependent Variable	xfreq
Observations Used	21

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	18	18.3525	1.0196
Scaled Deviance	18	18.3525	1.0196
Pearson Chi-Square	18	17.4615	0.9701
Scaled Pearson X2	18	17.4615	0.9701
Log Likelihood		5134.4977	

Algorithm converged.

```
/*Here is the 3-moment fit for Y.*/
proc genmod data=llin;
output out=llinouty p=ypp;
model yfreq=score score2 score3 /link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.LLIN
Distribution	Poisson
Link Function	Log
Dependent Variable	yfreq
Observations Used	21

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	17	20.2373	1.1904
Scaled Deviance	17	20.2373	1.1904
Pearson Chi-Square	17	18.3453	1.0791
Scaled Pearson X2	17	18.3453	1.0791
Log Likelihood		5108.5027	

Algorithm converged.

Appendix D

Example 1—Observed and Fitted Frequencies

Observed and fitted frequencies are from SAS PROC GENMOD and from von Davier et al. (2004, p. 102).

```
data llinoutx;set llinoutx;keep score xfreq xpp;
proc print data=llinoutx noobs;run;
```

score	xfreq	xpp	von Davier et al.'s fitted frequencies
0	1	3.300	3.30
1	3	6.435	6.44
2	8	11.767	11.77
3	25	20.175	20.17
4	30	32.433	32.43
5	64	48.889	48.89
6	67	69.099	69.10
7	95	91.572	91.57
8	116	113.787	113.79
9	124	132.575	132.58
10	156	144.832	144.83
11	147	148.356	148.36
12	120	142.490	142.49
13	129	128.321	128.32
14	110	108.355	108.35
15	86	85.790	85.79
16	66	63.688	63.69
17	51	44.332	44.33
18	29	28.935	28.93
19	15	17.707	17.71
20	11	10.161	10.16

```
data llinouty;set llinouty;keep score yfreq ypp;
proc print data=llinouty noobs;run;
```

score	yfreq	ypp	von Davier et al.'s fitted frequencies
0	0	1.706	1.71
1	4	3.775	3.77
2	11	7.650	7.65
3	16	14.242	14.24
4	18	24.436	24.44
5	34	38.752	38.75
6	63	56.978	56.98
7	89	77.905	77.91
8	87	99.354	99.35
9	129	118.542	118.54
10	124	132.723	132.72
11	154	139.868	139.87
12	125	139.154	139.15
13	131	131.097	131.10
14	109	117.307	117.31
15	98	100.000	100.00
16	89	81.457	81.46
17	66	63.596	63.60
18	54	47.732	47.73
19	37	34.545	34.54
20	17	24.180	24.18

Appendix E

Example 1—Comparing the Fitted and Observed Moments for *X*

```
data xactual;set llinoutx;
do i=1 to 1000*xfreq;
output;
end;
drop i;
proc means data=xactual mean std skew kurt;
var score;
title 'Moments based on the actual frequencies of X.';
run;
```

Moments based on the actual frequencies of X

The MEANS Procedure
Analysis Variable: score

Mean	Std Dev	Skewness	Kurtosis
10.8183070	3.8058570	0.0025777	-0.4678344

```
data xfitted;set llinoutx;
do i=1 to 1000*xpp;
output;
end;
drop i;
proc means data=xfitted mean std skew kurt;
var score;
title 'Moments based on the fitted frequencies of X.';
run;
```

Moments based on the fitted frequencies of X

The MEANS Procedure

Analysis Variable: score

Mean	Std Dev	Skewness	Kurtosis
10.8183085	3.8058348	-0.0648419	-0.3010408

Appendix F

Example 1—Comparing the Fitted and Observed Moments for *Y*

```
data yactual;set llinouty;
do i=1 to 1000*yfreq;
output;
end;
drop i;
proc means data=yactual mean std skew kurt;
var score;
title 'Moments based on the actual frequencies of Y.';
run;
```

Moments based on the fitted frequencies of Y

The MEANS Procedure

Analysis Variable: score

Mean	Std Dev	Skewness	Kurtosis
11.5931271	3.9342663	-0.0626866	-0.4988359

```
data yfitted;set llinouty;
do i=1 to 1000*ypp;
output;
end;
drop i;
proc means data=yfitted mean std skew kurt;
var score;
title 'Moments based on the fitted frequencies of Y.';
run;
```

Moments based on the fitted frequencies of Y

The MEANS Procedure
Analysis Variable: score

Mean	Std Dev	Skewness	Kurtosis
11.5931343	3.9342516	-0.0626788	-0.4277965

Appendix G

Example 2—Entering the Dataset and Defining the Score Functions

The dataset includes all possible *X*-*Y* score combinations. For space considerations, only the first and last *X*-*Y* combinations are shown in this appendix.

```
data llinbi;
input x y freq;
cards;
0 0 0
0 1 0
0 2 1
0 3 0
0 4 0
……..
20 14 0
20 15 0
20 16 2
20 17 3
20 18 3
20 19 2
20 20 0
;
/*This defines the relevant score functions corresponding
to the desired moments.*/

data llinbi;set llinbi;
x2=x**2;
y2=y**2;
x3=x**3;
y3=y**3;
xy=x*y;
```

Appendix H

Example 2—Fitting the Bivariate Model

```
/*Here is the bivariate model.*/
proc genmod data=llinbi;
output out=llinoutb p=fpp;
model freq=x y x2 y2 x3 y3 xy/link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.LLINBI
Distribution	Poisson
Link Function	Log
Dependent Variable	freq
Observations Used	441

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	433	242.7285	0.5606
Scaled Deviance	433	242.7285	0.5606
Pearson Chi-Square	433	232.7659	0.5376
Scaled Pearson X2	433	232.7659	0.5376
Log Likelihood		1889.4990	

Algorithm converged.

Appendix I

Example 2—The Marginal Fitted Frequencies for *X* and *Y*

The marginal fitted frequencies for *X* and *Y* are from SAS PROC GENMOD and from von Davier et al. (2004, p. 118).

```
proc means data=llinoutb noprint;
var fpp;
class x;
output out=llinbifx sum=sfy;run;
data llinbifx;set llinbifx;drop _TYPE_ _FREQ_;
/*Table 8.3*/
proc print data=llinbifx noobs;run;
```

x	sfy	von Davier et al.'s fitted frequencies
0	2.30	2.3
1	5.17	5.17
2	10.47	10.47
3	19.22	19.22
4	32.32	32.32
5	50.01	50.01
6	71.57	71.57
7	95.03	95.03
8	117.40	117.4
9	135.26	135.26
10	145.70	145.7
11	147.07	147.07
12	139.44	139.44
13	124.46	124.46
14	104.81	104.81
15	83.44	83.44
16	62.90	62.9
17	44.93	44.93
18	30.39	30.39
19	19.42	19.42
20	11.67	11.67

```
proc means data=llinoutb noprint;
var fpp ;
class y;
output out=llinbify sum=sfx;run;
data llinbify;set llinbify;drop _TYPE_ _FREQ_;
/*Table 8.3*/
proc print data=llinbify noobs;run;
```

y	sfx	von Davier et al.'s fitted frequencies
0	2.29	2.29
1	5.27	5.27
2	10.86	10.86
3	20.31	20.31
4	34.68	34.68
5	54.38	54.38
6	78.55	78.55
7	104.78	104.78
8	129.34	129.34
9	147.99	147.99
10	157.19	157.19
11	155.24	155.24
12	142.76	142.76
13	122.39	122.39
14	97.87	97.87
15	72.94	72.94
16	50.49	50.49
17	32.25	32.25
18	18.83	18.83
19	9.93	9.93
20	4.67	4.67

Appendix J

Example 2—Comparing the Observed and Fitted Moments

```
/*Observed moments.*/
data xyobs;set llinoutb;
do i=1 to 1000*freq;
output;
end;
drop i;
proc means data=xyobs mean std skew kurt;
title 'Moments based on the actual frequencies of X and
Y.';
var x y;
run;
```

Moments based on the actual frequencies of X and Y

The MEANS Procedure

Variable	Mean	Std Dev	Skewness	Kurtosis
x	10.8183070	3.8058570	0.0025777	-0.4678344
y	10.3888507	3.5866342	-0.0055610	-0.5156722

```
/*Fitted Moments.*/
data xyfit;set llinoutb;
do i=1 to 1000*fpp;
output;
end;
drop i;
proc means data=xyfit mean std skew kurt;
title 'Moments based on the fitted frequencies of X and
Y.';
var x y;
run;
```

Moments based on the fitted frequencies of X and Y

The MEANS Procedure

Variable	Mean	Std Dev	Skewness	Kurtosis
x	10.8184227	3.8055450	0.0026940	-0.3406657
y	10.3888881	3.5863219	-0.0055181	-0.2597099

Appendix K

Example 2—Comparing the Observed and Fitted Cross-moment

```
/*Observed Cross-Moment.*/
proc corr data=xyobs cov noprob;
title 'Covariances and Correlations based on the actual X-Y
frequencies.';
var x y;
run;
```

Covariances and Correlations based on the actual X-Y frequencies

The CORR Procedure

2 Variables: x y

Covariance Matrix

	x	y
x	14.48454750	10.58270279
y	10.58270279	12.86394511

Pearson Correlation Coefficients

	x	y
x	1.00000	0.77528
y	0.77528	1.00000

```
/*Fitted Cross-Moment.*/
proc corr data=xyfit cov noprob;
title 'Covariances and Correlations based on the fitted X-Y
frequencies.';
var x y;
run;
```

Covariances and Correlations based on the fitted X-Y frequencies

The CORR Procedure

2 Variables: x y

Covariance Matrix

	x	y
x	14.48217263	10.58250193
y	10.58250193	12.86170494

Pearson Correlation Coefficients

	x	y
x	1.00000	0.77539
y	0.77539	1.00000

Appendix L

Example 3—Entering the Dataset and Defining the Score Functions

```
/*Calling in the Dataset from an Excel file named
ht2000p159.xls.*/
PROC IMPORT OUT= WORK.HTEX3
            DATAFILE= "C:\ht2000p159.xls"
            DBMS=EXCEL2000 REPLACE;
     GETNAMES=YES;
RUN;


/*Defining an Indicator variable for the X scores that have
teeth.*/
data htex3;set htex3;
if x=47.5 then tx=1;
if x=41.5 then tx=1;
if x=37.5 then tx=1;
if x=31.5 then tx=1;
if x=27.5 then tx=1;
if x=21.5 then tx=1;
if x=17.5 then tx=1;
if x=11.5 then tx=1;
if x=7.5 then tx=1;
if x=1.5 then tx=1;
if x=-2.5 then tx=1;
run;


/*Defining the relevant score functions for Models 1-3.*/
data htex3;set htex3;
x2=x**2;
x3=x**3;
x4=x**4;
x5=x**5;
y2=y**2;
y3=y**3;
y4=y**4;
y5=y**5;
xy=x*y;
x2y=x2*y;
xy2=x*y2;
x2y2=x2*y2;
if tx=. then tx=0;
run;
```

Appendix M

Example 3—Fitting the First Model

```
/*Fitting Model 1 (16 parameters).*/
proc genmod data=htex3;
output out=ht3mod1 p=pred1;
class tx;
model freq=x x2 x3 x4 x5 tx x*tx x2*tx x3*tx x4*tx x5*tx y
y2 y3 y4 y5/link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.HTEX3
Distribution	Poisson
Link Function	Log
Dependent Variable	freq
Observations Used	551

Class Level Information

Class	Levels	Values
tx	2	0 1

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	534	24187.2590	45.2945
Scaled Deviance	534	24187.2590	45.2945
Pearson Chi-Square	534	30739.9815	57.5655
Scaled Pearson X2	534	30739.9815	57.5655
Log Likelihood		94229.6611	

Algorithm converged.

Appendix N

Example 3—Fitting the Second Model

```
/*The second model.*/
proc genmod data=htex3;
output out=ht3mod2 p=pred2;
class tx;
model freq=x x2 x3 x4 x5 tx x*tx x2*tx x3*tx x4*tx x5*tx y
y2 y3 y4 y5 xy /link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.HTEX3
Distribution	Poisson
Link Function	Log
Dependent Variable	freq
Observations Used	551

Class Level Information

Class	Levels	Values
tx	2	0 1

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	533	577.8977	1.0842
Scaled Deviance	533	577.8977	1.0842
Pearson Chi-Square	533	556.0665	1.0433
Scaled Pearson X2	533	556.0665	1.0433
Log Likelihood		-1.79769E308	

ERROR: The mean parameter is either invalid or at a limit of its range for some observations.

Appendix O

Example 3—Fitting the Third Model

```
/*The third model.*/
proc genmod data=htex3;
output out=ht3mod3 p=pred3;
class tx;
model freq=x x2 x3 x4 x5 tx x*tx x2*tx x3*tx x4*tx x5*tx y
y2 y3 y4 y5 xy x2y xy2 x2y2/link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.HTEX3
Distribution	Poisson
Link Function	Log
Dependent Variable	freq
Observations Used	551

Class Level Information

Class	Levels	Values
tx	2	0 1

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	530	435.4766	0.8217
Scaled Deviance	530	435.4766	0.8217
Pearson Chi-Square	530	410.4314	0.7744
Scaled Pearson X2	530	410.4314	0.7744
Log Likelihood		106105.5523	

Algorithm converged.

Appendix P

Example 4—Entering the Dataset and Defining the Score Functions

```
data xy;
input x y freq;
cards;
1 1 8
1 2 15
1 3 12
1 4 23
1 5 11
2 1 0
2 2 1
2 3 4
2 4 10
2 5 9
3 1 0
3 2 0
3 3 4
3 4 4
3 5 6
4 1 0
4 2 0
4 3 0
4 4 5
4 5 4
5 1 0
5 2 0
5 3 0
5 4 0
5 5 5
;

/*The frequencies that are zero are impossible X-Y
combinations that must be eliminated before the smoothing
model is fit (X cannot be greater than Y).*/
data xy; set xy;
if freq=0 then delete;

/*Defining the score functions.*/

data xy; set xy;
x2=x**2;
y2=y**2;
xy=x*y;
```

Appendix Q

Example 4—Fitting the First Model

```
/*Model 1*/
proc genmod data=xy;
output out=xyffit1 p=xyffit;
model freq=x x2 y y2 /link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.XY
Distribution	Poisson
Link Function	Log
Dependent Variable	freq
Observations Used	15

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	10	13.4550	1.3455
Scaled Deviance	10	13.4550	1.3455
Pearson Chi-Square	10	11.9491	1.1949
Scaled Pearson X2	10	11.9491	1.1949
Log Likelihood		149.6674	

Algorithm converged.

Appendix R

Example 4—Fitting the Second Model

```
/*Model 2*/
proc genmod data=xy;
output out=xyffit2 p=xyffit;
model freq=x x2 y y2 xy /link=log dist=p type3;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.XY
Distribution	Poisson
Link Function	Log
Dependent Variable	freq
Observations Used	15

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	9	10.1326	1.1258
Scaled Deviance	9	10.1326	1.1258
Pearson Chi-Square	9	9.5425	1.0603
Scaled Pearson X2	9	9.5425	1.0603
Log Likelihood		151.3286	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	3.0399	0.7594	1.5515	4.5284	16.02	<.0001
x	1	-2.1072	0.5382	-3.1620	-1.0523	15.33	<.0001
x2	1	0.0960	0.0871	-0.0746	0.2667	1.22	0.2700
y	1	0.7877	0.4290	-0.0531	1.6286	3.37	0.0663
y2	1	-0.1468	0.0721	-0.2881	-0.0055	4.15	0.0417
xy	1	0.2614	0.1467	-0.0262	0.5490	3.17	0.0749
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

Appendix S

Example 4—The Observed (freq) and Fitted (xyffit) Frequencies

The Observed (freq) and Fitted (xyffit) Frequencies from Models 1 and 2 Y are from SAS PROC GENMOD and from Holland and Thayer (2000, p. 181).

```
/*Fits from Model 1*/
proc print data=xyffit1 noobs;run;
```

x	y	freq	x2	y2	xy	xyffit	Holland and Thayer's fitted frequencies
1	1	8	1	1	1	6.9597	6.96
1	2	15	1	4	2	11.3114	11.31
1	3	12	1	9	3	15.4183	15.42
1	4	23	1	16	4	17.6259	17.63
1	5	11	1	25	5	16.8991	16.90
2	2	1	4	4	4	4.7763	4.78
2	3	4	4	9	6	6.5104	6.51
2	4	10	4	16	8	7.4426	7.44
2	5	9	4	25	10	7.1357	7.14
3	3	4	9	9	9	4.0499	4.05
3	4	4	9	16	12	4.6297	4.63
3	5	6	9	25	15	4.4388	4.44
4	4	5	16	16	16	4.2427	4.24
4	5	4	16	25	20	4.0678	4.07
5	5	5	25	25	25	5.4917	5.49

```
/*Fits from Model 2*/
proc print data=xyffit2 noobs;run;
```

x	y	freq	x2	y2	xy	xyffit	Holland and Thayer's fitted frequencies
1	1	8	1	1	1	6.8968	6.90
1	2	15	1	4	2	12.6751	12.68
1	3	12	1	9	3	17.3671	17.37
1	4	23	1	16	4	17.7405	17.74
1	5	11	1	25	5	13.5106	13.51
2	2	1	4	4	4	3.4671	3.47
2	3	4	4	9	6	6.1696	6.17
2	4	10	4	16	8	8.1848	8.18
2	5	9	4	25	10	8.0952	8.10
3	3	4	9	9	9	2.6558	2.66
3	4	4	9	16	12	4.5758	4.58
3	5	6	9	25	15	5.8776	5.88
4	4	5	16	16	16	3.0999	3.10
4	5	4	16	25	20	5.1711	5.17
5	5	5	25	25	25	5.5130	5.51