Quality and Quantity, 21: 49–57 (1987)

Models for the probability of concordance in cross-classification tables

ALAN AGRESTI, JOHN SCHOLLENBERGER & DENNIS WACKERLY
Department of Statistics, University of Florida, Gainesville, Florida 32611, USA

Keywords: concordance, logit model, probit model, ordinal data, Kendall's tau

Abstract For cross-classification tables having an ordinal response variable, logit and probit models are formulated for the probability that a pair of subjects is concordant. For multidimensional tables, generalized models are given for the probability that the response at one setting of explanatory variables exceeds the response at another setting. Related measures of association are discussed for two-way tables.

1. Introduction

The notion of concordance or discordance of a pair of observations has been important in the development of measures of association for ordinal variables. For instance, Kendall's tau equals the difference between the probabilities of concordance and discordance, for a randomly selected pair of observations. For ordinal categorical data, Goodman and Kruskal's gamma equals this difference, conditional on the event that the pair is untied on both variables. Other generalizations of Kendall's tau for categorical data include tau-*b*, tau-*c*, and Somers' *d*.

The main focus in this article is on logit and probit models for the relative numbers of concordant and discordant pairs, for cross-classification tables in which the response variable is ordinal. Section 2 deals with the case in which there is a single explanatory variable, the categories of which are assigned scores. An example is discussed in which the logit of a concordance probability for mental health status is modeled as a simple function of parental socioeconomic status. Section 3 introduces multidimensional generalizations in which qualitative and quantitative explanatory variables can appear. The fitting of these models using weighted least squares is described in an appendix. Section 4 presents model-related measures of association that correspond to generalized correlation measures for ordinal and interval variables.

2. Models for two-way tables

Consider an *rxc* cross-classification table in which the column variable, Y, is an ordinal response variable and for which scores $x_1 < x_2 < \dots < x_r$ can be assigned to the levels of the row (independent) variable, X. For example, Y might measure a subject's views concerning the legalization of abortion (should be available on demand, should be legal only in restricted cases, should never be legal), whereas X might measure the subject's attained education level.

For the models we consider, the response is a function of the conditional probabilities on Y, given X. Hence, it is natural to assume a product multinomial sampling model. Specifically, we assume that independent samples of sizes $\{n_i\}$ are taken at the various levels of X, and that π_{ij} denotes the probability that when $X = x_i$, a subject is classified in the jth level of Y. Note that $\sum_j \pi_{ij} = 1$ for $i = 1, \ldots, r$.

Let Y_i denote the response for a subject selected at random in row i, and let Y_j denote an independent observation on the response for a subject selected at random in row j. For $x_i < x_j$, let $P_c(x_i, x_j)$ denote the probability that the pair (Y_i, Y_j) satisfies $Y_i < Y_j$, given that $Y_i \neq Y_j$. That is, $P_c(x_i, x_j)$ is the probability of concordance for the pair, conditional on the event that the responses differ. Hence,

$$P_c(x_i, x_j) = \frac{\sum\sum_{b>a} \pi_{ia}\pi_{jb}}{\sum\sum_{b>a} \pi_{ia}\pi_{jb} + \sum\sum_{b<a} \pi_{ia}\pi_{jb}}. \qquad [2.1]$$

Similarly, the conditional probability of discordance for this pair of X-values is $P_d(x_i, x_j) = 1 - P_c(x_i, x_j)$.

In some cases, it would be informative to describe how the relative numbers of concordant and discordant pairs depend on the values (x_i, x_j) for the possible pairings of the explanatory variables. The concordance or discordance of a pair can be treated as a quasi-binary response. Paralleling the development of linear models for various transformations of the probability of "success" for binary variables, we model transformations of the probability of concordance.

To motivate the choice of transformation for the probability of concordance, suppose that there is an underlying continuum such that for each fixed value of X, Y has a normal distribution with mean $\gamma_0 + \gamma_1 X$ and variance σ^2. If Φ denotes the standard normal cumulative distribution function, then for any pair of values $x_i < x_j$,

$$P_c(x_i, x_j) = \Phi\left(\frac{\gamma_1(x_j - x_i)}{\sigma\sqrt{2}}\right).$$

Hence, the probability of concordance depends on the x-values only through their difference, and the probit (inverse normal) transformation is linearly related to this difference. Note also that $P_c(x_i, x_j)$ is a monotonic function of $(x_j - x_i)$ and that $P_c(x_i, x_j) \to 1/2$ as $(x_j - x_i) \to 0$.

Table 1. Cross-classification of mental health status by parent's socioeconomic status.

Parent's socioeconomic status	Mental Health Status			
	Well	Mild Symptom Formation	Moderate Symptom Formation	Impaired
A (high)	64	94	58	46
B	57	94	54	40
C	57	105	65	60
D	72	141	77	94
E	36	97	54	78
F (low)	21	71	54	71

In many applications it would seem reasonable to assume that there is an underlying distribution for which $P_c(x_i, x_j)$ is approximately a monotonic function of $x_j - x_i$ such that $P_c(x_i, x_j) \to 1/2$ as $x_j - x_i \to 0$. Then a simple model that may be adequate for the cross-classification table is

$$F_{ij} = \beta(x_j - x_i), \quad x_i < x_j \qquad [2.2]$$

where F_{ij} is a monotone transformation from (0, 1) onto $(-\infty, \infty)$. For instance, natural choices are the probit transformation $F_{ij} = \Phi^{-1}(P_c(x_i, x_j))$ or the logit transformation $F_{ij} = \log[P_c(x_i, x_j)/(1 - P_c(x_i, x_j))]$. The assumption that $P_c(x_i, x_j) \to 1/2$ as $x_j - x_i \to 0$ implies that no constant term is needed in the model. For model [2.2] when the x-values are u units apart, the probability of concordance is $\Phi(\beta u)$ for the probit model and $\exp(\beta u)/(1 + \exp(\beta u))$ for the logit model.

If model [2.2] holds, then

$$F_{ij} + F_{jk} = F_{ik}, \quad \text{for } 1 \leqslant i < j < k \leqslant r. \qquad [2.3]$$

It follows that there are $r - 1$ linearly independent pairs of comparisons. Since model [2.2] has only one parameter, it has $r - 2$ residual degrees of freedom. The appendix shows how this model can be fitted using weighted least squares.

The data in Table 1 were analyzed using standard loglinear and logit models by Goodman (1979) and by Agresti (1984, pp. 134–135). Here we treat mental health status as an ordinal response, and we will use model [2.2] to estimate the probability that a subject at one parent's SES level has better mental health than a subject at some other parent's SES level. We used integer scoring $\{x_i = i\}$ for the levels of parent's SES, but other monotonic choices of scores gave similar results.

Cox (1970, pp 26–29) notes that the logit and probit transformations are in reasonable agreement for values of the probability of concordance in the range of 0.1 through 0.9. Thus, we restrict our illustration to the logit transformation in this discussion. Using methods discussed in the appendix, we obtained a

Table 2. Comparison of observed concordance probabilities with values (in parentheses) predicted by model [2.2], for Table 1.

Parent's SES levels	B	C	D	E	F
A	0.496	0.543	0.561	0.616	0.669
	(0.536)	(0.571)	(0.605)	(0.639)	(0.671)
B		0.548	0.566	0.622	0.677
		(0.536)	(0.571)	(0.605)	(0.639)
C			0.519	0.575	0.630
			(0.536)	(0.571)	(0.605)
D				0.555	0.610
				(0.536)	(0.571)
E					0.556
					(0.536)

goodness-of-fit statistic of $\chi^2 = 3.36$ for model [2.2], based on $df = r - 2 = 4$ degrees of freedom. The parameter estimate $\hat{\beta} = 0.142$ has standard error 0.023, implying that there is very strong evidence that the concordance probabilities exceed the discordance probabilities; in other words, higher parent's SES tends to correspond to better mental health status.

Using model [2.2], we obtain an estimated probability of concordance of

$$\hat{P}_c(x_i, x_j) = \frac{\exp[0.142(x_j - x_i)]}{1 + \exp[0.142(x_j - x_i)]}.$$

The estimated probabilities are 0.536, 0.571, 0.605, 0.639, 0.671 for the parent's SES distances 1, 2, 3, 4, and 5. For instance, the probability that SES level $i + 1$ corresponds to higher mental health status than SES level i is estimated to be $0.536/0.464 = 1.15$ times higher than the probability that it corresponds to lower mental health status. The estimated probabilities are compared to their observed sample values in Table 2. The difference is less than 0.05 for all 15 comparisons. Based on the goodness-of-fit test and this comparison, it appears that model [2.2] adequately models concordance probabilities for these data.

3. Models for multidimensional tables

For $x_i \neq x_j$, let $P(x_i, x_j)$ denote the probability that $Y_i < Y_j$, given that $Y_i \neq Y_j$. that is, $P(x_i, x_j) = P_c(x_i, x_j)$ if $x_i < x_j$ and $P(x_i, x_j) = P_d(x_j, x_i)$ if $x_i > x_j$. Thus, model (2.2) can be expressed in terms of any (x_i, x_j) (rather than only $x_i < x_j$) by defining F_{ij} to be the probit or logit transformation of $P(x_i, x_j)$. This type of notation is useful when the explanatory variable is a vector, which is the case considered in this section. Specifically, suppose that the cross-classification table has ordinal response variable Y and a vector $\boldsymbol{X}$ of $k \geqslant 2$ explanatory variables. We consider two types of models. The first, a direct generalization of (2.2), describes how $P(Y_j > Y_i \mid Y_i \neq Y_j)$ depends on values $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ of $\boldsymbol{X}$. The second describes how the conditional association between Y and one element of $\boldsymbol{X}$ varies according to the values of the other elements of $\boldsymbol{X}$.

Let Y_i and Y_j be independent observations on Y at the values $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ of $\boldsymbol{X}$. As in Section 2, we can model how $P(Y_j > Y_i \mid Y_i \neq Y_j)$ depends on $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. Let F_{ij} denote the logit or probit transformation of $P(Y_j > Y_i \mid Y_i \neq Y_j)$. If it is reasonable to assume that for each fixed value of $\boldsymbol{X}$, the underlying distribution of Y is approximately normal with mean $\gamma_0 + \boldsymbol{\gamma}'\boldsymbol{X}$ and constant variance, then a natural generalization of model (2.2) is

$$F_{ij} = \boldsymbol{\beta}'(\boldsymbol{x}_j - \boldsymbol{x}_i). \qquad [3.1]$$

Suppose we take $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ such that $\boldsymbol{x}_i' = (x_i, \boldsymbol{w}_i')$ and $\boldsymbol{x}_j' = (x_j, \boldsymbol{w}_j')$ with $x_i \neq x_j$ but $\boldsymbol{w}_i' = \boldsymbol{w}_j'$. Then F_{ij} represents a transformed probability of concordance for the conditional (Y, X) association, given $\boldsymbol{W}$. Model [3.1] is rather restrictive, since it implies that this probability is identical at all $\boldsymbol{w}$ values.

If our primary interest is directed at how conditional (Y, X) concordance probabilities depend on a vector of covariates $\boldsymbol{W}$, a different form of model may be more reasonable. Let $P(x_i, x_j; \boldsymbol{w})$ denote the probability that $Y_i < Y_j$ given that $Y_i \neq Y_j$, when $\boldsymbol{x}_i' = (x_i, \boldsymbol{w}')$ and $\boldsymbol{x}_j' = (x_j, \boldsymbol{w}')$; i.e., $\boldsymbol{w}$ is the common value of $\boldsymbol{W}$ in both members of the pair. Let $F_{ij}(\boldsymbol{w})$ denote the logit or probit transformation of $P(x_i, x_j; \boldsymbol{w})$. Then a simple generalization of model [2.2] that describes how the conditional (Y, X) concordance probability depends on $\boldsymbol{W}$ is given by

$$F_{ij}(\boldsymbol{w}) = (\alpha + \boldsymbol{\beta}'\boldsymbol{w})(x_j - x_i). \qquad [3.2]$$

Given $\boldsymbol{X} = \boldsymbol{x}$, the transformation of $P(x_i, x_j; \boldsymbol{w})$ is again directly proportional to the distance between the X-values. For a given distance between X-values, the transformation of $P(x_i, x_j; \boldsymbol{w})$ is a linear function of $\boldsymbol{W}$. Note that in model [3.2], unlike model [3.1], both members of a pair must have the same $\boldsymbol{W}$ value.

Model [3.2] is analogous to a Bradley-Terry paired comparisons model proposed by Semenya and Koch (1980) (see also Semenya et al., 1983) in which the preference parameters have a linear pattern. For the special case of $c = 2$ responses, the logit formulation of model [2.2] is equivalent to the loglinear model with linear-by-linear association, which corresponds to Goodman's (1979) uniform association model when the $\{x_i\}$ are equal-interval. When $c = 2$ it also is equivalent to standard linear logit models. When $c > 2$, models discussed here are not equivalent to loglinear models or other models that have been proposed for ordinal responses (e.g., see McCullagh 1980 and Agresti 1983). A disadvantage of the concordance models, relative to these others, is that cell probabilities are not determined by the model parameters. Hence, the models do not provide information about structural aspects such as stochastic orderings on the response. If concordance probabilities *are* of interest, though, an advantage of these models is the direct modelling of such probabilities.

All these models can be fitted using weighted least squares. Details are rather cumbersome, and are presented in the appendix. For several examples and additional model formulations for the multidimensional case, see Schollenberger (1982). Listings of programs for implementing these procedures which use the Statistical Analysis Systems (SAS) procedure MATRIX can also be found there, and are available upon request from the authors.

4. Association measures

Most ordinal measures of association give summary descriptions of the relative numbers of concordant and discordant pairs. Let S denote the sign function,

$$S(u)=\begin{cases}-1, & u<0\\ 0, & u=0\\ 1, & u>0.\end{cases}$$

A pair of observations with $x_i < x_j$ is concordant if $S(Y_j - Y_i) = 1$, and discordant if $S(Y_j - Y_i) = -1$. We can regard model (2.2) as a way of describing how well $S(Y_j - Y_i)$ can be predicted based on the distance $x_j - x_i$. Related association measures can be defined that, like that model, utilize the ordinal nature of Y and the quantitative nature of X.

Suppose there are n observations on (X, Y). Daniels (1944) defined a generalized correlation coefficient

$$G=\frac{\sum a_{ij}b_{ij}}{\left[\left(\sum a_{ij}^2\right)\left(\sum b_{ij}^2\right)\right]^{1/2}},$$

where a_{ij} is an x-score and b_{ij} is a y-score for the (i, j)-th pair of observations, such that $a_{ij} = -a_{ji}$ and $b_{ij} = -b_{ji}$. For the scoring $a_{ij} = x_j - x_i$ and $b_{ij} = y_j - y_i$, G is the Pearson correlation coefficient. For the scoring $a_{ij} = S(x_j - x_i)$ and $b_{ij} = S(y_j - y_i)$, G is Kendall's tau for fully-ranked data and it is Kendall's tau-b for ordered categorical data.

Corresponding to model [2.2], it seems natural to take $a_{ij} = x_j - x_i$ and $b_{ij} = S(y_j - y_i)$ in G. Some properties of this interval-ordinal measure are as follows:

1. For an $r \times r$ table with $P_{ii} = 1/r$, $G = [2(r+1)/3r]^{1/2}$.
2. For a sample of size n with continuous variables, $G \leqslant [2(n+1)/3n]^{1/2}$.

3. If (X, Y) is distributed bivariate normal with correlation ρ, then $\gamma = \sqrt{2/\pi}\ \rho$, where

$$\gamma = \frac{E\left[(X_j - X_i)S(Y_j - Y_i)\right]}{\left[\operatorname{Var}(X_j - X_i)\right]^{1/2}\left[\operatorname{Var}(S(Y_j - Y_i))\right]^{1/2}}.$$

4. An interval-ordinal measure that is a compromise between the Pearson and Spearman correlations is given by G with $a_{ij} = x_j - x_i$ and $b_{ij} = R(y_j) - R(y_i)$, where R is the rank function. When the observations are fully ranked on Y, this measure is the constant multiple $(3n/2(n+1))^{1/2}$ times the value of G obtained using the sign score for Y.

Unfortunately, these compromise measures do not seem as useful as their ordinal-ordinal or interval-interval analogues. For instance, $S(Y_j - Y_i)$ cannot be perfectly correlated with $X_j - X_i$ when either X or Y has more than two distinct values, so $|G|$ cannot equal 1.0 for tables of size greater than 2×2. A more thorough discussion of these measures, together with the derivation of their asymptotic variances using the delta method, is contained in Schollenberger (1982).

Appendix: WLS estimation of model parameters

The models we presented have the form $\boldsymbol{F} = \boldsymbol{U\beta}$, where $\boldsymbol{U}$ is a design matrix and $\boldsymbol{F}$ is the vector of the F_{ij} obtained at all the appropriate pairings of settings of explanatory variables. The models can be fitted using the weighted least squares methodology for categorical data, as described in Grizzle, Starmer and Koch (1969).

Let $\hat{\boldsymbol{F}}$ and $\hat{F}_{ij}$ denote $\boldsymbol{F}$ and F_{ij} calculated for the sample proportions. For instance, for the logit transformation and the two-way rxc table of cell counts $\{n_{ij}\}$, $\hat{\boldsymbol{F}}$ has the $r(r-1)/2$ elements

$$\hat{\boldsymbol{F}} = (\hat{F}_{12}, \hat{F}_{13}, \ldots, \hat{F}_{1r}, \ldots, \hat{F}_{r-1,r}),$$

with

$$\hat{F}_{ij} = \log\left[\sum_{b>a}\sum n_{ia}n_{jb} \Big/ \sum_{b<a}\sum n_{ia}n_{jb}\right].$$

However, there are dependencies among the $\{\hat{F}_{ij}\}$ leading to a singular asymptotic covariance matrix. To avoid this singularity, we use a smaller set of responses transformed from the $\{\hat{F}_{ij}\}$. For simplicity, we discuss this transformation for the two-way table and model [2.2].

If model [2.2] holds, then

$$\boldsymbol{F} = \boldsymbol{Z}\boldsymbol{F}_T,$$

where $\boldsymbol{F}_T' = (F_{12}, F_{13}, \ldots, F_{1r})$ and

$$\boldsymbol{Z}' = \begin{array}{c} \begin{array}{ccccccccccc} 1,2 & & & & 1,r & 2,3 & & & 2,r & & r-1,r \end{array} \\ \left[\begin{array}{ccccc|cccc|c|c} 1 & 0 & 0 & \ldots & 0 & -1 & -1 & \ldots & -1 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 & 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & & -1 \\ 0 & 0 & 0 & \ldots & 1 & 0 & 0 & & 1 & & 1 \end{array}\right] \end{array}$$

Alternatively, $F_T = TF$, where $T = (Z'Z)^{-1}Z'$. This suggests the $r-1$ transformed responses

$$\hat{G} = T\hat{F}$$

for the model $G = (TU)\beta$. This transition can also be motivated by a Bradley-Terry model for the comparisons of responses, and it was suggested for a related purpose by Semenya and Koch (1980).

Let $\pi' = (\pi_1', \ldots, \pi_r')$, where $\pi_i' = (\pi_{i1}, \ldots, \pi_{ic})$ with $\Sigma_j \pi_{ij} = 1$. Let $V(\hat{\pi}_i) = (1/n_{i+})[D(\pi_i) - \pi_i \pi_i']$, where $D(\pi_i)$ is the cxc diagonal matrix with the elements of π_i on the diagonal. The covariance matrix of $\hat{\pi}$ is the $rc \times rc$ block diagonal matrix $V(\hat{\pi})$ that has the $\{V(\hat{\pi}_i)\}$ matrices on the diagonal. If H denotes the $r(r-1)/2 \times rc$ matrix of partial derivations $[\partial F_{ij}/\partial \pi_{ab}]$, the asymptotic covariance matrix of $\hat{F}$ is $HV(\hat{\pi})H'$, and the asymptotic covariance matrix of $\hat{G}$ is

$$V(\hat{G}) = THV(\hat{\pi})H'T'.$$

Depending on the response function in the model, H will have different forms. If $A_{ij} = \Sigma\Sigma_{v<u} \pi_{iv}\pi_{ju}$ and $B_{ij} = \Sigma\Sigma_{v>u} \pi_{iv}\pi_{ju}$, the logit response has the form $F_{ij} = \log[A_{ij}/B_{ij}]$ and the probit response has the form $F_{ij} = \Phi^{-1}(A_{ij}/(A_{ij} + B_{ij}))$. Define

$$A'_{i(b),j} = \partial A_{ij}/\partial \pi_{ib} = \sum_{b<u} \pi_{ju}, \quad A^*_{i,j(b)} = \partial A_{ij}/\partial \pi_{jb} = \sum_{v<b} \pi_{iv},$$
$$B'_{i(b),j} = \partial B_{ij}/\partial \pi_{ib} = \sum_{b>u} \pi_{ju}, \quad B^*_{i,j(b)} = \partial B_{ij}/\partial \pi_{jb} = \sum_{v>b} \pi_{iv}.$$

Then for the logit response, a row in H has elements

$$\frac{\partial F_{ij}}{\partial \pi_{ib}} = \frac{A'_{i(b),j}}{A_{ij}} - \frac{B'_{i(b),j}}{B_{ij}} \qquad b = 1, \ldots, c$$
$$\frac{\partial F_{ij}}{\partial \pi_{jb}} = \frac{A^*_{i,j(b)}}{A_{ij}} - \frac{B^*_{i,j(b)}}{B_{ij}} \qquad b = 1, \ldots, c$$
$$\frac{\partial F_{ij}}{\partial \pi_{ab}} = 0 \qquad \text{if } a \neq i,\ a \neq j.$$

If ϕ denotes the standard normal density function, then for the probit response,

$$\frac{\partial F_{ij}}{\partial \pi_{ib}} = \frac{B_{ij}A'_{i(b),j} - A_{ij}B'_{i(b),j}}{(A_{ij} + B_{ij})^2 \phi\left[\Phi^{-1}\left(\dfrac{A_{ij}}{A_{ij} + B_{ij}}\right)\right]} \qquad b = 1, \ldots, c$$
$$\frac{\partial F_{ij}}{\partial \pi_{jb}} = \frac{B_{ij}A^*_{i,j(b)} - A_{ij}B^*_{i,j(b)}}{(A_{ij} + B_{ij})^2 \phi\left[\Phi^{-1}\left(\dfrac{A_{ij}}{A_{ij} + B_{ij}}\right)\right]}, \qquad b = 1, \ldots, c$$
$$\frac{\partial F_{ij}}{\partial \pi_{ab}} = 0 \qquad \text{if } a \neq i \text{ and } a \neq j.$$

Let $\hat{V}(\hat{G})$ denote the sample value of $V(\hat{G})$. For the model $G = (TU)\beta$, it follows that the WLS estimate of β is

$$b = \left(U'T'\hat{V}(\hat{G})^{-1}TU\right)^{-1} U'T'\hat{V}(\hat{G})^{-1}\hat{G}.$$

It is easily seen that the same value $\boldsymbol{b}$ is obtained for several other ways of forming $\hat{\boldsymbol{G}}$, for instance for $\boldsymbol{T}$ matrices corresponding to $\boldsymbol{F}_T' = (F_{12}, F_{23}, \ldots, F_{r-1,r})$ or $\boldsymbol{F}_T' = (F_{1i}, F_{2i}, \ldots, F_{ir})$.

The goodness of fit of the model can be tested using the statistic

$$\chi^2 = (\hat{\boldsymbol{G}} - \boldsymbol{TUb})'\hat{\boldsymbol{V}}(\hat{\boldsymbol{G}})^{-1}(\hat{\boldsymbol{G}} - \boldsymbol{TUb}).$$

For the transformed version of model [2.2], there are $r-1$ response functions and 1 parameter, so χ^2 has a null asymptotic chi-squared distribution with $df = r-2$. Hypotheses of the form $\boldsymbol{C\beta} = \boldsymbol{O}$ can be tested using the statistic

$$(\boldsymbol{Cb})'\left[\boldsymbol{C}\left(\boldsymbol{U}'\boldsymbol{T}'\hat{\boldsymbol{V}}(\hat{\boldsymbol{G}})^{-1}\boldsymbol{TU}\right)^{-1}\boldsymbol{C}'\right]^{-1}(\boldsymbol{Cb}),$$

for which the null asymptotic chi-squared distribution has degrees of freedom equal to the rank of C.

If we consider multidimensional tables, the above approach is easily generalized. For models like [3.1] that are not of the conditional type, we adopt the following convention. If the i-th explanatory variable has r_i levels, $i = 1, \ldots, k$, we simply view the problem in the $r \times c$ context with $r = r_1 r_2 \ldots r_k$. The same transformations discussed earlier in this section lead to the desired results. For conditional models like [3.2] we make the corresponding transformation within each level of the control variables. For instance, suppose $X^{(2)}$ is a control variable that has r_2 levels and $X^{(1)}$ is an interval variable with r_1 levels. Then, for each value of $X^{(2)}$ we make the transformation given by $\boldsymbol{T}$ with $r = r_1$, leading to $r_1 - 1$ response functions for each of the r_2 values of $X^{(2)}$. Thus, in this case, we would analyze a set of $r_2(r_1 - 1)$ response functions.

References

Agresti, A. (1983). A Survey of Strategies for Modeling Cross-Classifications Having Ordinal Variables. *J. Amer. Statist. Assoc., 78*, 184–198.

Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. New York: Wiley.

Cox, D.R. (1970). *Analysis of Binary Data*. London: Chapman and Hall.

Daniels, H.E. (1944). The Relation between Measures of Correlation in the Universe of Sample Permutations. *Biometrika, 33*, 129–125.

Goodman, L.A. (1979). Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories. *J. Amer. Statist. Assoc., 74*, 537–552.

Grizzle, J.E., Starmer, C.F., and Koch, G.G. (1969). Analysis of Categorical Data by Linear Models. *Biometrics, 25*, 489–504.

McCullagh, P. (1980). Regression Models for Ordinal Data. *J. Roy. Statist. Soc., B 42*, 109–142.

Schollenberger, J.M. (1982). Categorical Data Analysis with an Ordinal Response Variable and Interval Explanatory Variables. Ph.D. Thesis, University of Florida.

Semenya, K., and Koch, G.G. (1980). Compound Function and Linear Model Methods for the Multivariate Analysis of Ordinal Categorical Data. Chapel Hill: University of North Carolina Institute of Statistics Mimeo Series No. 1323.

Semenya, K., Koch, G.G., Stokes, M.E., and Forthofer, R.N. (1983). Linear Models Methods for Some Rank Function Analyses of Ordinal Categorical Data, *Communications in Statistics, A 12*, 1277–1298.