MULTILEVEL MODEL FOR A SAMPLE OF CONTINGENCY TABLES: A SIMULATION STUDY

Minerva Montero[1], Ernestina Castell[2], Mario M. Ojeda[3]
[1] Instituto de Cibernética, Matemática y Física, La Habana, Cuba
[2] Universidad de La Habana, La Habana, Cuba
[3] Universidad Veracruzana, Xalapa, México

RESUMEN:
Frecuentemente en las investigaciones científicas, los problemas a estudiar conducen al análisis de una muestra de tablas de contingencia. La complicación estadística surge a partir de la estructura jerárquica de los datos ya que generalmente se violan las hipótesis de independencia y observaciones idénticamente distribuidas, necesarias para el uso de muchos de los métodos estadísticos estándar. Algunos procedimientos han sido desarrollados para el análisis de modelos multinivel con respuesta binaria. Montero, Castell y Ojeda (2001) propusieron utilizar el enfoque GSK para analizar datos categóricos en una muestra de tablas de contingencia. El algoritmo de estimación se basa en los mínimos cuadrados generalizados. En este trabajo se usa un estudio de simulación para explorar algunas propiedades del procedimiento propuesto. El estudio examina los efectos de diferentes tamaños de muestra y magnitudes de la varianza de nivel-2 sobre la precisión de las estimaciones de los efectos fijos y aleatorios. También se investiga la precisión de los errores estándar.

Palabras claves: tablas de contingencia, modelos multinivel, enfoque GSK

ABSTRACT:
Samples of contingency tables frequently occur in many fields of the scientific research. The statistical complexities in analyzing these data result from the hierarchical structure of the data set because the standard assumptions of independence and identically distributed observations, commonly required in the use of many of the standard statistical methods, are generally violated. Several approximate procedures have been developed for the analysis of multilevel models with binary response. We propose the use of the GSK (Grizzle, Starmer and Koch, 1969) approach in order to fit multilevel models related to a sample of contingency tables. The estimation algorithm is based on iterative generalized least squares. In this paper, a simulation study is used to explore some properties of the proposed procedure. The study examines the effects of different sample sizes and 2-level variance size on the accuracy of the estimates of the fixed and random parameters. We also investigate the accuracy of the standard errors.

Key words: contingency tables, multilevel models, GSK approach

1. INTRODUCTION

Frequently, in many fields of scientific research, many practical problems of interest lead to the analysis of a sample of contingency tables. In many of these situations the tables have been obtained from different studies, as in the case of meta-analysis (Glass, 1976, Hedges and Olkin, 1985), where it is of interest to combine the information to obtain a pattern of heterogeneity or association. Other examples of interest include the analysis of panel data (Hsiao, 1986; Hamerle and Ronning 1995) where the contingency tables are studied longitudinally; or as in the case of small areas estimation (Rudas, 1986. 1998), where it is required to estimate any probability associated with the table.

Statistical complexities in many of these situations result from the hierarchical structure of the data. The observations within each table are generally correlated, thus the standard assumptions of independence and identically distributed observations, required in many of the standard statistical methods, are generally violated. An appropriated approach that takes into account dependence across as well as within tables is known as multilevel modeling. Multilevel models (Goldstein, 1995), are also referred as random coefficient models (Longford, 1995) or hierarchical models (Bryk and Raudenbush, 1992).

Standard models generally assume a common effect for all tables. In practice, however, we are interested in a relationship, which is usually not homogeneous among several tables. When analyzing a sample of contingency tables following multilevel analysis some parameters in a generalized linear model can differ for the different tables. Various models have been developed according with this basic idea, for example, the hierarchical generalized linear models (HGLMs) of Lee and Nelder (1996, 2001), which combine generalized linear models (GLMs) with random effects in the linear predictor, Generalized linear mixed models (GLMMs) (Breslow and Clayton, 1993), which assume Gaussian random effects and form a subclass of HGLMs. Other approaches also make use of generalized linear models to develop techniques of multilevel analysis for proportions (Longford, 1994; Goldstein, 1991). For proportions very close to 0 or 1 and a small number of groups, Bayesian estimation using Gibbs sampling have been developed (for a brief description see Goldstein, 1995).

This paper presents how the generalized least squares method of categorical data analysis, also known as the GSK approach (Grizzle, Starmer and Koch, 1969), can be used as strategy for the fit of multilevel models related to a multidimensional contingency table sample. Dependencies between the observations are modeled via random effects. In this paper the validity of the proposed procedure is explored by means of the logit function for a binary response in a simulation study.
The asymptotic properties of the estimation methods used in multilevel models presuppose that the sample size must be large. There are some simulation studies, which investigate the effects of sample size on the accuracy of the estimates for multilevel models (Kreft, 1996, Longford, 1994, Mok, 1999). Several studies suggest that variance size of the random effects also affects the precision of the estimation procedure (Rodriguez and Goldman, 1995, 2001), but there are relatively very few simulation studies on this topic and only a small amount of research work investigates the accuracy of the standard error used to test the specific model parameters. A review of the existing research can be consulted in Hox and Mass (2002).

In this paper, the simulation study is used to examine the effects of different sample sizes and variance distributions on the accuracy of the estimates given by the procedure presented in this paper. We also investigate the accuracy of the standard error.

2. A MULTILEVEL MODEL FOR PROPORTIONS

We consider a 2-level hierarchical data structure. Suppose a sample of *J* contingency tables (level-2 units) where the rows of each table, called subpopulations, represent *I* levels (level-1 units) of an explanatory variable or combinations of levels of several explanatory variables. Random samples of size n_{ij} $(i=1,\cdots,I;\quad j=1,\cdots J)$ are selected from rows. The responses are classified according to *r* categories with n_{ilj} $(l=1,2,\cdots,r)$ denoting the number of elements classified in the *l*-th response category for the *i*-th subpopulation of the *j*-th table.

Let $\boldsymbol{\pi}_j = (\boldsymbol{\pi}'_{1j},\boldsymbol{\pi}'_{2j},\cdots\boldsymbol{\pi}'_{Ij})'$, where $\boldsymbol{\pi}_{ij} = (\pi_{i1j},\pi_{i2j},\cdots,\pi_{irj})'$ with $\sum_l \pi_{ilj} = 1$, represents a vector of probabilities for the *j*-th table.

Each set of probabilities has $r-1$ linearly independent elements.

Let $\boldsymbol{F}(\boldsymbol{\pi}_j) = [F_1(\boldsymbol{\pi}_j),F_2(\boldsymbol{\pi}_j),\cdots,F_a(\boldsymbol{\pi}_j)]'$ be a vector of $a \le I(r-1)$ functions of $\boldsymbol{\pi}_j$.

The goal of this paper is centered in the logit function but different types of functions may be represented in a relatively simple manner using matrix notation. For practical purposes, many functions are interesting; however, Forthoper and Koch, 1973, demonstrated that a wide range of problems in categorical data analysis could be considered within a few general classes of functions. For example, loglinear and logit functions are special cases of $\boldsymbol{F}(\boldsymbol{\pi}_j) = \boldsymbol{B}_j log(\boldsymbol{E}_j \boldsymbol{\pi})$ for certain matrices $\boldsymbol{B}_j$ and $\boldsymbol{E}_j$.

By analyzing tables where the *I* samples of the *J* tables are independent, Grizzle, Starmer and Koch (1969) showed that once the function has been specified it can be used as dependent variable in a linear model. Applying the same idea to a sample of contingency tables when the *I* samples of the *J* tables are not independent we developed a separate level-1 linear model for each of the *J* level-2 units.
Let $\boldsymbol{p}_j$ be the vector of observed proportions, given in the same way as π_j. Thus, $\boldsymbol{F}(\boldsymbol{p}_j)$ denote the observed proportion functions. Logit response models have response functions of the form $\boldsymbol{F}(\boldsymbol{p}_j) = \boldsymbol{B}_j\, log(\boldsymbol{p}_j)$, where the elements of $\boldsymbol{B}_j$ are the coefficients of the natural logarithms of the vector $\boldsymbol{p}_j$.

The level- 1 models for the observed proportions are of the form:

$$\boldsymbol{F}(\boldsymbol{p}_j) = \boldsymbol{X}_j \boldsymbol{\beta}_j + \boldsymbol{e}_j \;, \qquad j = 1,2,\cdots,J, \tag{1}$$

where $\boldsymbol{X}_j$ is a $a \times t$ design matrix of know constants having rank t, $\boldsymbol{\beta}_j$ is a $t \times 1$ vector of unknown parameters and $\boldsymbol{e}_j$ is a $a \times 1$ vector of random errors at level 1, where $\mathrm{E}(\boldsymbol{e}_j) = \boldsymbol{0}$ and $\mathrm{Var}(\boldsymbol{e}_j) = \boldsymbol{\Omega}_{\boldsymbol{e}_j}$.
Note $\boldsymbol{\beta}_j$ can vary through tables. The variability of the *J* coefficients $(\beta_{1k},\cdots,\beta_{Jk})$ pertaining to the *k*-th variable $(k = 1,\cdots,t)$ can be explained by an additional set of explanatory variables $Z_1, Z_2, \cdots, Z_q$ measured at level 2. It then follows that:

$$\beta_{jk} = \boldsymbol{Z}'_{jk} \boldsymbol{\Gamma}_k + u_{jk} \quad , \qquad j = 1,2,\cdots,J \tag{2}$$

where $\boldsymbol{Z}_{jk}$ are the values of the q_k variables at level 2 in the *j*-th table, $\boldsymbol{\Gamma}_k$ is the $q_k \times 1$ vector of the coefficients associated with explanatory variables at level 2 and u_{jk} are the non observable level 2 random errors with zero mean and finite variance.
Equation (2) can be more succinctly expressed as:

$$\boldsymbol{\beta}_j = \boldsymbol{Z}_j \boldsymbol{\Gamma} + \boldsymbol{u}_j \;, \qquad j = 1,2,\cdots,J \tag{3}$$

where $\boldsymbol{Z}_j = \mathrm{diag}(\boldsymbol{Z}'_{j1},\cdots \boldsymbol{Z}'_{jt})$ is a $t \times \mathrm{Q}$ block-diagonal matrix, $\mathrm{Q} = q_1 + q_2 + \cdots q_t$, $\boldsymbol{\Gamma} = (\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \cdots, \boldsymbol{\Gamma}_t)$ is the $\mathrm{Q} \times 1$ vector of coefficients and $\boldsymbol{u}_\mathrm{j}$ is the $t \times 1$ vector of level 2 random errors, where $\mathrm{E}(\boldsymbol{u}_j) = \boldsymbol{0}$ and $\mathrm{Var}(\boldsymbol{u}_j) = \boldsymbol{\Omega}_{\boldsymbol{u}_j}$.
Substituting equation (3) into equation (1) we obtain a single expression for the model given by

$$\boldsymbol{F}(\boldsymbol{p}_j) = \boldsymbol{X}_j \boldsymbol{Z}_j \boldsymbol{\Gamma} + \boldsymbol{X}_j \boldsymbol{u}_j + \boldsymbol{e}_j \,, \tag{4}$$

where we assume that the errors between levels are mutually independent.
Model (4) can be expressed compactly as:

$$\boldsymbol{F}(\boldsymbol{p}) = \boldsymbol{A\Gamma} + \boldsymbol{Xu} + \boldsymbol{e} \tag{5}$$

where $\boldsymbol{F}(\boldsymbol{p}) = \left(\boldsymbol{F}(\boldsymbol{p}_1)', \boldsymbol{F}(\boldsymbol{p}_2)', \cdots, \boldsymbol{F}(\boldsymbol{p}_J)'\right)'$, $\boldsymbol{A} = \left(\boldsymbol{Z}_1'\boldsymbol{X}_1', \boldsymbol{Z}_2'\boldsymbol{X}_2', \cdots, \boldsymbol{Z}_J'\boldsymbol{X}_J'\right)'$. The index j is redundant to $\boldsymbol{X}_j$, since such a variable is constant within tables, $\boldsymbol{X} = \operatorname{diag}(\boldsymbol{X}_j)$ a diagonal block matrix with $\boldsymbol{X}_j$ in the j-th diagonal block, $\boldsymbol{e} = \left(\boldsymbol{e}_1', \cdots, \boldsymbol{e}_J'\right)'$ and $\boldsymbol{u} = \left(\boldsymbol{u}_1', \cdots, \boldsymbol{u}_J'\right)'$

The assumptions are summarized below:

$$\mathrm{E}(\boldsymbol{e}) = \mathrm{E}(\boldsymbol{u}) = \boldsymbol{0}$$

$$\mathrm{Var}(\boldsymbol{ee}') = \boldsymbol{\Omega_e}, \quad \mathrm{Var}(\boldsymbol{uu}') = \boldsymbol{\Omega_u},$$

where

$$\boldsymbol{\Omega_e} = \operatorname{diag}\left(\boldsymbol{1}_{a\times 1} \otimes \boldsymbol{\Omega}_{\boldsymbol{e_j}}\right) \text{ and } \boldsymbol{\Omega_u} = \operatorname{diag}\left(\boldsymbol{1}_{t\times 1} \otimes \boldsymbol{\Omega}_{\boldsymbol{u_j}}\right)$$

$$\mathrm{Cov}(\boldsymbol{e}, \boldsymbol{u}) = \boldsymbol{0}$$

We can then say that the corresponding variance-covariance matrix has the following general form:

$$\boldsymbol{V_F} = \boldsymbol{X\Omega_u X}' + \boldsymbol{\Omega_e}$$

Note, unlike the formulation under independence, the approach presented here accounts for dependence across tables as well as within tables, however, the model (5) can be considered as a special case of the general lineal model if we write:

$$\boldsymbol{F}(\boldsymbol{p}) = \boldsymbol{A\Gamma} + \boldsymbol{e}^*$$

where $\boldsymbol{e}^* = \boldsymbol{Xu} + \boldsymbol{e}$, $\mathrm{E}(\boldsymbol{e}^*) = \boldsymbol{0}$ and $\mathrm{Cov}(\boldsymbol{e}^*, \boldsymbol{e}^*) = \boldsymbol{V_F}$

If the variance-covariance matrix is known, the parameter vector $\boldsymbol{\Gamma}$ is estimated by generalized least squares (GLS),

$$\hat{\boldsymbol{\Gamma}} = \left(\boldsymbol{A}'\boldsymbol{V_F}^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}'\boldsymbol{V_F}^{-1}\boldsymbol{F}(\boldsymbol{p}) \tag{6}$$

When $\boldsymbol{V_F}$ is unknown it is a common practice to substitute an estimate $\hat{V}_F$ for it in the expression (6). We carry out an iterative procedure. The procedure starts with initial estimates of the fixed parameters. We propose those from a generalized least squares (GLS) fit for categorical data ignoring the random errors at level 2. (see appendix for the estimation of $\boldsymbol{V_F}$ when the *I* subpopulation of the *J* tables are independents). Once a suitable starting value for the fixed parameters is obtained our development continues similarly to the GLS in Goldstein (1987) .In each iteration we use the current estimates of the fixed and random parameters until an appropriate value of convergence is reached.

3. SIMULATION STUDIES

Generally, estimates and tests for the regression coefficients in multilevel modeling are accurate for samples of moderate sizes, but estimates and tests for the variances are not. Following the estimation procedure presented in this paper a simulation framework is used to distinguish the effects of the sample size and the level-2 variance on the estimates of parameters of a multilevel model. We are particularly interested in the multilevel logistic regression model, but the procedure can be extended to other link functions.

3.1 *The simulation model*

In order to illustrate the behavior of the procedure presented in this paper we use a simple 2-level model, with a single dichotomous explanatory variable at the level 1, given by the following equation:

$$f_{ij} = logit(\pi_{ij}) = log(\pi_{ij}) - log(1 - \pi_{ij}) = \gamma_{00} + \gamma_{10}x_{ij} + u_j x_{ij}, \quad (7)$$

The level-2 random effects u_j are assumed to follow a multivariate normal distribution each with mean zero and variance σ_u^2.

The simulation work reported here describes the distribution of the fixed and the random parameters based on many replications of artificial data sets with known parameters.
In Table 1 are the 35 different designs considered in the simulation study. Observe that there are 5 different numbers of contingency tables and 7 different subpopulation sizes. All the designs are balanced, so that the subpopulation sizes n_{ij} are the same for every $i\ j$.

INSERTAR TABLA 1

The designs are classified in 5 different types of designs, depending on the number of tables (Table 1). Each cell of Table 1 represents the total sample size.

One small and one large level-2 variance were assumed ($\sigma_{\boldsymbol{u}}^2 = 0.5$ and $\sigma_{\boldsymbol{u}}^2 = 1.0$). Thus, there are 35×2 = 70 different design conditions and for each condition 500 data sets were generated. The values of the parameters γ_{00} and γ_{10} were set to 0.5 and 1.0, respectively, in data generating. The explanatory variable is fixed following the requirements of the 2×2 contingency tables. $logit(\pi_{ij})$ is obtained adding the fixed part and level-2 random errors. Finally, the values of the variable n_{ij} (used for obtain p_{ij}) are generated from a binomial distribution with parameter π_{ij}.

3.2 Results

Estimates of the fixed and random parameters (γ_{00}, γ_{10} and $\sigma_{\boldsymbol{u}}^2$) were obtained for 500 simulations under the different conditions of the designs. The estimation procedure converged in all 35 000 simulated data sets. The mean of the estimates values, the mean of estimated standard errors and the standard deviation (empirical standard error) of the fixed and random effect estimates over the 500 runs for each design condition are presented in Tables 2, 3 and 4. The results are presented separately according with the two values used for the 2-level variance. The case of variance equal 0.5 is called Study I, and the case where variance is equal 1 is called Study II.

From the parameter estimates over 500 simulations of the design conditions, the following measures were also computed: the (signed) bias, the empirical sampling variance and the empirical Mean Square Error (MSE).

Box-plots of the distributions of the 500 estimates of the fixed and random parameters for each design are shown in Figures 1, 2 and 3. In each figure, the display a) represents the distributions of the estimates in Study I, and the display b) those of Study II.

We summarize the results separately for each of the estimated fixed and random parameters, but in this paper we focus our attention on the estimates of the random parameter.

3.2.1 Estimates of the fixed parameters, γ_{00} and γ_{10} .

Displays 1.a) y 1.b) are very similar. In all cases the median centers the true value. In all design types, the sampling variance decreases noticeably as the subpopulation size increases from 10 to about 25, and the estimates are closer to 0.5 when the subpopulation size is relatively large. Indeed, given the number of tables, by increasing the subpopulation size the variance decreases. This shows that, for the fixed part of the model, the benefit in efficiency by increasing the number of tables is relatively small.

The tendency shown in the Box-plots 2.a) and 2.b) does not differ substantially from previous results, but the variation of the estimates is noticeably larger than in study I. The median is closer to the true value (1.0).

In general, the results from all designs are consistent: As sample size increases, the estimated values approximate to the true value, that is: the bias tends to zero.

3.2.2 Estimates of the random parameter, $\sigma_{\boldsymbol{u}}^2$.

Box-plots 3.a) and 3.b) show that the random parameter is overestimated, except for samples with large subpopulation sizes. In almost all designs the median is above the true value, there are even, in the study I, some Box-plots, completely over the value 0.5. Box-plots based on the samples with large number of tables and small subpopulation size, such as D-10 and E-10 designs, do not contain the true value, but these box-plots are considerably shorter. For the Study II, designs with small subpopulation sizes, like the type A, produce widely different values. In general, given a design type, the increasing of the subpopulation size improves the median of the estimates, approaching to the true value.
In order to complete the comparisons, Figures 4, 5 and 6 display the plots of the bias, sampling variance and MSE of $\hat{\sigma}_u^2$ against subpopulation size, respectively. Also here, display a) is for Study I and display b) is for Study II.

The greater part of designs types are biased upwards (Figure 4.a). In study I, the bias approaches zero rapidly and theses are essentially indistinguishable given a subpopulation size. The biases in the study II (Figure 4.b) are also nearly identical when, given a subpopulation size, the number of tables is greater than 50 (Designs types C, D and E). All design types for subpopulation sizes that are greater than or equal to 25 tend to be consistent. Unexpectedly, biases reported in type A designs are smaller than the others.

Type A designs have evidently larger sampling variances than the other designs (Figure 5.a and 5.b). The difference is more noticeable in study II. It is obvious that designs with more tables have smaller variances; however, the difference in precision of the estimates for designs with 50 or more tables is not substantial, independently of the subpopulation size. For example, note that difference between the variances of type D and E designs is essentially indistinguishable, mainly in study I. When we compare both studies, the results show that, in general, the estimates in study I tend to be more consistent than in study II.

The favorable position of designs with more tables and larger subpopulation size is evident from the MSE of the random parameter, which is plotted against subpopulation size in Figures 6.a) and 6.b). The difference in the MSE is not substantial for a number of tables greater than 25, especially in study I.
To complete the comparisons an interesting observation was made on both studies (I and II). For designs with the same total sample size, it appeared that even when the sample size is considerably large, designs involving large subpopulation size and fewer tables tend to be less biased than designs involving greater amount tables and small subpopulation sizes. To illustrate this situation, Figure 7 shows the trend of the contribution of each estimated value of σ_u^2 over the replicates for three designs (E-50, E-100, C-

100) in study I. Note that, for example, the design C-100 (total sample size = 10000) is less biased than the design E-50 (total sample size = 10000). A comparison of this last design with another of greater sample size (E-100, total sample size = 20000) reveals that the performance of the estimators remains nearly the same. Concluding, there would be little gain from using a large total sample size. Only when the subpopulation size is large the estimates improve by reducing the upward bias. The comparisons in the case of the study II are similar.

The normal probability plots show that the distribution of estimates for the simulations do not differ substantially between studies I and II, therefore, only a selection of these plots for the random parameter in the study I is presented in this paper. Figure 8 reveals that by increasing the number of tables, the behavior of the estimates improves considerably. Except for a few outliers, the plots for the random parameter are reasonably consistent with the expected asymptotic normality.

3.2.3 Estimated Standard Error

The averages of the estimated standard errors of the fixed effects are reasonably close to those calculated empirically from the replications. The differences from the random parameter are noticeable.

To assess the accuracy of the standard error, for each parameter in each simulated data set, the 95% confidence interval was obtained using the asymptotic standard normal distribution (Goldstein, 1995; Longford, 1993). We found that for fixed parameters nearly all of the intervals include the true value.

Tables 2, 3 and 4 also give interesting information with reference to the relative size of the bias and variance. Note that all estimations of the fixed effects are located within 1 standard error of the true value. For the designs with small level-2 variance, however, the bias in the random parameter estimates is of the order of 1 to 6 standard errors. For the designs with large level-2 variance the bias is of the order of 1 to 11 standard errors.

CONCLUSIONS

The procedure presented in this paper can be used to analyze a sample of contingency tables. This method provides an alternative to the multilevel modeling for hierarchical categorical data. We focused on the application of the proposed procedure to logit response models but this approach is more general and can be used for model other functions of the probabilities.

The purpose of the simulation study presented in this paper was to investigate the effects of sample size and the influence of different variance size on estimating the parameters (regression coefficients and level-2 variance) given by the proposed procedure. The simulation results show that:

The estimates for fixed parameters are accurate for samples of moderate size, but estimates of the variance are not. For both studies (I and II) the level-2 variance is overestimated. The main characteristic regarding bias is that it decreases by increasing subpopulation size.

The analysis presented in this paper shows that for the two studies reported here we need a large number of observations to reach a reasonable performance of the estimator in terms of bias and efficiency, especially regarding the estimation of the random parameter. On the other hand, no large differences are expected among designs with 50 or more tables. The level-2 variance estimates require large subpopulation sizes; but the number of tables is less important.
In summary, the estimator for the parameters exhibits greater reduction in bias and gain in efficiency by increasing the sub-population size more than increasing number of tables. Moreover, among the two studies considered, the proposed procedure can be expected to perform better when the level-2 variance is small.

The dependence of the bias on the subpopulation size needs further researches, including more complex models and extreme data sets. It does seem, however, that the bias for the categorical data arises principally from the relatively small subpopulation size per contingency table. An additional analysis of models including imbalance is also necessary, to be sure if accuracy is always found under large sample.

ACKNOWLEDGEMENTS

The simulation study was performed while the first author was visiting the Abdus Salam International Center for Theoretical Physics, Trieste, in your capacity of Associate.
We would like to thank Dr. Jesús Sánchez for several helpful comments with respect to the revision of the original manuscript.

Appendix: ***Estimation of*** $\boldsymbol{V_F}$ ***considering independence.***

If we consider the *I* subpopulations of the *j*-th table as being uncorrelated with one another a consistent estimator for the covariance matrix of $\boldsymbol{p_j}$ is the $IC \times IC$ matrix:

$$\boldsymbol{V_j}(\boldsymbol{p_j}) = \operatorname{diag}\left(\boldsymbol{V}_{1j}(\boldsymbol{p}_{1j}), \boldsymbol{V}_{2j}(\boldsymbol{p}_{2j}), \cdots, \boldsymbol{V}_{Ij}(\boldsymbol{p}_{Ij})\right),$$

with the matrices

$$\boldsymbol{V_{ij}}(\boldsymbol{p_{ij}}) = \frac{1}{n_{ij}}\left[\boldsymbol{D}_{p_{ij}} - \boldsymbol{p_{ij}}\boldsymbol{p_{ij}}'\right], \qquad (i = 1, \cdots, I)$$

where $\boldsymbol{D}_{p_{ij}}$ a $C \times C$ matrix diagonal with elements of the vector $\boldsymbol{p_{ij}}$ on the main diagonal.

We assume that $\boldsymbol{F_j}$ has continuous second order partial derivatives in an open region containing $\boldsymbol{\pi_j}$. A consistent estimator for the covariance matrix of $\boldsymbol{F_j}$ is the $a \times a$ matrix:

$$\hat{\boldsymbol{V}}_{F_j} = \boldsymbol{H_j}\left[\boldsymbol{V_j}(\boldsymbol{p_j})\right]\boldsymbol{H_j}'$$

where $H = \left[\partial F(\pi_j)/\partial \pi_j : \pi_j = p_j\right]$ is the $a \times IC$ matrix of first partial derivatives of the functions $\boldsymbol{F_j}$ evaluated on $\boldsymbol{p_j}$.

Observations from different tables are mutually independent and, if no function combines probabilities from more than one population, this independence is maintained through the transformation. Thus, the covariance between observations from different tables is zero, and the estimated covariance matrix of the ***F*** has the form***:***

$$\hat{\boldsymbol{V}}_{F} = \operatorname{diag}\left(\hat{\boldsymbol{V}}_{F_1}, \hat{\boldsymbol{V}}_{F_2}, \cdots, \hat{\boldsymbol{V}}_{F_J}\right).$$

Note: A consistent estimator for the covariance matrix of the function logit (Forthofer and Koch, 1973) is the matrix:

$$\hat{\boldsymbol{V}}_{F_j} = \boldsymbol{B}_j\boldsymbol{D}_j^{-1}\left[\boldsymbol{V}_j(\boldsymbol{p}_j)\right]\boldsymbol{D}_j^{-1}\boldsymbol{B}_j' \quad ,$$

where the main diagonal of $\boldsymbol{D}_j$ contains the elements of the vector $\boldsymbol{p}_j'$.

REFERENCES

Breslow, N. E. and Clayton, D. G. (1993): Approximate inference in generalized linear models. *Journal of the American Statistical Association*, **88,** 9-25**.**

Bryk. A.S. and Rraudenbush. S.W., 1992: *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications, Thousand Oaks, California, USA.

Forthoper, R. N. and Koch, G. G. (1973): An analysis for compounded functions of categorical data. *Biometrics*. 29, 143-157.

Hamerle, A. and Honning, G. (1995): Panel Analysis for Qualitative Variables. in: G. Arminger, C.C. Clogg und M.E. Sobel (Herausgeber): *A Handbook for Statistical Modeling in the Social and Behavioral Sciences.* New York: Plenum, 401-451.

Hedges, L. V. and Olkin, I (1985): *Statistical Methods for Meta–analysis.* Academic Press, New York.

Hox, J.J. and Maas, C.J.M. (2002): Sample sizes for multilevel modeling. In: J. Blasius, J. Hox , E. de Leeuw & P. Schmidt (eds.) (2002). *Social Science Methodology in the New Millennium*. Proceedings of the Fifth International Conference on Logic and Methodology. Second expanded edition. Opladen, RG: Leske + Budrich Verlag (CD-ROM).

Hsiao, C. (1986): *Analysis of Panel Data*. Cambridge University Press, New York.
Glass, G. V. (1976): Primary, secondary and meta–analysis of research. *Educational Researcher*, **5**, 3-8.
Goldstein, H. (1987). Multilevel models in educational and social research. London: Griffinn.
Goldstein, H. (1991): Nonlinear multilevel models, with an application to discrete response data. *Biometrika.* **78**. 1. pp 45-51.
Goldstein, H. (1995): *Multilevel Statistical Models.* 2^{nd}. Ed. Halsted Press, New York.
Grizzle, J.E., Starmerc. F. and Koch G., (1969): Analysis of categorical data by linearmodels. *Biometrics*, **25**, 489-504.
Kreft, I.G.G. (1996): Are Multilevel Techniques Necessary? An overview, including Simulation Studies …………….
Lee Y. and Nelder J.A., (1996): Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society*, Series *B*. **58**: 619-678.
Lee Y. and Nelder J.A., (2001): Hierarchical generalized linear models: a synthesis of generalized linear model, random-effect models and structured dispersions. *Biometrika*, **88**, 987-1006.
Longford, N.T., (1993): *Random coefficient models*. Oxford: Clarendon Press.
Longford, N.T., (1994): Logistic regression with random coefficients. *Computational Statistics and Data Analysis*, **97**, 1-15.
Longford, N.T., (1995): Random coefficient models, in: G. Arminger, C. C. Clogg, M. E. Sobel (eds.), *Handbook of Statistiscal Modeling for the Social and Behavioral Sciences*, Plenum Press, New York, 519-577.
Mok, M. (1999): Sample size requirements for 2-level designs in educational research….
Rodriguez, G. and Goldman, N. (1995): An assessment of estimation procedures for multilevel models with binary response. *Statistics in Medicine*, **13**, 541-51.
Rodriguez, G. and Goldman, N. (2001): Improved estimation procedures for multilevel models with binary response: a case-study. *Journal of the Royal Statistical Society*, Series A, **164.** Part 2, pp. 339-335.
Rudas, T. (1986): A Monte Carlo comparison of the small sample behavior of the Pearson, the likelihood ratio and the Cressie-Read statistics. *Journal of Statistical Computation and Simulation*, **24**, 107-120.
Rudas, T. (1998): *Odds Ratios in the Analysis of Contingency Tables*. Sage Publications, Thousand Daks, USA.