This research was supported in part by National Insitutes of Health, Institute of General Medical Sciences through Grant GM-70004-04, by the National Institute of Child Health and Human Development through Grant HD-00371-02, and by the U. S. Bureau of the Census through Joint Statistical Agreement JSA 74-2.

A GENERAL METHODOLOGY FOR THE ANALYSIS OF EXPERIMENTS WITH REPEATED MEASUREMENT OF CATEGORICAL DATA

by

Gary G. Koch, Jean L. Freeman, Daniel H. Freeman
Department of Biostatistics
University of North Carolina at Chapel Hill

and

Robert G. Lehnen
Department of Political Science
University of Houston

Institute of Statistics Mimeo Series No. 961

November 1974

1. INTRODUCTION

The investigation of theories in the social and biological sciences frequently results in data arrayed as multi-dimensional contingency tables. Such tables represent a complex underlying structure about which the scientist wishes to evaluate certain research questions. Thus, the precise formulation of corresponding hypotheses within the framework of a general statistical methodology is a necessary prerequisite for accomplishing this objective in an informative manner.

Experiments (or surveys) with repeated measurements on subjects are a class of research designs which frequently involve such conceptually complex data arrays. Examples of such investigations are commonplace in the literature and include the following:

i. Longitudinal follow-up studies of the health status of subjects who are being treated for chronic diseases including situations such as change-over designs in clinical trials (e.g., the examples in Section 3.1, 3.2, and 3.3 and in Koch, Imrey, and Reinfurt [20]);

ii. Longitudinal panel surveys of political opinion or economic status (e.g., Lehnen and Koch [24]);

iii. Intensity of agreement surveys dealing with comparative attitudes towards different policy questions (e.g., Lehnen and Koch [25]);

iv. Preference surveys involving complete or partial rankings of two or more alternative issues or items (e.g., the example in Section 3.4).

In all of these situations, the most common research design can be regarded as involving samples from s sub-populations of subjects with each

subject selected from each sub-population being exposed to d different measurement conditions (e.g., treatments or a set of stimuli like comparative attitudinal policy questions) and classified for each in terms of a response with L levels. However, it should be recognized that the underlying structure of such studies can become extremely complex if data collection is incomplete in the sense that different subsets of the d conditions (or set of stimuli) are measured on different subjects. This topic is outside the scope of this paper, but is discussed in detail in Koch, Imrey, and Reinfurt [20] and Lehnen and Koch [24].

The questions of substantive interest which are associated with experiments (or surveys) with repeated measurements are as follows:

1. Are there any differences among the sub-populations with respect to the distribution of the responses to the d conditions?
2. Are there any differences among the distributions of responses to the d conditions within each of the respective sub-populations?
3. Are there any differences among the sub-populations with respect to differences among the distributions of responses to the d conditions?

As indicated in Koch and Reinfurt [22], the questions (1) - (3) are directly analogous to the hypotheses of "no whole-plot effects," "no split-plot effects," and "no whole-plot x split-plot" interaction in standard split-plot experiments as described in Anderson and Bancroft [1], Federer [10], or Steel and Torrie [29]. Here, however, the conceptual formulation of such hypotheses must be undertaken in terms of an underlying (s x r) contingency table where $r = L^d$ represents the number of possible multivariate response profiles. This topic is discussed in considerable detail in Section 2.2 where it is indicated that the first order marginal distributions of response for each

of the d measurement conditions contain most of the relevant information for dealing with the questions (1) - (3). Test statistics for such hypotheses and the estimation of parameters for underlying linear regression models are obtained through weighted least squares computations by methods as described in Grizzle, Starmer, and Koch [16] (hereafter GSK). Four examples which illustrate various aspects of the scope of such analyses are discussed in Section 3.

2. METHODOLOGY

The purpose of this section is the presentation of a general statistical methodology for repeated measurement experiments (or surveys) which involve multivariate categorical data. For this purpose, Section 2.1 is concerned with summarizing the most important aspects of the matrix operations associated with the GSK approach to the analysis of complex contingency tables. These procedures are then applied to repeated measurement experiments in Section 2.2. In this context, attention is primarily focused on the formulation of certain hypotheses of interest and specifying the relevance of these hypotheses to the experimental (or survey) conditions under consideration. Thus, the resulting methodology represents a categorical data analogue to more well-known counterparts for quantitative data like multivariate analysis of variance as described by Cole and Grizzle [8] and Morrison [27] in the parametric case and multivariate rank analysis as described by Koch [17,18] in the non-parametric case. Finally, several computational strategies for effectively dealing with certain problems associated with the manipulation of large contingency tables are given in Section 2.3.

2.1. General Framework

Let $j = 1,2,\ldots,r$ index a set of categories which correspond to r, possibly multivariate, response profiles associated with a specific set of dependent variable(s) of interest. Similarly, let $i = 1,2,\ldots,s$ index a set of categories which correspond to distinct sub-populations as defined in terms of pertinent independent variables. If samples of size n_i where $i = 1,2,\ldots,s$ are independently selected from the respective sub-populations, then the resulting data can be summarized in an (s x r) contingency table as shown in Table 1 where n_{ij} denotes the frequency of response category j in the sample from the i-th sub-population.

1. OBSERVED CONTINGENCY TABLE

Sub-population	Response profile categories				
	1	2	...	r	Total
1	n_{11}	n_{12}	...	n_{1r}	n_1
2	n_{21}	n_{22}	...	n_{2r}	n_2
...	...	...	...	...	...
s	n_{s1}	n_{s2}	...	n_{sr}	n_s

The vector $\underset{\sim}{n}_i$ where $\underset{\sim}{n}_i' = (n_{i1}, n_{i2}, \ldots n_{ir})$ will be assumed to follow the multinomial distribution with parameters n_i and $\underset{\sim}{\pi}_i' = (\pi_{i1}, \pi_{i2}, \ldots, \pi_{ir})$, where π_{ij} represents the probability that a randomly selected element from the i-th population is classified in the j-th response category. Thus, the relevant product multinomial model is

$$\phi = \prod_{i=1}^{s} \{n_i! \prod_{j=1}^{r} [\pi_{ij}^{n_{ij}}/n_{ij}!]\} \qquad (2.1)$$

with the constraint

$$\sum_{j=1}^{r} \pi_{ij} = 1 \text{ for } i = 1,2,\ldots,s. \tag{2.2}$$

Let $\underset{\sim}{p}_i = (n_{\underset{\sim}{i}}/n_i)$ be the (r x 1) vector of observed proportions associated with the sample from the i-th sub-population and let $\underset{\sim}{p}$ be the (sr x 1) compound vector defined by $\underset{\sim}{p}' = (\underset{\sim}{p}_1', \underset{\sim}{p}_2', \ldots, \underset{\sim}{p}_s')$. The vector $\underset{\sim}{p}$ represents the unrestricted maximum likelihood estimator of $\underset{\sim}{\pi}$ where $\underset{\sim}{\pi}' = (\underset{\sim}{\pi}_1', \underset{\sim}{\pi}_2', \ldots, \underset{\sim}{\pi}_s')$. A consistent estimator for the covariance matrix of $\underset{\sim}{p}$ is given by the (sr x sr) block diagonal matrix $\underset{\sim}{V}(\underset{\sim}{p})$ with the matrices

$$\underset{\sim}{V}_i(\underset{\sim}{p}_i) = \frac{1}{n_i}[D_{\underset{\sim}{p}_i} - \underset{\sim}{p}_i\underset{\sim}{p}_i'], \tag{2.3}$$
$$(rxr)$$

for i = 1,2,...,s on the main diagonal; here, $D_{\underset{\sim}{p}_i}$ is an (r x r) diagonal matrix with elements of the vector $\underset{\sim}{p}_i$ on the main diagonal.

Let $F_1(\underset{\sim}{p}), F_2(\underset{\sim}{p}), \ldots, F_u(\underset{\sim}{p})$ be a set of u functions of $\underset{\sim}{p}$ which pertain to some aspect of the relationship between the distribution of the response profiles and the nature of the sub-populations. Each of these functions is assumed to have continuous partial derivatives through order two with respect to the elements of $\underset{\sim}{p}$ within an open region containing $\underset{\sim}{\pi} = \underset{\sim}{E}\{\underset{\sim}{p}\}$. If $\underset{\sim}{F} \equiv \underset{\sim}{F}(\underset{\sim}{p})$ is defined by

$$\underset{\sim}{F}' = [\underset{\sim}{F}(\underset{\sim}{p})]' = [F_1(\underset{\sim}{p}), F_2(\underset{\sim}{p}), \ldots, F_u(\underset{\sim}{p})], \tag{2.4}$$

then a consistent estimator for the covariance matrix of $\underset{\sim}{F}$ is the (u x u) matrix

$$\underset{\sim}{V}_{\underset{\sim}{F}} = \underset{\sim}{H}[\underset{\sim}{V}(\underset{\sim}{p})]\underset{\sim}{H}', \tag{2.5}$$

where $\underset{\sim}{H} = [d\underset{\sim}{F}(\underset{\sim}{x})/d\underset{\sim}{x} \mid \underset{\sim}{x} = \underset{\sim}{p}]$ is the (u x sr) matrix of first partial deriva-

tives of the functions $\underset{\sim}{F}$ evaluated at $\underset{\sim}{p}$. In all applications, the functions comprising $\underset{\sim}{F}$ are chosen so that $\underset{\sim}{V}_{\underset{\sim}{F}}$ is asymptotically nonsingular.

The function vector $\underset{\sim}{F}$ is a consistent estimator of $\underset{\sim}{F}(\underset{\sim}{\pi})$. Hence, the variation among the elements of $\underset{\sim}{F}(\underset{\sim}{\pi})$ can be investigated by fitting linear regression models by the method of weighted least squares. This phase of the analysis can be characterized by writing

$$\underset{\sim}{E}_{\underset{\sim}{A}}\{\underset{\sim}{F}\} \equiv \underset{\sim}{E}_{\underset{\sim}{A}}\{\underset{\sim}{F}(\underset{\sim}{p})\} = \underset{\sim}{F}(\underset{\sim}{\pi}) = \underset{\sim}{X}\ \underset{\sim}{\beta}, \tag{2.6}$$

where $\underset{\sim}{X}$ is a pre-specified (u x t) design (or independent variable) matrix of known coefficients with full rank $t \leq u$, $\underset{\sim}{\beta}$ is an unknown (t x 1) vector of parameters, and "$\underset{\sim}{E}_{\underset{\sim}{A}}$" means "asymptotic expectation."

The model (2.6) implies the existence of a full rank [(u-t) x u] matrix $\underset{\sim}{C}$ which is orthogonal to $\underset{\sim}{X}$ such that

$$\underset{\sim}{E}_{\underset{\sim}{A}}\{\underset{\sim}{C}\ \underset{\sim}{F}\} = \underset{\sim}{C}\ \underset{\sim}{X}\ \underset{\sim}{\beta} = \underset{\sim}{0}_{(u-t),1}, \tag{2.7}$$

where $\underset{\sim}{0}_{(u-t),1}$ is a [(u-t) x 1] vector of 0's. The equations (2.7) represent the set of constraints on the vector $\underset{\sim}{F}(\underset{\sim}{\pi})$ which are implied by the model (2.6). Thus, it follows that a consistent estimator for the covariance matrix of the [(u-t) x 1] transformed functions vector $\underset{\sim}{G} = \underset{\sim}{C}\ \underset{\sim}{F}$ is the [(u-t) x (u-t)] matrix $\underset{\sim}{V}_{\underset{\sim}{G}} = \underset{\sim}{C}\ \underset{\sim}{V}_{\underset{\sim}{F}}\ \underset{\sim}{C}'$. As a result, an appropriate test statistic for the goodness of fit of the model (2.6) is

$$Q = Q(\underset{\sim}{X},\underset{\sim}{F}) = \underset{\sim}{G}'\ \underset{\sim}{V}_{\underset{\sim}{G}}^{-1}\ \underset{\sim}{G} = \underset{\sim}{F}'\ \underset{\sim}{C}'\ [\underset{\sim}{C}\ \underset{\sim}{V}_{\underset{\sim}{F}}\ \underset{\sim}{C}']^{-1}\ \underset{\sim}{C}\ \underset{\sim}{F}, \tag{2.8}$$

which is approximately distributed according to the χ^2-distribution with D.F. = (u-t) if the sample sizes $\{n_i\}$ are sufficiently large that the elements of the vector $\underset{\sim}{F}$ have an approximate multivariate normal distribution as a consequence of Central Limit Theory. Such test statistics are known as

Wald [30] statistics and various aspects of their application to a broad range of problems involving the analysis of multivariate categorical data are discussed in Bhapkar and Koch [6,7] and GSK.

On the other hand, the actual manner in which statistics like (2.8) are applied in practice involves a weighted least squares computational algorithm which is justified on the basis of the fact that

$$Q = \underset{\sim}{F}' \underset{\sim}{C}' [\underset{\sim}{C} \underset{\sim}{V}_{\underset{\sim}{F}} \underset{\sim}{C}']^{-1} \underset{\sim}{C} \underset{\sim}{F} \equiv (\underset{\sim}{F} - \underset{\sim}{X} \underset{\sim}{b})' \underset{\sim}{V}_{\underset{\sim}{F}}^{-1} (\underset{\sim}{F} - \underset{\sim}{X} \underset{\sim}{b}), \tag{2.9}$$

where $\underset{\sim}{b} = (\underset{\sim}{X}' \underset{\sim}{V}_{\underset{\sim}{F}}^{-1} \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{V}_{\underset{\sim}{F}}^{-1} \underset{\sim}{F}$ is a BAN estimator for $\underset{\sim}{\beta}$ based on the linearized modified χ^2_1-statistic of Neyman [28]. In view of this identity both Q and $\underset{\sim}{b}$ are regarded as having reasonable statistical properties in samples which are sufficiently large for applying Central Limit Theory to the functions $\underset{\sim}{F}$. With these considerations in mind, it then can be noted that

$$\underset{\sim}{V}_{\underset{\sim}{b}} = (\underset{\sim}{X}' \underset{\sim}{V}_{\underset{\sim}{F}}^{-1} \underset{\sim}{X})^{-1} \tag{2.10}$$

is a consistent estimator for the covariance matrix of $\underset{\sim}{b}$.

If the model (2.6) does adequately characterize the vector $\underset{\sim}{F}(\underset{\sim}{\pi})$, tests of linear hypotheses pertaining to the parameters $\underset{\sim}{\beta}$ can be undertaken by standard multiple regression procedures. In particular, for a general hypothesis of the form,

$$H_0: \quad \underset{\sim}{C} \underset{\sim}{\beta} = \underset{\sim}{0}_{q,1}, \tag{2.11}$$

where $\underset{\sim}{C}$ is a known (q x t) matrix of full rank $q \leq t$ and $\underset{\sim}{0}_{q,1}$ is a (q x 1) vector of 0's, a suitable test statistic is

$$\begin{aligned} Q_{\underset{\sim}{C}} = Q_{\underset{\sim}{C}}(\underset{\sim}{X},\underset{\sim}{F}) &= \underset{\sim}{b}' \underset{\sim}{C}' [\underset{\sim}{C} \underset{\sim}{V}_{\underset{\sim}{b}} \underset{\sim}{C}']^{-1} \underset{\sim}{C} \underset{\sim}{b} \\ &= \underset{\sim}{b}' \underset{\sim}{C}' [\underset{\sim}{C} (\underset{\sim}{X}' \underset{\sim}{V}_{\underset{\sim}{F}}^{-1} \underset{\sim}{X})^{-1} \underset{\sim}{C}']^{-1} \underset{\sim}{C} \underset{\sim}{b} \end{aligned} \tag{2.12}$$

which has approximately a chi-square distribution with D.F. = q in large samples under H_0.

In this framework, the test statistic $\underset{\sim}{Q}_C$ reflects the amount by which the goodness of fit Wald statistic (2.8) would increase if the model (2.6) were simplified (or reduced) by substitutions based on the additional constraints implied by (2.11). Thus, these methods permit the total variation within $\underset{\sim}{F}(\underset{\sim}{\pi})$ to be partitioned into specific sources and hence represent a statistically valid analysis of variance for the corresponding estimator functions $\underset{\sim}{F}$.

Predicted values for $\underset{\sim}{F}(\underset{\sim}{\pi})$ based on the model (2.6) can be calculated from

$$\hat{\underset{\sim}{F}} = \underset{\sim}{X}\,\underset{\sim}{b} = \underset{\sim}{X}\,(\underset{\sim}{X}'\,\underset{\sim}{V}_{\underset{\sim}{F}}^{-1}\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\,\underset{\sim}{V}_{\underset{\sim}{F}}^{-1}\,\underset{\sim}{F}. \tag{2.13}$$

Consistent estimators for the variances of the elements of $\hat{\underset{\sim}{F}}$ can be obtained from the diagonal elements of

$$\underset{\sim}{V}_{\hat{\underset{\sim}{F}}} = \underset{\sim}{X}\,(\underset{\sim}{X}'\,\underset{\sim}{V}_{\underset{\sim}{F}}^{-1}\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'. \tag{2.14}$$

The predicted values $\hat{\underset{\sim}{F}}$ not only have the advantage of characterizing essentially all the important features of the variation in $\underset{\sim}{F}(\underset{\sim}{\pi})$, but also represent better estimators than the original function statistics $\underset{\sim}{F}$ since they are based on the data from the entire sample as opposed to its component parts. Moreover, they are descriptively advantageous in the sense that they make trends more apparent and permit a clearer interpretation of the relationship between $\underset{\sim}{F}(\underset{\sim}{\pi})$ and the variables comprising the columns of $\underset{\sim}{X}$.

As indicated in GSK, two classes of functions $\underset{\sim}{F}$ pertain to most applications which are currently discussed in the literature. These are linear functions of the type

$$\underset{\sim}{F}(\underset{\sim}{p}) = \underset{\sim}{A}\ \underset{\sim}{p} = \underset{\sim}{a} \tag{2.15}$$

where $\underset{\sim}{A}$ is a known (u x sr) matrix and log-linear functions of the type

$$\underset{\sim}{F}(\underset{\sim}{p}) = \underset{\sim}{K}\ [\underset{\sim\sim}{\log_e}\ (\underset{\sim}{A}\ \underset{\sim}{p})] = \underset{\sim}{K}\ [\underset{\sim\sim}{\log_e}\ (\underset{\sim}{a})] = \underset{\sim}{f} \tag{2.16}$$

where $\underset{\sim}{K}$ is a known (k x u) matrix, $\underset{\sim}{A}$ is as defined in (2.15), and $\underset{\sim\sim}{\log_e}$ transforms a vector to the corresponding vector of natural logarithms. On the basis of (2.5), the estimated covariance matrix for the linear functions in (2.15) is

$$\underset{\sim}{V}_{\underset{\sim}{a}} = \underset{\sim}{A}\ [\underset{\sim}{V}(\underset{\sim}{p})]\ \underset{\sim}{A}'; \tag{2.17}$$

and for the log-linear functions in (2.16) is

$$\underset{\sim}{V}_{\underset{\sim}{f}} = \underset{\sim}{K}\ \underset{\sim}{D}_{\underset{\sim}{a}}^{-1}\ \underset{\sim}{A}\ [\underset{\sim}{V}(\underset{\sim}{p})]\ \underset{\sim}{A}'\ \underset{\sim}{D}_{\underset{\sim}{a}}^{-1}\ \underset{\sim}{K}', \tag{2.18}$$

where $\underset{\sim}{D}_{\underset{\sim}{a}}$ is a diagonal matrix with elements of the vector $\underset{\sim}{a}$ on the main diagonal.

More generally, Forthofer and Koch [12] consider an extended class of compounded logarithmic, exponential, and linear functions which includes complex ratio estimates like rank correlation coefficients, survival rates derived from life tables (see Koch, Johnson, and Tolley [21]), and log log functions (see Freeman, Freeman, and Koch [13]). Finally, Koch and Tolley [23] discuss the application of this general approach to implicitly defined functions of $\underset{\sim}{p}$ in the context of the estimation of bacteria density in serial dilution experiments. Thus, all aspects of this methodology can be directed at implicit functions which are based on maximum likelihood estimation equations corresponding to preliminary or intermediate (as opposed to final) models with a priori assumed validity; in other words, models in which the likelihood (2.1) initially (i.e., prior to any data analysis) satisfies both (2.2) as well as certain other constraints analogous to (2.7).

2.2. Formulation For Repeated Measurement Experiments

For repeated measurement experiments, each subject is observed under each of d different conditions (e.g., treatments, policy questions, time points, etc.), and the corresponding responses are each classified in terms of L categories. Thus, there are $r = L^d$ possible multivariate response profiles.

In accordance with the general framework in Section 2.1, these response profiles will be indexed by a vector subscript $\underset{\sim}{j} = (j_1, j_2, \ldots, j_d)$ where $j_g = 1,2,\ldots,L$ for $g = 1,2,\ldots,d$. As a result, $\pi_{i\underset{\sim}{j}} = \pi_{ij_1,j_2,\ldots,j_d}$ represents the joint probability of response profile $\underset{\sim}{j}$ for randomly selected subjects from the i-th sub-population. If there are no differences among the sub-populations, then the $\{\pi_{i\underset{\sim}{j}}\}$ satisfy the hypothesis,

$$H_{SJ}: \quad \pi_{1\underset{\sim}{j}} = \pi_{2\underset{\sim}{j}} = \ldots = \pi_{s\underset{\sim}{j}} \quad \text{for all } \underset{\sim}{j}; \tag{2.19}$$

similarly, if there are no differences among the d conditions in a strict overall sense, then the $\{\pi_{i\underset{\sim}{j}}\}$ satisfy the hypothesis of <u>total symmetry</u>

$$H_{CJ}: \quad \pi_{i\underset{\sim}{j}} = \pi_{i,\underset{\sim}{z}(\underset{\sim}{j})} \quad \text{for all } \underset{\sim}{j} \text{ and } i = 1,2,\ldots,s, \tag{2.20}$$

where $\underset{\sim}{z}(\underset{\sim}{j})$ is any permutation of $\underset{\sim}{j}$. These hypotheses, together with certain no interaction formulations like those discussed in Bhapkar and Koch [6,7] can be tested by means of the general GSK methodology outlined in Section 2.1. However, these statistical tests are of limited practical interest since the quantities which provide the most obvious indication of the relative effects associated with the respective measurement conditions and sub-populations are the first order marginal probabilities

$$\phi_{igk} = \sum_{\underset{\sim}{j} \text{ with } j_g=k} \cdots \sum \pi_{ij_1,j_2,\ldots,j_d} \qquad \begin{matrix} i = 1,2,\ldots,s \\ \text{for } g = 1,2,\ldots,d \\ k = 1,2,\ldots,L \end{matrix} . \tag{2.21}$$

Here, ϕ_{igk} represents the probability of the k-th response category for the g-th condition in the i-th sub-population. If there are no differences among the sub-populations in the sense of (2.19), then the $\{\phi_{igk}\}$ satisfy the hypothesis,

$$H_{SM}: \quad \phi_{1gk} = \phi_{2gk} = \ldots = \phi_{sgk} \quad \text{for } \begin{matrix} g = 1,2,\ldots,d \\ k = 1,2,\ldots,L \end{matrix} ; \tag{2.22}$$

while if there are no differences among the conditions in the sense of (2.20), then the $\{\phi_{igk}\}$ satisfy the hypothesis of first order marginal symmetry (homogeneity)

$$H_{CM}: \quad \phi_{i1k} = \phi_{i2k} = \ldots = \phi_{idk} \quad \text{for } \begin{matrix} i = 1,2,\ldots,s \\ k = 1,2,\ldots,L \end{matrix} . \tag{2.23}$$

Thus, it logically follows that

$$H_{SM} \text{ false} \longrightarrow H_{SJ} \text{ false} \tag{2.24}$$

$$H_{CM} \text{ false} \longrightarrow H_{CJ} \text{ false} \tag{2.25}$$

which means that H_{SJ} (or H_{CJ}) must be rejected whenever H_{SM} (or H_{CM}) is rejected. Alternatively, situations where H_{SM} (or H_{CM}) is true, regardless of whether H_{SJ} (or H_{CJ}) is true, would be interpreted as involving no gross (macro) differences among the sub-populations (conditions). Thus, the hypotheses H_{SM} and H_{CM} are of interest in their own right with respect to investigating the effects of sub-populations and conditions on the response distribution under study.

Moreover, H_{SM} (or H_{CM}) have greater logical relevance for this purpose than H_{SJ} (or H_{CJ}). This point of view can be justified by noting that the additional constraints on the π_{ij} which are associated with H_{SJ} (or H_{CJ}) but not H_{SM} (or H_{CM}) involve relatively complicated equations which do not have a straightforward interpretation. Also, these constraints are partially confounded with those which pertain to the patterns of association (i.e., lack

of independence) among the responses to the respective measurement conditions since the hypothesis,

$$H_{JI}: \quad \pi_{ij_1,j_2,\ldots,j_d} = \prod_{g=1}^{d} \phi_{igj_g} \qquad \text{for all } \underset{\sim}{j} \text{ and } i = 1,2,\ldots,s, \quad (2.26)$$

of joint independence in conjunction with H_{SM} (or H_{CM}) necessarily imply H_{SJ} (or H_{CJ}). In repeated measurement experiments, such association is present to the extent that underlying subject effects cause measurements under two different conditions on the same subject to be relatively more similar than corresponding measurements on different subjects. This type of dependence among responses may be even further complicated by factors pertaining to the adjacency of measurements in time and space (e.g., in certain longitudinal studies, measurements from adjacent time points are often more related to each other than those for distant time points). Since such association can be inherently present without having any real bearing on the interpretation of sub-population and condition effects, it then becomes reasonably appropriate to filter the corresponding constraints out of H_{SJ} and H_{CJ}; but this, of course, yields H_{SM} and H_{CM}. Thus, for most practical purposes, consideration can often be entirely restricted to H_{SM} and H_{CM} in repeated measurement experiments (or surveys). Finally, statistical tests directed at H_{SM} and H_{CM} tend to have better asymptotic convergence properties (with respect to the valid application of Central Limit Theory arguments) in moderately large samples than those directed at H_{SJ} and H_{CJ} since they involve fewer constraint equations (i.e., degrees of freedom). This aspect of the analysis of repeated measurement experiments (or surveys) will be discussed in further detail in Section 2.3. Otherwise, the remainder of this section will be concerned with additional hypotheses involving the

$\{\phi_{igk}\}$ and their subsequent analysis.

The hypothesis of no interaction between conditions and sub-populations can be formulated for the $\{\phi_{igk}\}$ in an additive sense in terms of the model

$$H_{AM}: \quad \phi_{igk} = \mu_k + \xi_{i*k} + \tau_{*gk} \quad \text{for } \begin{matrix} i = 1,2,\ldots,s \\ g = 1,2,\ldots,d, \\ k = 1,2,\ldots,L \end{matrix} \tag{2.27}$$

where μ_k is an overall mean associated with the k-th response category, ξ_{i*k} is an effect due to the i-th sub-population, and τ_{*gk} is an effect due to the g-th condition and where it is usually understood that the $\{\mu_k\}$, $\{\xi_{i*k}\}$, and $\{\tau_{*gk}\}$ satisfy the following types of constraints

$$\sum_{k=1}^{L} \mu_k = 1, \quad \sum_{k=1}^{L} \xi_{i*k} = 0, \quad \sum_{k=1}^{L} \tau_{*gk} = 0,$$

$$\sum_{i=1}^{s} \xi_{i*k} = 0, \quad \sum_{g=1}^{d} \tau_{*gk} = 0, \tag{2.28}$$

as a consequence of (2.2) and certain parameter identifiability considerations (i.e., removal of logical redundancies among parameters). If the model (2.27) is appropriate for a particular experimental situation, then the hypothesis H_{SM} in (2.22) implies

$$H_{SM|AM}: \quad \xi_{1*k} = \xi_{2*k} = \ldots = \xi_{s*k} = 0 \quad \text{for } k = 1,2,\ldots,L, \tag{2.29}$$

and the hypothesis H_{CM} in (2.23) implies

$$H_{CM|AM}: \quad \tau_{*1k} = \tau_{*2k} = \ldots = \tau_{*dk} = 0 \quad \text{for } k = 1,2,\ldots,L. \tag{2.30}$$

Other hypotheses of interest can be formulated in an analogous manner in order to account for any inherent structure associated with the s sub-populations or the d conditions (e.g., trends or separate roles of two or more underlying factors). Also, in some cases, it may be more appropriate to

work with multiplicative no interaction models in which the additivity relationships like (2.27) pertain to the $\{\log \phi_{igk}\}$ (for additional discussion see, Bhapkar and Koch [6,7] and Darroch [9]. An example of this type of situation is discussed in Section 3.3.

All of these consideratiosn can be extended somewhat further if the response categories $k = 1,2,\ldots,L$ are ordinally scaled with progressively larger intensities. In this event, the effects of the respective sub-populations and measurement conditions can be compared in terms of one or more of the following cumulative marginal probability functions:

$$\theta_{igk} = \sum_{\ell=(k+1)}^{L} \phi_{ig\ell} \quad \text{for } \begin{matrix} i = 1,2,\ldots,s \\ g = 1,2,\ldots,d, \end{matrix} \tag{2.31}$$

where $k = 1,2,\ldots,(L-1)$. Here θ_{igk} represents the probability that the response intensity is strictly greater than the k-th category for the g-th condition and the i-th population. If the relative substantive importance of the θ_{igk} can be characterized by non-negative weights $W_1,W_2,\ldots,W_{(L-1)}$ (at least one of which is non-zero), then attention can be directed at summary indexes:

$$\eta_{ig} = \sum_{k=1}^{(L-1)} W_k \theta_{igk} = \sum_{k=1}^{L} a_k \phi_{igk} \quad \text{for } \begin{matrix} i = 1,2,\ldots,s \\ g = 1,2,\ldots,d, \end{matrix} \tag{2.32}$$

where $a_1 = 0$ and $a_k = \sum_{\ell=1}^{(k-1)} W_\ell$ for $k = 2,\ldots,L$. Thus, η_{ig} can also be regarded as a mean score for the g-th condition and the i-th sub-population with respect to an underlying numerical scaling $a_1,a_2,\ldots,a_L$ of the L response categories. In this context, the $\{\eta_{ig}\}$ are equivalent to mean scores derived from strictly quantitatively scaled response variables as discussed by Bhapkar [5]. Hence, if there are no differences among the sub-populations

in the sense of H_{SM} in (2.22), then the $\{\eta_{ig}\}$ satisfy the hypothesis:

$$H_{SAM}: \quad \eta_{1g} = \eta_{2g} = \ldots = \eta_{sg} \quad \text{for } g = 1,2,\ldots,d; \tag{2.33}$$

while if there are no differences among the conditions in the sense of H_{CM} in (2.23), then the $\{\eta_{ig}\}$ satisfy the hypothesis

$$H_{CAM}: \quad \eta_{i1} = \eta_{i2} = \ldots = \eta_{id} \quad \text{for } i = 1,2,\ldots,s. \tag{2.34}$$

The logical relationship between H_{SAM} (or H_{CAM}) and H_{SM} (or H_{CM}) is directly analogous to the relationship previously discussed between H_{SM} (or H_{CM}) and H_{SJ} (or H_{CJ}). In particular, it follows that

$$H_{SAM} \text{ false} \longrightarrow H_{SM} \text{ false} \tag{2.35}$$

$$H_{CAM} \text{ false} \longrightarrow H_{CM} \text{ false} \tag{2.36}$$

which means that H_{SM} (or H_{CM}) must be rejected whenever H_{SAM} (or H_{CAM}) is rejected. Alternatively, situations where H_{SAM} (or H_{CAM}) is true, regardless of whether H_{SM} (or H_{CM}) is true, would be interpreted as involving no average gross (macro) differences among the sub-populations (conditions). Thus, the hpotheses H_{SAM} and H_{CAM} are of interest in their own right with respect to investigating the effects of sub-populations and conditions on the response distribution under study. Moreover, H_{SAM} (or H_{CAM}) have greater logical relevance for this purpose than H_{SM} (or H_{CM}), particularly if the scores $a_1,a_2,\ldots,a_L$ provide a reasonably meaningful and valid quantitative measure of the intensity of the response. Further justification for this point of view can be derived by noting that the additional constraints on the $\{\phi_{igk}\}$ which are associated with H_{SM} (or H_{CM}) but not H_{SAM} (or H_{CAM}) involve relatively complicated equations which are difficult to interpret because they pertain more to the "shape" of the response distribution as

opposed to its "location." For example, if H_{SM} (or H_{CM}) are false in the sense that differences among the sub-populations (or conditions) are progressively increasing with respect to the probabilities of the response categories k = 1,2,...,L so that the "weaker (smaller k) responses" are more likely for some sub-populations (or conditions) while the "stronger (larger k) responses" are more likely in others, then H_{SAM} (or H_{CAM}) will be correspondingly contradicted since the mean scores $\{\eta_{ig}\}$ based on (2.32) will tend to be smaller in the former and larger in the latter. Although it is preferable for the weights $W_1, W_2, \ldots, W_{(L-1)}$ (or alternatively, the scores $a_1, a_2, \ldots, a_L$) to have a strong substantive scaling basis, these same conclusions still apply to almost the same extent if these weights are equal; i.e.,

$$W_1 = W_2 = \ldots = W_{(L-1)} = 1, \tag{2.37}$$

so that

$$a_1 = 0,\ a_1 = 1,\ a_2 = 2,\ \ldots,\ a_L = (L-1) \tag{2.38}$$

represent an equally spaced scale. Thus, for most practical purposes, consideration can often be restricted to some reasonably appropriate formulation of H_{SAM} and H_{CAM} in repeated measurement experiments (or surveys) involving ordinally scaled response categories. Finally, statistical tests directed at H_{SAM} and H_{CAM} tend to have better asymptotic convergence properties (with respect to the valid application of Central Limit Theory arguments) in moderately large samples than those directed at H_{SM} and H_{CM} since they involve fewer constraint equations (i.e., degrees of freedom). As mentioned previously, this aspect of the analysis of repeated measurement experiments (or surveys) will be discussed in further detail in Section 2.3.

Other hypotheses can be formulated in terms of the $\{\eta_{ig}\}$. For example, if there is no interaction between conditions and sub-populations in the

sense of (2.27), then the $\{\eta_{ig}\}$ satisfy the model

$$H_{AMA}: \quad \eta_{ig} = \bar{\mu}_{\cdot} + \bar{\xi}_{i*\cdot} + \bar{\tau}_{*g\cdot} \quad \text{for } \begin{matrix} i = 1,2,\ldots,s \\ g = 1,2,\ldots,d \end{matrix}, \tag{2.39}$$

with the constraints

$$\sum_{i=1}^{s} \bar{\xi}_{i*\cdot} = 0, \quad \sum_{g=1}^{d} \bar{\tau}_{*g\cdot} = 0. \tag{2.40}$$

Here $\bar{\mu}_{\cdot}$ is an overall mean, $\bar{\xi}_{i*\cdot}$ is an effect due to the i-th sub-population, and $\bar{\tau}_{*g\cdot}$ is an effect due to the g-th condition. If the model (2.39) is appropriate for a particular experimental situation, then the hypothesis H_{SAM} in (2.33) implies

$$H_{SAM|AMA}: \quad \xi_{1*\cdot} = \xi_{2*\cdot} = \ldots = \xi_{s*\cdot} = 0, \tag{2.41}$$

and the hypothesis H_{CAM} in (2.34) implies

$$H_{CAM|AMA}: \quad \tau_{*1\cdot} = \tau_{*2\cdot} = \ldots = \tau_{*d\cdot} = 0. \tag{2.42}$$

The hypothesis (2.35) is directly analogous to the hypothesis of no whole-plot by split-plot interaction in standard split-plot (or repeated measurement) experiments involving a univariate quantitatively scaled response variable. Similarly, the hypotheses (2.41) and (2.42) are analogous to the hypothesis of no whole-plot main effects and the hypothesis of no split-plot main effects. Thus, all the models and hypotheses considered in this section represent straightforward extensions of split-plot (repeated measurement) analysis of variance to multivariate categorical data.

Statistical tests for the hypotheses previously discussed in this section as well as the estimation of corresponding model parameters can be undertaken within the general framework of Section 2.1 by specifying the appropriate set of functions $\underset{\sim}{F}$ and operator matrices (eg., $\underset{\sim}{A}$, $\underset{\sim}{K}$, $\underset{\sim}{X}$, and $\underset{\sim}{C}$'s). For this purpose,

1. The $\underset{\sim}{A}$ matrix specifies additive operations and can be used to generate linear functions of the observed compound proportion vector $\underset{\sim}{p}$ which are unbiased estimators for certain sets of cell probabilities like the $\{\pi_{ij}\}$, marginal probabilities like the $\{\phi_{igk}\}$ and/or mean scores like the $\{\eta_{ig}\}$. In this regard, the rows of $\underset{\sim}{A}$ represent the respective functions while the columns are the corresponding coefficients of the elements of the compound vector $\underset{\sim}{p}$ which produce such functions.
2. The $\underset{\sim}{K}$ matrix specifies multiplicative operations (i.e., additive on the $\log_e$ scale) and in combination with the appropriate $\underset{\sim}{A}$ matrix can be used to generate asymptotically unbiased estimators for functions involving the $\{\log_e \pi_{ij}\}$ or the $\{\log_e \phi_{igk}\}$.
3. The $\underset{\sim}{X}$ matrix has the same role here as it does in standard multiple regression; i.e., it indicates the manner in which the variation among a certain set of functions $\underset{\sim}{F}(\underset{\sim}{\pi})$ of cell probabilities can be characterized as linear functions of a smaller set of unknown parameters. Thus, the columns of $\underset{\sim}{X}$ correspond to the components of $\underset{\sim}{\beta}$; and the rows of $\underset{\sim}{X}$ specify the appropriate linear combination of the components of the parameter vector $\underset{\sim}{\beta}$ which apply to the respective elements of $\underset{\sim}{F}(\underset{\sim}{\pi})$.
4. The $\underset{\sim}{C}$ matrix also has the same role as in standard multiple regression; i.e., it indicates which linear functions of the param-

eters are equal to zero in accordance with a particular hypothesis which is to be tested.

Although the construction of these matrices is reasonably straightforward for repeated measurement experiments (or surveys) associated with specific applications, the corresponding expressions for the general framework which has been discussed here are uninformatively complicated and tedious. For this reason, such mathematical details will be discussed in terms of the special case where there are s = 2 sub-populations, d = 2 measurement conditions, and an ordinal response with L = 3 categories which can be quantitatively scaled in terms of scores a_1, a_2, a_3 respectively. The functions required to test H_{SJ} in (2.19) and H_{CJ} in (2.20) can be generated (with proper account given to (2.2)) by using

$$\underset{16x18}{\underset{\sim}{A}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \otimes \underset{\sim}{I}_2, \tag{2.43}$$

where $\otimes$ denotes Kronecker product and $\underset{\sim}{I}_u$ denotes the (u x u) identity matrix.

If $\underset{\sim}{X} = \underset{\sim}{I}_{16}$ is used, then

$$\underset{8x16}{\underset{\sim}{C}_{SJ}} = [\underset{\sim}{I}_8 \vdots -\underset{\sim}{I}_8] \tag{2.44}$$

generates a test statistic for H_{SJ} via (2.12) while

$$\underset{6x16}{\underset{\sim}{C}_{CJ}} = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix} \otimes \underset{\sim}{I}_2 \tag{2.45}$$

similarly generates a test statistic for H_{CJ}.

The functions required to test H_{SM} in (2.22) and H_{CM} in (2.23) can be generated (with proper account given to (2.2)) by using

$$\underset{8\times 18}{\underset{\sim}{A}} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \otimes \underset{\sim}{I}_2. \tag{2.46}$$

If $\underset{\sim}{X} = \underset{\sim}{I}_8$ is used, then

$$\underset{4\times 8}{\underset{\sim}{C}_{SM}} = [\underset{\sim}{I}_4 \,\vdots\, -\underset{\sim}{I}_4] \tag{2.47}$$

produces a test statistic for H_{SM} via (2.12) while

$$\underset{4\times 8}{\underset{\sim}{C}_{CM}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \otimes \underset{\sim}{I}_2 \tag{2.48}$$

similarly produces a test statistic for H_{CM}. The hypothesis H_{AM} of no interaction in (2.27) can be tested by the goodness of fit statistic (2.8) which corresponds to the model

$$\underset{8\times 6}{\underset{\sim}{X}} = \underset{\sim}{I}_2 \otimes \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \quad \text{with} \quad \underset{6\times 1}{\underset{\sim}{\beta}} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \xi_{1*1} \\ \xi_{1*2} \\ \tau_{*11} \\ \tau_{*12} \end{bmatrix}. \tag{2.49}$$

If the model (2.49) corresponding to H_{AM} can be presumed to hold, then the hypotheses $H_{SM|AM}$ and $H_{CM|AM}$ can be tested by using

$$\underset{2\times 6}{\underset{\sim}{C}_{SM|AM}} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \underset{2\times 6}{\underset{\sim}{C}_{CM|AM}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.50}$$

respectively.

The functions required to test H_{SAM} in (2.33) and H_{CAM} in (2.34) can be generated by using

$$\underset{4x18}{\underset{\sim}{A}} = \begin{bmatrix} a_1 & a_1 & a_1 & a_2 & a_2 & a_2 & a_3 & a_3 & a_3 \\ a_1 & a_2 & a_3 & a_1 & a_2 & a_3 & a_1 & a_2 & a_3 \end{bmatrix} \otimes \underset{\sim}{I}_2. \tag{2.51}$$

If $\underset{\sim}{X} = \underset{\sim}{I}_4$ is used, then

$$\underset{2x4}{\underset{\sim}{C}_{SAM}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \underset{2x4}{\underset{\sim}{C}_{CAM}} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \tag{2.52}$$

produce test statistics via (2.12) for H_{SAM} and H_{CAM} respectively. The hypothesis H_{AMA} of no interaction in (2.39) can be tested by the goodness of fit statistic (2.8) which corresponds to the model

$$\underset{4x3}{\underset{\sim}{X}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \quad \text{with } \underset{\sim}{\beta} = \begin{bmatrix} \bar{\mu} \\ \bar{\xi}_{1*\cdot} \\ \bar{\tau}_{*1\cdot} \end{bmatrix}. \tag{2.53}$$

If the model (2.53) corresponding to H_{AMA} can be presumed to hold, then the hypotheses $H_{SAM|AMA}$ and $H_{CAM|AMA}$ can be tested by using

$$\underset{1x3}{\underset{\sim}{C}_{SAM|AMA}} = [0\ 1\ 0] \quad \text{and} \quad \underset{1x3}{\underset{\sim}{C}_{CAM|AMA}} = [0\ 0\ 1] \tag{2.54}$$

respectively.

Other aspects of the specification of these operator matrices will be discussed in terms of the examples in Section 3. In particular, the example in Section 3.3 illustrates the use of the $\underset{\sim}{K}$-matrix operator (in conjunction with an $\underset{\sim}{A}$-matrix analogous in structure to (2.46)) to form appropriate log-linear functions of the $\{\phi_{igk}\}$.

2.3. Special Considerations For Large Tables And Zero Cells

Unless both the number of conditions d and the number of response categories L are small (eg., $d \leq 3$, $L \leq 3$), the number of possible multivariate response profiles $r = L^d$ can become very large. In this event, there are two sources of potential difficulty:

i. The manipulation of very large contingency tables and operator matrices can become very expensive with respect to both computer time as well as programmer effort.

ii. For each sub-population $i = 1,2,\ldots,s$, some of the r possible response profiles $\underset{\sim}{j}$ will not necessarily be observed in the respective samples so that the corresponding cell frequencies $n_{i\underset{\sim}{j}}$ are zero.

As noted in Section 2.2, the hypotheses of primary interest in repeated measurement experiments (or surveys) involve the first order marginal probabilities $\{\phi_{igk}\}$. Unbiased estimators for these quantities can be obtained by applying an $\underset{\sim}{A}$ matrix analogous to (2.46) to the hypothetical data in Table 1. Alternatively, these same identical estimators can also be generated by forming appropriate functions of the raw data associated with each subject and then computing the corresponding across subject arithmetic means within each of the respective sub-populations. In particular, let

$$y_{igkm} = \begin{cases} 1 & \text{if the m-th subject in the sample from the i-th sub-population has the response for the g-th condition classified into the k-th category} \\ 0 & \text{otherwise} \end{cases} \qquad (2.55)$$

where $i = 1,2,\ldots,s$; $g = 1,2,\ldots,d$; $k = 1,2,\ldots,(L-1)$; and $m = 1,2,\ldots,n_i$. If

$$\bar{y}_{igk} = \frac{1}{n_i} \sum_{m=1}^{n_i} y_{igkm}, \qquad (2.56)$$

$$\bar{\underset{\sim}{y}}_i' = (\bar{y}_{i11},\ldots,\bar{y}_{i1,(L-1)},\ldots,\bar{y}_{id1},\ldots,\bar{y}_{id,(L-1)}), \tag{2.57}$$

$$\bar{\underset{\sim}{y}} = (\bar{\underset{\sim}{y}}_1',\bar{\underset{\sim}{y}}_2',\ldots,\bar{\underset{\sim}{y}}_s'), \tag{2.58}$$

then $\bar{\underset{\sim}{y}} \equiv \underset{\sim}{a}$ where $\underset{\sim}{a} = \underset{\sim}{A}\,\underset{\sim}{p}$ as defined in (2.15). Moreover, the covariance matrix $\underset{\sim}{V}_{\bar{\underset{\sim}{y}}}$ of $\bar{\underset{\sim}{y}}$ is block diagonal with the matrices

$$\underset{\sim}{V}_{\bar{\underset{\sim}{y}}_i} = \frac{1}{n_i^2} \sum_{m=1}^{n_i} (\underset{\sim}{y}_{im} - \bar{\underset{\sim}{y}}_i)(\underset{\sim}{y}_{im} - \bar{\underset{\sim}{y}}_i)' \tag{2.59}$$

where

$$\underset{\sim}{y}_{im}' = (y_{i11m},\ldots,y_{i1,(L-1),m},\ldots,y_{id1m},\ldots,y_{id,(L-1),m}) \tag{2.60}$$

on the main diagonal; and

$$\underset{\sim}{V}_{\bar{\underset{\sim}{y}}} \equiv \underset{\sim}{V}_{\underset{\sim}{a}} = \underset{\sim}{A}\,\underset{\sim}{V}\,[(\underset{\sim}{p})]\,\underset{\sim}{A}' \tag{2.61}$$

in (2.17). Since sd(L-1) is usually moderate in size, this method of computation of the estimators for the first order marginal probabilities is reasonably straightforward and efficient. In addition, it can be readily linked with other algorithms for performing logarithmic operations as shown in (2.16) and (2.18) as well as weighted least squares regression analysis. Finally, this approach can be easily extended to deal with situations where the response is ordinal and can be quantitatively scaled in terms of scores $a_1,a_2,\ldots,a_L$. In this case, the functions formed from the raw data on each subject are simply the observed values of the scored responses to each condition; i.e.,

$$y_{igm} = a_k \qquad \begin{array}{l}\text{if m-th subject in the sample from the i-th}\\ \text{sub-population has the response for the g-th}\\ \text{condition classified into the k-th category}\end{array} \tag{2.62}$$

for $i = 1,2,\ldots,s$; $g = 1,2,\ldots,d$; and $m = 1,2,\ldots,n_i$. Then

$$\bar{y}_{ig} = \frac{1}{n_i} \sum_{m=1}^{n_i} y_{igm} \tag{2.63}$$

represents the observed mean score associated with the i-th sub-population and g-th condition; and both the $\{\bar{y}_{ig}\}$ and the corresponding covariance matrix as obtained by operations like those in (2.59) are identical to what would be obtained by applying an $\underset{\sim}{A}$ matrix operation analogous to (2.51) to the hypothetical data in Table 1. Thus, this computational strategy represents an effective way of dealing with the problems associated with (i).

For repeated measurement experiments, the potential tendency for some of the n_{ij} to be zero does not cause any real problems except when such zero frequencies induce singularities in the estimated covariance matrix $\underset{\sim}{V}_{\underset{\sim}{F}}$ in (2.5) for the function vector $\underset{\sim}{F}$ which is to be analyzed, or otherwise restrict the extent to which Central Limit Theory arguments can be applied to the distribution of $\underset{\sim}{F}$. With respect to the hypotheses H_{SJ} and H_{CJ}, Central Limit Theory arguments cannot be applied with confidence to the estimators $p_{ij} = (n_{ij}/n_i)$ of the π_{ij} unless most of the $n_{ij} \geq 5$.

This condition, however, is usually not satisfied for situations involving moderately large samples (i.e., $n_i \geq 100$) except when both L and d are small (i.e., $d \leq 3$, $L \leq 3$). Thus, in such cases, the potential presence of many zero frequencies implies that these hypotheses cannot be validly tested by the general methodological approach given in this paper.

On the other hand, if attention can be restricted to hypotheses like H_{SM} and H_{CM} or H_{SAM} and H_{CAM} which involve the first order marginal probabilities $\{\phi_{igk}\}$, then the sample size requirements are considerably less severe. In particular, Central Limit Theory can be applied with reasonable

validity to estimates of the ϕ_{igk} obtained either by $\underset{\sim}{A}$ matrix operations like (2.46) or by direct computations like (2.56) provided the overall within sub-population sample sizes $n_i \geq 25$ and most of the first order marginal frequencies

$$n_{igk} = \sum_{\underset{\sim}{j} \text{ with } j_g=k} \cdots \sum n_{ij_1,j_2,\ldots,j_d} \geq 5. \tag{2.64}$$

Since $\sum_{k=1}^{L} n_{igk} = n_i$, the conditions (2.64) tend to hold in most situations where the sample sizes $n_i \geq 25$ and where the number of response categories is small (i.e., $L \leq 3$).

However, if L is not small, then cases where (2.64) tends not to hold can be handled by pooling certain of the response categories together so that (2.64) is roughly satisfied by the pooled frequencies. Of course, attempts should be made to base such pooling on conceptual similarities among the response categories as well as to prevent such pooling from masking any obvious differences among either sub-populations or among conditions. Otherwise, the logical relationship between the interpretation of the hypotheses H_{SM} (or H_{CM}) for the pooled categories vs. the original categories is analogous to the logical relationship between H_{SAM} (or H_{CAM}) and H_{SM} (or H_{CM}). Alternatively, many of the situations where L is large involve ordinally scaled response categories, and hence, the quantities of interest are estimates of mean score functions $\{\eta_{ig}\}$ which are obtained either by $\underset{\sim}{A}$ matrix operations like (2.51) or by direct computations like (2.63). Central Limit Theory can be applied with reasonable validity to these statistics provided the overall within sub-population sample sizes, $n_i \geq 25$, and the corresponding observed response distributions for each measurement condition are not degen-

erate in the sense of being concentrated in a single response category. This latter requirement is adequately satisfied for most purposes if there exist response categories k(i,g) for each sub-population x condition combination such that,

$$\sum_{\ell=1}^{k(i,g)} n_{igk} \geq 5, \qquad \sum_{\ell=(k(i,g)+1)}^{L} n_{igk} \geq 5. \tag{2.65}$$

The previous remarks deal with the extent to which the presence of zero frequencies affects the types of functions to which Central Limit Theory arguments can be applied in a valid manner. Moreover, when such functions are chosen judiciously, the resulting estimated covariance matrix $\underset{\sim}{V}_{\underset{\sim}{F}}$ usually will be nonsingular . However, if $\underset{\sim}{V}_{\underset{\sim}{F}}$ is singular, then a number of technical difficulties arise in applying the methodology of Section 2.1 because matrix inversion of $\underset{\sim}{V}_{\underset{\sim}{F}}$ can no longer be performed.

There are several ways of dealing with these potential singularities. One approach is simply either to delete functions from $\underset{\sim}{F}$ or to pool two or more functions within $\underset{\sim}{F}$ together by forming appropriate averages in such a way that the estimated covariance matrix $\underset{\sim}{V}_{\underset{\sim}{F}_R}$ for the reduced function vector $\underset{\sim}{F}_R$ is nonsingular. On the other hand, when the sample sizes n_i are very large (e.g., $n_i \geq 1000$), such reductions in the dimension of $\underset{\sim}{F}$ may be unnecessarily conservative. For these situations, the vector $\underset{\sim}{F}$ can be analyzed by the methodology of Section 2.1 provided that the estimated covariance matrix $\underset{\sim}{V}_{\underset{\sim}{F}}$ is replaced by a suitably similar, nonsingular, pseudo estimated covariance matrix $\underset{\sim}{V}^{*}_{\underset{\sim}{F}}$.

If both L and d are small (e.g., $d \leq 3$, $L \leq 3$), such a nonsingular matrix $\underset{\sim}{V}^{*}_{\underset{\sim}{F}}$ can be obtained by applying the same operations used to form $\underset{\sim}{V}_{\underset{\sim}{F}}$ to the pseudo full (s x r) contingency table (like Table 1) in which

some of the "0" frequencies are replaced by an appropriate small number like "($\frac{1}{2}$)". The motivation behind this strategy are the findings of Berkson [4] who investigated its effects on estimation involving the logit (i.e., natural logarithm of the ratio of the proportions) associated with r = 2 responses. For the case where $r = L^d$ is large, the properties of this rule are largely unknown; and thus the recommendation here is not to replace all "0" frequencies by ($\frac{1}{2}$), but to modify only those which are necessary to the construction of a nonsingular estimated covariance matrix $\underset{\sim}{V}^*_{\underset{\sim}{F}}$ for $\underset{\sim}{F}$. Other aspects of the analysis of $\underset{\sim}{F}$ then proceed as described in Section 2.1, except that $\underset{\sim}{V}^*_{\underset{\sim}{F}}$ is used as the estimated covariance matrix instead of $\underset{\sim}{V}_{\underset{\sim}{F}}$. Finally, in some cases the separate determination of $\underset{\sim}{F}$ and $\underset{\sim}{V}^*_{\underset{\sim}{F}}$ from different "observed" contingency tables and subsequent linkage is a computational nuisance. For this reason, the pseudo function vector $\underset{\sim}{F}^*$, which is obtained from the pseudo contingency table by the same operations used to form $\underset{\sim}{F}$, is often analyzed with respect to $\underset{\sim}{V}^*_{\underset{\sim}{F}}$ as opposed to $\underset{\sim}{F}$ since this only involves computations on the pseudo contingency table. For most practical purposes, the analyses of $\underset{\sim}{F}$ and $\underset{\sim}{F}^*$ are essentially equivalent and the choice between them is mostly a matter of personal taste.

The previously described method for constructing $\underset{\sim}{V}^*_{\underset{\sim}{F}}$ is not really practical when $d \geq 4$, $L \geq 4$ because the full contingency table is too large to manipulate efficiently. Moreover, in these cases, attention is primarily directed at the hypotheses pertaining to the first order marginal probabilities $\{\phi_{igk}\}$ for which unbiased estimators $\{\bar{y}_{igk}\}$ are determined via (2.55) and (2.56) and the corresponding estimated covariance matrix $\underset{\sim}{V}_{\underset{\sim}{\bar{y}}}$ is determined via (2.59). Unfortunately, if $\underset{\sim}{V}_{\underset{\sim}{\bar{y}}}$ is singular, considerable effort may be required within this computational framework to construct a nonsingular, pseudo estimated covariance matrix $\underset{\sim}{V}^*_{\underset{\sim}{\bar{y}}}$ which can be used in the analysis

of the functions $\underset{\sim}{F} = \underset{\sim}{\bar{y}}$. For this purpose, the basic strategy is to continue to augment the raw data associated with the respective subjects with pseudo data vectors for additional pseudo subjects which are then given a smaller weight like ($\frac{1}{2}$) than the raw data vectors associated with the real subjects in the computations (2.59) until the required nonsingular matrix $\underset{\sim}{\underset{\sim}{V}}^{*}_{\underset{\sim}{y}}$ is determined. If the singularities in $\underset{\sim}{\underset{\sim}{V}}_{\underset{\sim}{y}}$ are relatively obvious (eg., the estimated variances for certain of the individual $\bar{y}_{igk}$ are zero), this method of determining $\underset{\sim}{\underset{\sim}{V}}^{*}_{\underset{\sim}{y}}$ can be undertaken in a reasonably straightforward manner which involves a minimum number of pseudo data vectors. However, if the singularities in $\underset{\sim}{\underset{\sim}{V}}_{\underset{\sim}{y}}$ are more subtle (eg., the estimated variances for certain second order differences $(y_{igk}-y_{ig'k}-y_{igk'}+y_{ig'k'})$ are zero), then the efficient construction of $\underset{\sim}{\underset{\sim}{V}}^{*}_{\underset{\sim}{y}}$ requires some additional considerations. For this purpose, let $\underset{\sim}{C}$ be a block diagonal matrix with matrices $\underset{\sim}{C}_i$ on the main diagonal where the rows of $\underset{\sim}{C}_i$ constitute a basis for the vector space orthogonal to the columns of $\underset{\sim}{\underset{\sim}{V}}_{\underset{\sim}{y}_i}$ in (2.59). Then the estimated covariance matrix for the functions $\underset{\sim}{G}_i = \underset{\sim}{C}_i\underset{\sim}{\bar{y}}_i$ is $\underset{\sim}{V}_{\underset{\sim}{G}_i} = \underset{\sim}{C}_i\underset{\sim}{\underset{\sim}{V}}_{\underset{\sim}{y}_i}\underset{\sim}{C}_i' = \underset{\sim}{0}$ which means that the observed data vectors for the n_i subjects in the sample from the i-th subpopulation all satisfy the restrictions

$$\underset{\sim}{C}_i\underset{\sim}{y}_{im} = \underset{\sim}{G}_i \quad \text{for } m = 1,2,\ldots,n_i \tag{2.66}$$

where $i = 1,2,\ldots,s$. Thus, if for each of the restrictions in (2.66), a pseudo data vector which contradicts it is included in the augmented data set, then the resulting pseudo estimated covariance matrix $\underset{\sim}{\underset{\sim}{V}}^{*}_{\underset{\sim}{y}}$ is nonsingular. Thus, the analysis of $\underset{\sim}{\bar{y}}$ can proceed as described in Section 2.1 except that $\underset{\sim}{\underset{\sim}{V}}^{*}_{\underset{\sim}{y}}$ is used as the estimated covariance matrix instead of $\underset{\sim}{\underset{\sim}{V}}_{\underset{\sim}{y}}$.

Finally, from a general point of view, it is useful to note that the construction of a pseudo estimated covariance matrix $\underset{\sim}{V}^{*}_{\underset{\sim}{F}}$ is not necessary

for certain types of analyses of a pertinent function vector $\underset{\sim}{F}$. In these cases potential singularities in $\underset{\sim}{V}_{\underset{\sim}{F}}$ are bypassed by partitioning the function vector $\underset{\sim}{F}$ into functions

$$\underset{\sim}{F}_{\underset{\sim}{R}} = \begin{bmatrix} \underset{\sim}{F}_{\underset{\sim}{R}1} \\ \underset{\sim}{F}_{\underset{\sim}{R}2} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{R}_1 \underset{\sim}{F} \\ \underset{\sim}{R}_2 \underset{\sim}{F} \end{bmatrix}, \tag{2.67}$$

where the rows of $\underset{\sim}{R}_1$ constitute a basis for the vector space spanned by the columns of $\underset{\sim}{V}_{\underset{\sim}{F}}$, and the rows of $\underset{\sim}{R}_2$ constitute a basis for the vector space orthogonal to the columns of $\underset{\sim}{V}_{\underset{\sim}{F}}$. The estimated covariance matrix for $\underset{\sim}{F}_{\underset{\sim}{R}}$ is

$$\underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R}}} = \begin{bmatrix} \underset{\sim}{R}_1 \\ \underset{\sim}{R}_2 \end{bmatrix} \underset{\sim}{V}_{\underset{\sim}{F}} \begin{bmatrix} \underset{\sim}{R}_1' & \underset{\sim}{R}_2' \end{bmatrix} = \begin{bmatrix} \underset{\sim}{R}_1 \underset{\sim}{V}_{\underset{\sim}{F}} \underset{\sim}{R}_1' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R}1}} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R}2}} \end{bmatrix} \tag{2.68}$$

where $\underset{\sim}{0}$ denotes matrices of 0's with appropriate dimensions. Since $\underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R}1}}$ is nonsingular by construction, linear regression models can be fitted to the functions $\underset{\sim}{F}_{\underset{\sim}{R}1}$ by the method of weighted least squares as described in Section 2.1. It is also possible to formulate models indirectly for the functions $\underset{\sim}{F}_{\underset{\sim}{R}2}$. This is accomplished by identifying matrices $\underset{\sim}{C}$ such that $\underset{\sim}{C}\ \underset{\sim}{F}_{\underset{\sim}{R}2} = \underset{\sim}{0}$. Since the true covariance matrix $\underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R}2},0}$ for $\underset{\sim}{F}_{\underset{\sim}{R}2}$ is necessarily nonsingular even though the estimated covariance matrix is null, it follows that the pseudo-chi-square statistic

$$Q_{\underset{\sim}{C},0}(\underset{\sim}{F}_{\underset{\sim}{R}2}) = \underset{\sim}{F}_{\underset{\sim}{R}2}' \underset{\sim}{C}' \left[\underset{\sim}{C}\ \underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R}2},0}\ \underset{\sim}{C}'\right]^{-1} \underset{\sim}{C}\ \underset{\sim}{F}_{\underset{\sim}{R}2} = 0, \tag{2.69}$$

which can be interpreted to mean that the functions $\underset{\sim}{F}_{\underset{\sim}{R}2}$ satisfy the hypothesis

$$\underset{\sim}{E}_{\underset{\sim}{A}} \{\underset{\sim}{C}\ \underset{\sim}{F}_{\underset{\sim}{R}2}\} = \underset{\sim}{0}. \tag{2.70}$$

However, the hypothesis (2.70) implies that $\underset{\sim}{F}_{\underset{\sim}{R}2}$ can be characterized by the

model

$$\underset{\sim}{E}_{\underset{\sim}{A}} \{\underset{\sim}{F}_{\underset{\sim}{R2}}\} = \underset{\sim}{X}_2 \underset{\sim}{\beta}_2, \tag{2.71}$$

where the columns of $\underset{\sim}{X}_2$ are a basis for the vector space orthogonal to $\underset{\sim}{C}$. Since $\underset{\sim}{C}\ \underset{\sim}{F}_{\underset{\sim}{R2}} = 0$, the estimator $\underset{\sim}{b}_2$ of $\underset{\sim}{\beta}_2$ can be determined directly by

$$\underset{\sim}{b}_2 = (\underset{\sim}{X}_2' \underset{\sim}{X}_2)^{-1} \underset{\sim}{X}_2' \underset{\sim}{F}_{\underset{\sim}{R2}}. \tag{2.72}$$

From (2.72), it follows that the estimated covariance matrix for $\underset{\sim}{b}_2$ is

$$\underset{\sim}{V}_{\underset{\sim}{b}_2} = (\underset{\sim}{X}_2' \underset{\sim}{X}_2)^{-1} \underset{\sim}{X}_2' \underset{\sim}{V}_{\underset{\sim}{G}_2} \underset{\sim}{X}_2 (\underset{\sim}{X}_2' \underset{\sim}{X}_2)^{-1} = \underset{\sim}{0}. \tag{2.73}$$

Finally, the separate analyses of the functions $\underset{\sim}{G}_1$ and $\underset{\sim}{G}_2$ can be tied together through comparisons of linear combinations of appropriate sets of model parameters. However, only hypothesis testing can be undertaken, and this is limited to the extent that the corresponding estimated covariance matrices (which are based entirely on $\underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R1}}}$ since $\underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R2}}} = \underset{\sim}{0}$) for such comparisons must be nonsingular. Although the logical conclusions of such hypothesis testing define models which link $\underset{\sim}{F}_{\underset{\sim}{R1}}$ and $\underset{\sim}{F}_{\underset{\sim}{R2}}$ together, the estimation of the respective parameters and their covariance matrix is not straightforward within the scope of the methodology presented in this paper. In this context, the major reason for this problem is $\underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R2}}} = \underset{\sim}{0}$ suggests there is no inherent variance in the functions $\underset{\sim}{F}_{\underset{\sim}{R2}}$ which confounds the ease with which relative weights can be assigned to the functions $\underset{\sim}{F}_{\underset{\sim}{R1}}$ and $\underset{\sim}{F}_{\underset{\sim}{R2}}$ with respect to parameter estimation, particularly since the functions $\underset{\sim}{F}_{\underset{\sim}{R2}}$ really do have inherent variance as expressed by the conceptual existence of nonsingular $\underset{\sim}{V}_{\underset{\sim}{F}_{\underset{\sim}{R2},0}}$. These remarks thus provide further motivation for the use of a pseudo estimated covariance matrix as previously discussed.

In conclusion, the best solution to the technical difficulties associated with potential singularities in $\underset{\sim}{V}_{\underset{\sim}{F}}$ is to avoid them in the first

place. For many situations involving moderately large samples, this can be accomplished by restricting attention to hypotheses involving the first order marginal probabilities $\{\phi_{igk}\}$ (or mean scores $\{\eta_{ig}\}$) and working with estimates obtained through operations like (2.46) and (2.56) (or (2.51) and (2.63)). In other cases, the use of modular maximum likelihood estimates as discussed in Koch and Tolley [23] may be required in order to "smooth" away the singularities in $\underset{\sim}{V}_{\underset{\sim}{F}}$ through the assumption of certain a priori constraints on the likelihood (2.1). However, if singularities are still present in $\underset{\sim}{V}_{\underset{\sim}{F}}$ for the function vector $\underset{\sim}{F}$ obtained after taking these more rigorous strategies into account, then the analysis of $\underset{\sim}{F}$ can be effectively undertaken through the use of a nonsingular pseudo estimated covariance matrix $\underset{\sim}{V}^{*}_{\underset{\sim}{F}}$ as described here represents an effective computational technique. Thus, questions pertaining to zero cells and singularities are, for the most part, best resolved in the context of statistical judgments which suitably take into account, on an individual basis, the special features of each particular experimental (or survey) situation where such methods are to be applied.

3. EXAMPLES

Four examples are presented to illustrate the methodology outlined in Section 2. First of all, a simple experiment for the comparison of three drugs is considered in Section 3.1 to indicate the full range of hypotheses which can be tested and the relationships between the corresponding fitted models. Sections 3.2 and 3.3 are concerned with two typical, but relatively complex, medical experiments in which interaction is present among the effects of factors associated with the sub-populations and measurement conditions. Finally, Section 3.4 deals with the analysis of data resulting from a well known social science research tool which involves within subject rankings of several attitudes or policy questions.

3.1. A Single Population Drug Comparison Example

The hypothetical data in Table 2 have been previously analyzed in both GSK and Koch and Reinfurt [22] to illustrate the construction of test statistics analogous to that in (2.48) for the hypothesis H_{CM} of first order marginal symmetry (homogeneity) in (2.23). They are being reanalyzed here to demonstrate the relationships between the test statistics for the

2. TABULATION OF RESPONSES TO DRUGS A, B, AND C

Sub-population	Response profile for Drug A vs Drug B vs Drug C							
	F F F	F F U	F U F	F U U	U F F	U F U	U U F	U U U
Overall group	6	16	2	4	2	4	6	6
Observed proportions	0.13	0.35	0.04	0.09	0.04	0.09	0.13	0.13
Estimated s.e.	0.05	0.07	0.03	0.04	0.03	0.04	0.05	0.05
Joint model pred. prop.	0.13	0.35	0.04	0.10	0.04	0.10	0.10	0.13
Estimated s.e.	0.05	0.07	0.02	0.02	0.02	0.02	0.02	0.05
Marg. model pred. prop.	0.13	0.35	0.04	0.09	0.04	0.09	0.13	0.13
Estimated s.e.	0.05	0.07	0.03	0.03	0.03	0.03	0.05	0.05
Assoc. model pred. prop.	0.16	0.31	0.05	0.09	0.05	0.09	0.09	0.16
Estimated s.e.	0.04	0.06	0.01	0.01	0.01	0.01	0.03	0.04

F denotes favorable response; U denotes unfavorable response.

hypotheses H_{CJ}, H_{CM}, and H_{JI} in Section 2 and the respective predicted values from corresponding fitted models.

The experimental design for this example involves s = 1 sub-population, d = 3 conditions which represent three drugs (Drug A, Drug B, and Drug C), and L = 2 response categories (favorable F and unfavorable U). Thus, there are $r = L^d = 2^3 = 8$ possible multivariate response profiles. If $\pi_{j_1 j_2 j_3}$

represents the probability of response profile (j_1,j_2,j_3) for (Drug A, Drug B, Drug C), then the hypothesis H_{CJ} of <u>total symmetry</u> in (2.20) here corresponds to

$$H_{CJ}: \begin{matrix} \pi_{FFU} = \pi_{FUF} = \pi_{UFF} \\ \pi_{FUU} = \pi_{UFU} = \pi_{UUF} \end{matrix} . \tag{3.1}$$

The functions required to test this hypothesis can be generated (with proper account given to (2.2)) via

$$\underset{\sim}{F} = \underset{\sim}{A}\,\underset{\sim}{p} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.13 \\ 0.35 \\ 0.04 \\ 0.09 \\ 0.04 \\ 0.09 \\ 0.13 \\ 0.13 \end{bmatrix} = \begin{bmatrix} 0.13 \\ 0.35 \\ 0.04 \\ 0.09 \\ 0.04 \\ 0.09 \\ 0.13 \end{bmatrix} . \tag{3.2}$$

If $\underset{\sim}{X} = \underset{\sim 7}{I}$ is used, then a test statistic for H_{CJ} in (3.1) is obtained via (2.12) with

$$\underset{\sim CJ}{C}_{4x7} = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} . \tag{3.3}$$

The resulting $\underset{\sim}{Q}_C = 16.29$ with D.F. = 4 which implies that there are significant ($\alpha = .01$) differences among drug effects. The nature of these differences can be attributed to the relatively large magnitude of p_{FFU}, the observed proportion with favorable responses to Drug A and Drug B but an unfavorable response to Drug C. This conclusion can be formally justified by fitting the model associated with

$$\underset{\sim J}{X}_{7x4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{3.4}$$

From (2.8), the goodness of fit statistic for (3.4) is Q = 0.51 with D.F. = 3 which is non-significant (α = .25). Thus, this model provides a satisfactory characterization of the distribution of the joint probabilities $\{\pi_{j_1 j_2 j_3}\}$. Predicted values for these quantities, which are obtained for the model (3.4) via (2.13), are given in Table 2 with corresponding estimated standard errors based on (2.14) in the "joint model" row.

The hypothesis H_{CM} of first order marginal symmetry (homogeneity) in (2.23) is formulated for this example as

$$H_{CM}: \quad \phi_{AF} = \phi_{BF} = \phi_{CF} \tag{3.5}$$

where

$$\begin{aligned} \phi_{AF} &= \pi_{FFF} + \pi_{FFU} + \pi_{FUF} + \pi_{FUU} \\ \phi_{BF} &= \pi_{FFF} + \pi_{FFU} + \pi_{UFF} + \pi_{UFU} \\ \phi_{CF} &= \pi_{FFF} + \pi_{FUF} + \pi_{UFF} + \pi_{UUF} \end{aligned} \tag{3.6}$$

The functions required to test H_{CM} can be generated (with proper account given to (2.2)) via

$$\underset{\sim}{F} = \underset{\sim}{A}\,\underset{\sim}{p} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.13 \\ 0.35 \\ 0.04 \\ 0.09 \\ 0.04 \\ 0.09 \\ 0.13 \\ 0.13 \end{bmatrix} = \begin{bmatrix} 0.61 \\ 0.61 \\ 0.34 \end{bmatrix} . \tag{3.7}$$

If $X = I_3$ is used, then a test statistic for H_{CM} in (3.5) is obtained via (2.12) with

$$\underset{2x3}{C_{CM}} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} . \tag{3.8}$$

The resulting $Q_{\underset{\sim}{C}} = 6.58$ with D.F. = 2 which implies that there are signif-

icant ($\alpha = .05$) gross (macro) differences among the drug effects. In particular, the proportion with favorable responses for Drug C is less than those associated with Drug A and Drug B which are essentially the same. A model which reflects this conclusion is

$$\underset{\sim}{X}_M = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.9}$$

for which the goodness of fit statistic from (2.8) is $Q = 0.00$ with D.F. = 1. However, this model only pertains to the first order marginal probabilities $\{\phi_{gk}\}$. For this reason, there is occasionally some interest in fitting a model directly to the joint probabilities $\{\pi_{j_1 j_2 j_3}\}$ which induces (3.9) on the $\{\phi_{gk}\}$. This type of analysis can be undertaken by using an augmented $\underset{\sim}{A}$-matrix

$$\underset{7\times 8}{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{3.10}$$

whose rows include the $\underset{\sim}{A}$-matrix in (3.7) on the one hand and constitute a basis for the same vector space as the rows of the $\underset{\sim}{A}$-matrix in (3.2). These functions are then analyzed in terms of an augmented $\underset{\sim}{X}$ matrix

$$\underset{7\times 6}{\underset{\sim}{X}_{MJ}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}_2 & \underset{\sim}{0}_{3,4} \\ \underset{\sim}{0}_{4,2} & \underset{\sim}{I}_4 \end{bmatrix} \tag{3.11}$$

which applies (3.10) to the (3.7) functions and a non-restrictive identity matrix $\underset{\sim}{I}_4$ to the others. The goodness of fit test statistic (2.8) for this

model is Q = 0.00 with D.F. = 1 which is identical to that associated with (3.9) by construction. Predicted values for the joint probabilities, which are obtained for the model (3.11) via (2.13), are given in Table 2 with corresponding estimated standard errors based on (2.14) in the "marginal model" row.

Although the results for the "joint model" approach and the "marginal model" approach are reasonably similar, certain differences do exist between them; in particular, H_{CJ} is contradicted at the $\alpha = .01$ level while H_{CM} is contradicted at the $\alpha = .05$ level. However, as discussed in Section 2.2, such paradoxes are usually a consequence of the pattern of association among the responses to the three drugs as reflected by the hypothesis H_{JI} in (2.26). From Bhapkar and Koch [6,7], it follows that the hypothesis H_{JI} can alternatively be formulated for this example in terms of the following constraint equations

$$\log_e\{\pi_{FFF}+\pi_{FFU}\}-\log_e\{\pi_{FUF}+\pi_{FUU}\}-\log_e\{\pi_{UFF}+\pi_{UFU}\}+\log_e\{\pi_{UUF}+\pi_{UUU}\} = 0 \quad (3.12)$$

$$\log_e\{\pi_{FFF}+\pi_{FUF}\}-\log_e\{\pi_{FFU}+\pi_{FUU}\}-\log_e\{\pi_{UFF}+\pi_{UUF}\}+\log_e\{\pi_{UFU}+\pi_{UUU}\} = 0 \quad (3.13)$$

$$\log_e\{\pi_{FFF}+\pi_{UFF}\}-\log_e\{\pi_{FFU}+\pi_{UFU}\}-\log_e\{\pi_{FUF}+\pi_{UUF}\}+\log_e\{\pi_{FUU}+\pi_{UUU}\} = 0 \quad (3.14)$$

$$\log_e\{\pi_{FFF}\}-\log_e\{\pi_{FFU}\}-\log_e\{\pi_{FUF}\}+\log_e\{\pi_{FUU}\}-\log_e\{\pi_{UFF}\}+\log_e\{\pi_{UFU}\} +\log_e\{\pi_{UUF}\}-\log_e\{\pi_{UUU}\} = 0 \quad (3.15)$$

where (3.12) - (3.14) specify no first order interaction (or pairwise independence) between the response to Drug A and Drug B, Drug A and Drug C, and Drug B and Drug C respectively, and (3.15) specifies no second order interaction among the responses to the three drugs. In order to test H_{JI} in the sense of (3.12) - (3.15) as well as the logit transform H_{CML} of H_{CM} where

$$H_{CML}: \quad \log_e\{\phi_{AF}/(1-\phi_{AF})\} = \log_e\{\phi_{BF}/(1-\phi_{BF})\} = \log_e\{\phi_{CF}/(1-\phi_{CF})\}, \quad (3.16)$$

log-linear functions as defined in (2.16) are required. These can be generated by using the $\underset{\sim}{A}$-matrix in (3.17)

$$\underset{\sim}{A}_{26x8} = \begin{bmatrix} \underset{\sim}{A}_1 \\ 6x8 \\ \underset{\sim}{A}_2 \\ 12x8 \\ \underset{\sim}{A}_3 \\ 8x8 \end{bmatrix} \text{ with } \underset{\sim}{A}_{1\ 6x8} = \begin{bmatrix} 1&1&1&1&0&0&0&0 \\ 0&0&0&0&1&1&1&1 \\ 1&1&0&0&1&1&0&0 \\ 0&0&1&1&0&0&1&1 \\ 1&0&1&0&1&0&1&0 \\ 0&1&0&1&0&1&0&1 \end{bmatrix}, \underset{\sim}{A}_{2\ 12x8} = \begin{bmatrix} 1&1&0&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&0&0&0&1&1&0&0 \\ 0&0&0&0&0&0&1&1 \\ 1&0&1&0&0&0&0&0 \\ 0&1&0&1&0&0&0&0 \\ 0&0&0&0&1&0&1&0 \\ 0&0&0&0&0&1&0&1 \\ 1&0&0&0&1&0&0&0 \\ 0&1&0&0&0&1&0&0 \\ 0&0&1&0&0&0&1&0 \\ 0&0&0&1&0&0&0&1 \end{bmatrix}, \text{ and } \underset{\sim}{A}_3 = \underset{\sim}{I}_8 \quad (3.17)$$

and the $\underset{\sim}{K}$-matrix in (3.18).

$$\underset{\sim}{K} = \begin{bmatrix} 1&-1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&1&-1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&1&-1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&-1&-1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&-1&-1&1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&-1&-1&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&-1&-1&1&-1&1&1&-1 \end{bmatrix} \quad (3.18)$$

In this format, the first three functions correspond to estimators for the first order marginal logits specified in (3.18) for Drug A, Drug B, and Drug C respectively; and the last four functions correspond to estimators for the measures of association specified by the left hand sides of the constraint equations (3.12) - (3.15) respectively. These quantities and their estimated covariance matrix are given in (3.19).

$$\underset{\sim}{f} = \begin{bmatrix} 0.44 \\ 0.44 \\ -0.63 \\ 1.99 \\ -0.69 \\ -0.69 \\ 0.41 \end{bmatrix}, \quad \underset{\sim}{V}_f = \begin{bmatrix} 0.0913 & 0.0413 & -0.0152 & 0.0000 & 0.0000 & 0.0091 & 0.0000 \\ 0.0413 & 0.0913 & -0.0152 & 0.0000 & 0.0091 & 0.0000 & 0.0000 \\ -0.0152 & -0.0152 & 0.0958 & 0.0102 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0102 & 0.4621 & -0.0481 & -0.0481 & 0.0000 \\ 0.0000 & 0.0091 & 0.0000 & -0.0481 & 0.4000 & 0.1850 & 0.0000 \\ 0.0091 & 0.0000 & 0.0000 & -0.0481 & 0.1850 & 0.4000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 2.0625 \end{bmatrix} \quad (3.19)$$

If $\underset{\sim}{X} = \underset{\sim}{I}_7$ is used, then a test statistic for H_{JI} in (3.12) - (3.15) is obtained via (2.12) with

$$\underset{\sim}{C}_{JI} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} . \tag{3.20}$$

The resulting $Q_{\underset{\sim}{C}} = 9.49$ with D.F. = 4 which implies that significant ($\alpha = .05$) association exists among the responses to the three drugs. On the other hand, if each of the rows of $\underset{\sim}{C}_{JI}$ in (3.20) are tested individually, the only one which is significant ($\alpha = .01$) in its own right corresponds to the pairwise association between the Drug A response and the Drug B response. Other test statistics of interest in this framework are those for H_{CML} in (3.16) and an analogous logarithmic equivalent of H_{CJ}. The latter is obtained via (2.12) with

$$\underset{\sim}{C}_{CJL} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix} \tag{3.21}$$

for which $Q_{\underset{\sim}{C}} = 14.96$ with D.F. = 4; and the former with

$$\underset{\sim}{C}_{CML} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.22}$$

for which $Q_{\underset{\sim}{C}} = 5.95$ with D.F. = 2. Both of these results are consistent with those previously obtained in terms of (3.2) - (3.3) and (3.7) - (3.8) respectively. Here, however, it is possible to identify the additional constraints associated with H_{CJ} but not H_{CM} as being equivalent to the equality of the pairwise measures of association defined by the left hand sides of (3.12) - (3.14). This hypothesis can be tested via (2.12) with

$$\underset{\sim}{C}_{CPAL} = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix} \tag{3.23}$$

for which Q = 8.48 with D.F. = 2. As with the test of H_{JI} in (3.20), the significance (α = .05) of this result is a consequence of the significant positive Drug A vs. Drug B pairwise association in contrast to the non-significant negative pairwise associations for Drug A vs. Drug C and Drug B vs. Drug C. On the basis of these conclusions and those noted earlier with respect to H_{CM}, it follows that the nature of significant differences among the effects of the three drugs can be explained in terms of the existence of a relatively large proportion of subjects that have favorable responses to Drug A and Drug B but an unfavorable response to Drug C. A final model for the functions $\underset{\sim}{f}$ which formally reflects this structure is

$$\underset{\sim}{X}_{\underset{\sim}{f}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} . \qquad (3.24)$$

The goodness of fit statistic from (2.8) is Q = 1.72 with D.F. = 4 which is non-significant (α = .25). Thus, this model provides a satisfactory characterization of the distribution of the joint probabilities $\{\pi_{j_1 j_2 j_3}\}$. With this model, the calculation of predicted values for the $\{\pi_{j_1 j_2 j_3}\}$ and their estimated standard errors is considerably more difficult because it involves the solution of non-linear equations, and hence the specific computational details are beyond the scope of this paper. However, such quantities can be obtained in a reasonably straightforward manner by the implicit function approach of Koch and Tolley [23] and are given in Table 2 in the "association model" row.

3.2. A Complex Split Plot Contingency Table

The following hypothetical example arose from a two-period change-over design clinical trial pertaining to the comparison of three treatments (Drug A, Drug B, and Drug P) with respect to the occurrence of a favorable response associated with the relief of certain persistent symptoms of a particular chronic disease. In this setting, the patients were sub-divided into two groups according to age with 50 patients being assigned to each of three sequence sub-groups in each age group. The resulting data are shown in Table 3 where O denotes the older age group and Y denotes the younger age group.

3. TABULATION OF RESPONSES FOR TWO-PERIOD-CHANGE-OVER DESIGN

Age	Sequence	Response profile at time 1 vs time 2 FF	FU	UF	UU	Total
O	A:B	12	12	6	20	50
O	B:P	8	5	6	31	50
O	P:A	5	3	22	20	50
Y	B:A	19	3	25	3	50
Y	A:P	25	6	6	13	50
Y	P:B	13	5	21	11	50
F - Favorable, U - Unfavorable						

The A-matrix

$$\underset{\sim}{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \otimes \underset{\sim}{I}_6 \qquad (3.25)$$

generates estimates for the first order marginal probabilities of patients with a favorable response and its corresponding estimated covariance matrix shown in Table 4.

4. ESTIMATES OF FIRST ORDER MARGINAL PROBABILITIES OF A FAVORABLE RESPONSE

Age	Sequence	Period 1 Prob. fav. estimate	Period 1 Variance estimate	Period 2 Prob. fav. estimate	Period 2 Variance estimate	Covariance estimate
O	A:B	0.48	0.0050	0.36	0.0046	0.0013
O	B:P	0.26	0.0038	0.28	0.0040	0.0017
O	P:A	0.16	0.0027	0.54	0.0050	0.0003
Y	B:A	0.44	0.0049	0.88	0.0021	-0.0001
Y	A:P	0.62	0.0047	0.62	0.0047	0.0023
Y	P:B	0.36	0.0046	0.68	0.0044	0.0003

Let $\phi_{i_1 i_2 g}$ denote the probability of a favorable response for the i_1-th age group, i_2-th treatment, and g-th period. If there are no carry-over effects of the drug in Period 1 to the response in Period 2 in the sense described in Grizzle [15] or Koch [19], then the following model is of interest

$$\phi_{i_1 i_2 g} = \mu_{i_1} + \xi_{i_1 i_2} + \tau_{i_1 g} \tag{3.26}$$

where μ_{i_1} is an overall mean for the i_1-th age group, $\xi_{i_1 i_2}$ is an effect due to i_2-th treatment in i_1-th age group, and $\tau_{i_1 g}$ is an effect due to g-th period for i_1-th age group. This model can be fitted via the regression model in (3.27)

$$\underset{\sim}{E}\{\underset{\sim}{F}\} = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 = \begin{bmatrix} 1 & 1 & 0 & -1 & & & & \\ 1 & 0 & 1 & 1 & & & & \\ 1 & 0 & 1 & -1 & & & & \\ 1 & -1 & -1 & 1 & & & & \\ 1 & -1 & -1 & -1 & & & & \\ 1 & 1 & 0 & 1 & & & & \\ & & & & 1 & 0 & 1 & -1 \\ & & & & 1 & 1 & 0 & 1 \\ & & & & 1 & 1 & 0 & -1 \\ & & & & 1 & -1 & -1 & 1 \\ & & & & 1 & -1 & -1 & -1 \\ & & & & 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \xi_{0A} \\ \xi_{0B} \\ \tau_{02} \\ \mu_Y \\ \xi_{YA} \\ \xi_{YB} \\ \tau_{Y2} \end{bmatrix} \tag{3.27}$$

for which the goodness of fit statistic is Q = 0.24 with D.F. = 4. Statistical tests for certain hypotheses are given in Table 5. These results sug-

5. STATISTICAL TESTS FOR $\underset{\sim}{X}_1$ MODEL

Hypothesis	D.F.	Q_C
$\xi_{YA} = \xi_{0A}$, $\xi_{YB} = \xi_{0B}$	2	0.16
$\tau_{02} = \tau_{Y2}$	1	6.08
$\mu_0 = \mu_Y$	1	40.34

gest that the model can be simplified to reflect no age x treatment interaction as indicated in (3.28).

$$\underset{\sim}{E\{F\}} = \underset{\sim}{X}_2\underset{\sim}{\beta}_2 = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & -1 & 0 \\ 1 & 0 & -1 & -1 & 1 & 0 \\ 1 & 0 & -1 & -1 & -1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 1 & -1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_0 \\ \mu_Y \\ \xi_A \\ \xi_B \\ \tau_{02} \\ \tau_{Y2} \end{bmatrix} . \qquad (3.28)$$

For this model, the goodness of fit statistic is Q = 0.41 with D.F. = 6; the resulting estimated parameters and predicted values are shown in (3.29).

$$\begin{bmatrix} 1 & 0 & 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & -1 & 0 \\ 1 & 0 & -1 & -1 & 1 & 0 \\ 1 & 0 & -1 & -1 & -1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 1 & -1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.34 \\ 0.60 \\ 0.16 \\ -0.04 \\ 0.05 \\ 0.13 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.35 \\ 0.26 \\ 0.27 \\ 0.18 \\ 0.55 \\ 0.43 \\ 0.88 \\ 0.63 \\ 0.61 \\ 0.35 \\ 0.69 \end{bmatrix} \qquad (3.29)$$

Finally, statistical tests for hypotheses pertaining to this model are given in Table 6.

6. STATISTICAL TESTS FOR $\underset{\sim}{X}_2$ MODEL

Hypothesis	D.F.	Q_C
$\xi_A = \xi_B = 0$	2	49.21
$\mu_0 = \mu_Y$	1	40.37
$\tau_{02} = 0$	1	4.81
$\tau_{Y2} = 0$	1	31.16
$\tau_{02} = \tau_{Y2}$	1	6.37
$2\xi_A + \xi_B = 0$ $(\xi_A = \xi_P)$	1	47.55
$\xi_A + 2\xi_B = 0$ $(\xi_B = \xi_P)$	1	3.68
$\xi_A = \xi_B$	1	21.17

Thus, this analysis suggests that Drug A is significantly different from Drug B and Drug P which are relatively similar, period effects are significant for each group and interact with age in the sense of being of greater magnitude for the younger age group, and the overall mean for the younger age group is significantly greater than that for the older age group.

3.3. A Longitudinal Growth Curve Model Contingency Table

The following hypothetical example arose from a longitudinal follow-up study to compare a new drug and a standard drug with respect to the treatment of patients with both mild and severe diagnoses of a particular disease. In this regard, a patient's condition was graded as normal (N) or abnormal (A) at the end of 1 week, 2 weeks, and 4 weeks of continuous treatment with the results shown in Table 7.

7. TABULATION OF RESPONSES FOR LONGITUDINAL STUDY

Diagnosis	Treatment	Response profile at week 1 vs week 2 vs week 4 NNN	NNA	NAN	NAA	ANN	ANA	AAN	AAA	Total
Mild	Standard	16	13	9	3	14	4	15	6	80
Mild	New drug	31	0	6	0	22	2	9	0	70
Severe	Standard	2	2	8	9	9	15	27	28	100
Severe	New drug	7	2	5	2	31	5	32	6	90

The $\underset{\sim}{A}$-matrix

$$\underset{\sim}{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \otimes \underset{\sim}{I}_4 \qquad (3.30)$$

generates estimates for the first order marginal probabilities of normal (N)

and abnormal (A) for each week vs diagnosis vs treatment combination; and the $\underset{\sim}{K}$ matrix $\underset{\sim}{K} = [1\ -1] \otimes \underset{\sim}{I}_{12}$ generates their respective log ratios (or logits) as shown in Table 8 together with the corresponding estimated covar-

8. OBSERVED AND PREDICTED ESTIMATES FOR FIRST ORDER MARGINAL PROBABILITIES FOR NORMAL RESPONSE AND CORRESPONDING LOGITS

Diagnosis	Treatment	Week	Observed est. prob. normal	Est. s.e.	Observed est. logit	Est. s.e.	Predicted est. logit	Est. s.e.	Predicted est. prob. normal	Est. s.e.
Mild	Standard	1	0.51	0.06	0.05	0.22	-0.07	0.13	0.48	0.03
Mild	Standard	2	0.59	0.06	0.35	0.23	0.42	0.11	0.60	0.03
Mild	Standard	4	0.68	0.05	0.73	0.24	0.92	0.16	0.71	0.03
Mild	New drug	1	0.53	0.06	0.11	0.24	-0.07	0.13	0.48	0.03
Mild	New drug	2	0.79	0.05	1.30	0.29	1.38	0.15	0.80	0.02
Mild	New drug	4	0.97	0.02	3.53	0.71	2.84	0.25	0.94	0.01
Severe	Standard	1	0.21	0.04	-1.32	0.25	-1.35	0.13	0.21	0.02
Severe	Standard	2	0.28	0.04	-0.94	0.22	-0.86	0.10	0.30	0.02
Severe	Standard	4	0.46	0.05	-0.16	0.20	-0.36	0.15	0.41	0.04
Severe	New drug	1	0.18	0.04	-1.53	0.27	-1.35	0.13	0.21	0.02
Severe	New drug	2	0.50	0.05	0.00	0.21	0.10	0.12	0.53	0.03
Severe	New drug	4	0.83	0.04	1.61	0.27	1.56	0.21	0.82	0.03

iance matrix as shown in (3.31).

$$\underset{\sim}{V}_f = \begin{bmatrix} 5.00 & 1.27 & -0.76 & & & & & & & & & \\ 1.27 & 5.16 & -0.51 & & & & & & & & & \\ -0.76 & -0.51 & 5.70 & & & & & & & & & \\ & & & 5.73 & 0.94 & 3.12 & & & & & & \\ & & & 0.94 & 8.48 & -1.87 & & & & & & \\ & & & 3.12 & -1.87 & 51.47 & & & & & & \\ & & & & & & 6.03 & -0.56 & 0.08 & & & \\ & & & & & & -0.56 & 4.96 & -0.38 & & & \\ & & & & & & 0.08 & -0.38 & 4.03 & & & \\ & & & & & & & & & 7.60 & 0.34 & -0.81 \\ & & & & & & & & & 0.34 & 4.44 & 0.18 \\ & & & & & & & & & -0.81 & 0.18 & 8.00 \end{bmatrix} \times 10^{-2} \qquad (3.31)$$

Let $\lambda_{i_1 i_2 g}$ denote the large sample expected value of the logit corresponding to the i_1-th diagnosis, i_2-th treatment, and g-th week. If time is assumed to represent a metric which reflects dosage of the drugs under study,

then the linear logistic model with respect to log time represents a reasonable model by analogy to well known results from existing methodology for quantal bio-assays as discussed by Berkson [2,3,4] or Finney [11]. More specifically, we first consider the model (3.32)

$$\lambda_{i_1 i_2 g} = \mu_{i_1 i_2} + \gamma_{i_1 i_2} x_{i_1 i_2 g} \qquad \begin{array}{l} i_1 = 1\text{: Mild, } 2\text{: Severe} \\ i_2 = 1\text{: Standard, } 2\text{: New Drug} \\ g = 1\text{: Week 1, } 2\text{: Week 2, } 3\text{: Week 4} \end{array} \tag{3.32}$$

where $\mu_{i_1 i_2}$ represents an intercept parameter in reference to week 1 which is associated with i_1-th diagnosis and i_2-th treatment, $\gamma_{i_1 i_2}$ represents a corresponding continuous slope effect over time, and $x_{i_1 i_2 g}$ is the log to the base 2 of week (i.e., $x_{i_1 i_2 g} = 0,1,2$). In matrix notation, this model can be fitted via the regression model (3.33)

$$\underset{\sim}{E}_{\underset{\sim}{A}}\{\underset{\sim}{f}\} = \underset{\sim}{X}_1 \underset{\sim}{\beta}_1 = \begin{bmatrix} 1 & 0 & & & & & & \\ 1 & 1 & & & & & & \\ 1 & 2 & & & & & & \\ & & 1 & 0 & & & & \\ & & 1 & 1 & & & & \\ & & 1 & 2 & & & & \\ & & & & 1 & 0 & & \\ & & & & 1 & 1 & & \\ & & & & 1 & 2 & & \\ & & & & & & 1 & 0 \\ & & & & & & 1 & 1 \\ & & & & & & 1 & 2 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \gamma_{11} \\ \mu_{12} \\ \gamma_{12} \\ \mu_{21} \\ \gamma_{21} \\ \mu_{22} \\ \gamma_{22} \end{bmatrix} \tag{3.33}$$

for which the goodness of fit statistic is Q = 1.60 with D.F. = 4. The hypotheses and test statistics in Table 9 suggest differences exist among

9. STATISTICAL TESTS FOR $\underset{\sim}{X}_1$ MODEL

Hypothesis	D.F.	Q_C
$\mu_{11} = \mu_{12} = \mu_{21} = \mu_{22}$	3	44.47
$\gamma_{11} = \gamma_{12} = \gamma_{21} = \gamma_{22}$	3	29.46
$\mu_{11} = \mu_{12}$	1	0.01
$\mu_{21} = \mu_{22}$	1	0.17
$\gamma_{11} = \gamma_{21}$	1	1.29
$\gamma_{12} = \gamma_{22}$	1	0.16

the respective diagnosis x treatment patient groups with respect to both the intercept and slope parameters; and that such differences among the intercepts can be explained in terms of a diagnosis effect while such differences among the slopes can be explained in terms of a treatment effect. On the basis of these results, the original model can be simplified to that shown in (3.34)

$$\underset{\sim}{E}_{\underset{\sim}{A}}\{\underset{\sim}{f}\} = \underset{\sim}{X}_2\underset{\sim}{\beta}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} \mu_{1*} \\ \mu_{2*} \\ \gamma_{*1} \\ \gamma_{*2} \end{bmatrix} \qquad (3.34)$$

where μ_{i_1*} is an intercept parameter associated with the i_1-th diagnosis and γ_{*i_2} is a slope effect associated with the i_2-th treatment. For this model, the goodness of fit statistic is Q = 4.20 with D.F. = 8. The corresponding estimated parameter vector $\underset{\sim}{b}_2$ and its estimated covariance matrix $\underset{\sim}{V}_{\underset{\sim}{b}_2}$ are given in (3.35).

$$\underset{\sim}{b}_2 = \begin{bmatrix} -0.07 \\ -1.35 \\ 0.49 \\ 1.46 \end{bmatrix}, \quad \underset{\sim}{V}_{\underset{\sim}{b}_2} = \begin{bmatrix} 1.82 & 0.75 & -0.74 & -0.64 \\ 0.75 & 1.82 & -0.83 & -1.04 \\ -0.74 & -0.83 & 0.93 & 0.53 \\ -0.64 & -1.04 & 0.53 & 1.69 \end{bmatrix} \times 10^{-2} \qquad (3.35)$$

From these results, predicted logits as shown in Table 8 can be determined via (2.13). These can then be used to obtain the predicted values for the

first order marginal probabilities of normal (N) responses by reverse transformation. These quantities are also shown in Table 8. Estimated standard errors for these predicted values obtained through suitable manipulations of (2.14) are substantially smaller than those for the corresponding observed estimates, and thus reflect the extent to which the fitted model X_2 enhances statistical efficiency. Finally, the hypotheses and test statistics in Table 10 justify the conclusions that the effects of diagnosis are signficant

10. STATISTICAL TESTS FOR X_2 MODEL

Hypothesis	D.F.	Q_C
$\mu_{1*} = \mu_{2*}$	1	77.02
$\gamma_{*1} = \gamma_{*2}$	1	59.12
$\gamma_{*1} = 0$	1	26.35
$\gamma_{*2} = 0$	1	125.08

but do not interact with time. Drug effects are also significant, but are modulated in terms of different linear logistic trends over time. In other words, for both mild and severe diagnoses, a patient's condition becomes graded (N) relatively sooner with the new drug than the standard even though there is essentially no difference between the treatments at the end of one week.

3.4. Ranked Policy Preference Data

The following example is based on a survey which was administered to a national sample of adult Americans in order to elicit ranked preferences regarding the following seven tax alternatives: education (ED), environment (EN), cut taxes (CT), anti-poverty (P), foreign aid (F), income supplements (I), and health (H). Respondents were also classified into sub-populations

according to sex (male or female), ideology (conservative, liberal, in-between, no ideology) and tendency to criticize governmental tax policies (no, yes). Other aspects of this survey are described in Lehnen and Koch [26]. In order to analyze the relationship between sex x ideology x criticism and the ranked response profile of the seven tax alternatives, attention must be directed at a contingency table with 16 rows and 7! = 5040 columns. However, it is not necessary to generate this conceptual contingency table if the within group mean ranks for the respective alternatives can be viewed as a reasonable measure of preference since the corresponding $\underset{\sim}{A}$-matrix operation can be achieved by direct computation of these mean scores and their covariance matrix on the observed respondent-wise raw data by operations like those described in (2.55) - (2.61). For the sub-population of n = 169 respondents with ideology = in-between, sex = male, and criticism = no, this process gave rise to the results shown in Table 11.

11. MEAN RANK PREFERENCE AND COVARIANCE MATRIX FOR THE SUB-POPULATION WITH IDEOLOGY = IN-BETWEEN, SEX = MALE, CRITICISM = NO

Tax alternative	Mean rank preference	Estimated covariance matrix x 10^4						
		ED	EN	CT	P	F	I	H
ED	2.1	118	6	- 41	3	- 11	- 64	- 11
EN	3.7	6	172	- 24	- 32	23	-100	- 45
CT	3.7	- 41	- 24	260	- 87	- 34	- 29	- 45
P	4.5	3	- 32	- 87	156	- 9	- 10	- 21
F	6.1	- 11	23	- 34	- 9	89	- 45	- 13
I	4.4	- 64	-100	- 29	- 10	- 45	222	26
H	3.4	- 11	- 45	- 45	- 21	- 13	26	109

More completely, Table 12 presents mean ranks for each of the sex x

12. MEAN RANKS FOR EACH SEX X CRITICISM SUB-POPULATION WITHIN THE IDEOLOGY = IN-BETWEEN SUB-POPULATION

Sex	Criticism	n	Mean rank preferences							Friedman's χ^2 (D.F.=6)	Equal marg. mean Q_C (D.F.=6)
			ED	EN	CT	P	F	I	H		
M	No	169	2.1	3.7	3.7	4.5	6.1	4.4	3.4	325.6	815.7
M	Yes	75	2.2	3.1	3.6	5.0	6.3	4.8	3.1	191.7	1065.1
F	No	191	2.3	3.8	3.3	4.7	6.2	4.3	3.4	381.6	1365.5
F	Yes	52	2.5	3.3	2.5	5.1	6.7	4.4	3.4	153.7	1372.8

criticism sub-populations within the ideology = in-between sub-population together with corresponding Friedman [14] χ^2-statistics and equality of marginal mean score $\underset{\sim}{Q_C}$-statistics for the hypotheses of within group indifference (i.e., equality of mean ranks) in a sense analogous to H_{CAM} in (2.34). Here, it can be noted that the marginal mean $\underset{\sim}{Q_C}$-statistic is considerably larger than the Friedman χ^2-statistic. This results from the fact that the underlying variance of the mean ranks is greater when the hypothesis is true than when it is not true in the direction of a systematic pattern of preference, together with the fact that the Friedman statistic is based on estimates of variance under the hypothesis while the marginal mean $\underset{\sim}{Q_C}$-statistic is based on unrestricted estimates of variance derived from the underlying contingency table.

The effects of sex and criticism on the mean ranks associated with each alternative can be investigated in a univariate context by testing the coefficients of the regression model in (3.36).

$$\underset{\sim}{X}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{3.36}$$

The resulting $\underset{\sim}{Q_C}$ statistics with D.F. = 1 are indicated in Table 13. However, it

13. TEST STATISTICS Q_C WITH D.F. = 1 FOR REGRESSION PARAMETERS

	ED	EN	CT	P	F	I	H
Sex	2.2	1.2	14.4**	1.8	7.1*	2.0	1.7
Criticism	1.0	12.8**	4.4*	8.2**	13.1**	1.7	1.9
S x C	0.1	0.3	2.2	0.0	2.6	0.6	2.2

* means significant at $\alpha = .05$; ** means significant at $\alpha = .01$

should be recognized that the mean ranks satisfy a constraint in the sense that their sum is always (1 + 2 + 3 + 4 + 5 + 6 + 7) = 28. For this reason,

a simultaneous multivariate analysis of the effects of sex and criticism on the entire profile of response is more appropriate than univariate analyses of each tax alternative since it reflects the extent to which increases associated with one choice are offset by decreases associated with another. This type of analysis can be undertaken by deleting one of the tax alternatives from consideration (e.g., cut taxes CT) and handling the remaining six choices for each of the four groups simultaneously. Since the univariate analyses suggested no sex x criticism interaction on all of the choices, an appropriate preliminary model is given in (3.37)

$$\underset{\sim}{I}_6 \otimes \underset{\sim}{X}_1 = \begin{bmatrix} \underset{\sim}{I}_6 & \underset{\sim}{I}_6 & \underset{\sim}{I}_6 \\ \underset{\sim}{I}_6 & \underset{\sim}{I}_6 & -\underset{\sim}{I}_6 \\ \underset{\sim}{I}_6 & -\underset{\sim}{I}_6 & \underset{\sim}{I}_6 \\ \underset{\sim}{I}_6 & -\underset{\sim}{I}_6 & -\underset{\sim}{I}_6 \end{bmatrix} \cdot \begin{bmatrix} \underset{\sim}{\mu} \\ \underset{\sim}{\gamma} \\ \underset{\sim}{\xi} \end{bmatrix} \qquad (3.37)$$

where $\underset{\sim}{\mu}$ corresponds to overall mean parameters, $\underset{\sim}{\gamma}$ corresponds to sex effects, and $\underset{\sim}{\xi}$ corresponds to criticism effects. The goodness of fit statistic for this model was $Q = 8.3$ with D.F. = 6 for multivariate sex x criticism interaction on preference which supports the validity of the model in a multivariate context. For the hypothesis $\underset{\sim}{\gamma} = \underset{\sim}{0}$, $Q_C = 24.8$ with D.F. = 6 and for the hypothesis $\underset{\sim}{\xi} = \underset{\sim}{0}$, $Q_C = 42.6$ with D.F. = 6, both of which are significant ($\alpha = .01$). Moreover, the univariate tests suggest that sex effects can be primarily attributed to CT and F while criticism effects are associated with EN, CT, P, and F. Further investigation and analysis of these effects gave rise to the final model with seven parameters shown in (3.38) for which the goodness of fit statistic $Q = 22.2$ with D.F. = 17.

$$\underset{\sim}{X}_2 = \begin{bmatrix} 1 & & & & 0 \\ & 1 & & & -2 \\ & & 1 & & 2 \\ & & & 1 & 3 \\ & & & 1 & 0 \\ & & & 1 & 1 \\ 1 & & & & 0 \\ & 1 & & & 2 \\ & & 1 & & -2 \\ & & & 1 & -1 \\ & & & 1 & 0 \\ & & & 1 & 1 \\ 1 & & & & 0 \\ & 1 & & & -2 \\ & & 1 & & 2 \\ & & & 1 & 1 \\ & & & 1 & 0 \\ & & & 1 & -1 \\ 1 & & & & 0 \\ & 1 & & & 2 \\ & & 1 & & -2 \\ & & & 1 & -3 \\ & & & 1 & 0 \\ & & & 1 & -1 \end{bmatrix} \qquad (3.38)$$

The predicted values from this model for mean rank preferences are given in Table 14.

14. PREDICTED MEAN RANKS FOR EACH SEX X CRITICISM SUB-POPULATION WITHIN THE IDEOLOGY = IN-BETWEEN SUB-POPULATION

Sex	Criticism	ED	EN	CT	P	F	I	H
M	No	2.2	3.7	3.6	4.7	6.0	4.5	3.3
M	Yes	2.2	3.3	3.2	5.1	6.4	4.5	3.3
F	No	2.2	3.7	3.2	4.7	6.2	4.5	3.5
F	Yes	2.2	3.3	2.8	5.1	6.6	4.5	3.5

ACKNOWLEDGMENT

The authors would like to thank Rebecca Wesson for her conscientious typing of the manuscript.

REFERENCES

[1] Anderson, R. L. and Bancroft, T. A., Statistical Theory in Research, New York: McGraw-Hill Co. (Chapter 23), 1952.

[2] Berkson, J., "Application of the Logistic Function to Bioassay," Journal of the American Statistical Association, 39 (1944), 357-65.

[3] _____. "A Statistically Precise and Relatively Simple Method of Estimating the Bioassay with Quantal Response Based on the Logistic Function," Journal of the American Statistical Association, 48 (1953), 565-99.

[4] _____. "Maximum Likelihood and Minimum χ^2 Estimates of the Logistic Function," Journal of the American Statistical Association, 50 (March 1955), 130-62.

[5] Bhapkar, V. P., "On the Analysis of Contingency Tables with a Quantitative Response, " Biometrics, 24 (June 1968), 329-38.

[6] _____ and Koch, G. G., "Hypotheses of 'No Interaction' in Multidimensional Contingency Tables," Technometrics, 10 (February 1968), 107-23.

[7] _____ and Koch, G. G., "On the Hypotheses of 'No Interaction' in Contingency Tables," Biometrics, 24 (September 1968), 567-94.

[8] Cole, J. W. L. and Grizzle, J. E., "Application of Multivariate Analysis of Variance to Repeated Measurements Experiments," Biometrics, 22 (December 1966), 810-27.

[9] Darroch, J. N., "Multiplicative and Additive Interaction in Contingency Tables," Biometrika, 61 (August 1974), 207-14.

[10] Federer, W. T., Experimental Design: Theory and Application, New York: Macmillan Co. (Chapter 10), 1955.

[11] Finney, D. J., Statistical Method in Biological Assay, Second Edition, London: Charles Griffin and Company, Ltd., 1964.

[12] Forthofer, R. N. and Koch, G. G., "An Analysis for Compounded Functions of Categorical Data," Biometrics, 29 (March 1973), 143-57.

[13] Freeman, D. H., Freeman, J. L. and Koch, G. G., "A Modified χ^2 Approach for Fitting Weibull Models to Synthetic Life Tables," submitted to Biometrics.

[14] Friedman, M., "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," Journal of the American Statistical Association, 32 (1937), 675-99.

[15] Grizzle, J. E., "The Two-Period Change-over Design and Its Use in Clinical Trials," Biometrics, 21 (June 1965), 467-80.

[16] _____, Starmer, C. F., and Koch, G. G., "Analysis of Categorical Data by Linear Models," Biometrics, 25 (September 1969), 489-504.

[17] Koch, Gary G., "Some Aspects of the Statistical Analysis of 'Split Plot' Experiments in Completely Randomized Layouts," Journal of the American Statistical Association, 64 (June 1969), 485-505.

[18] _____, "The Use of Non-Parametric Methods in the Statistical Analysis of a Complex Split Plot Experiment," Biometrics,26 (March 1970), 105-28.

[19] _____, "The Use of Non-Parametric Methods in the Statistical Analysis of the Two-Period Change-over Design," Biometrics, 28 (June 1972), 577-84.

[20] _____, Imrey, P. B. and Reinfurt, D. W.,"Linear Model Analysis of Categorical Data With Incomplete Response Vectors," Biometrics, 28 (September 1972), 663-92.

[21] _____, Johnson, W. D., and Tolley, H. D., "A Linear Models Approach to the Analysis of Survival and Extent of Disease in Multidimensional Contingency Tables," Journal of the American Statistical Association, 67 (December 1972), 783-96.

[22] _____ and Reinfurt, D. W., "The Analysis of Categorical Data from Mixed Models," Biometrics, 27 (March 1971), 157-73.

[23] _____ and Tolley, H. D., "A Generalized Modified χ^2 Analysis of Data from a Complex Serial Dilution Experiment," Biometrics, 31 (March 1975).

[24] _____, "Analyzing Panel Data with Uncontrolled Attrition," The Public Opinion Quarterly, 38 (Spring 1974), 40-56.

[25] _____, "The Analysis of Categorical Data from Repeated Measurement Research Designs," to appear in Political Methodology, (Fall 1974).

[26] Lehnen, R. G., and Koch, G. G., "A Generalized Approach for the Analysis of Ranked Policy Preference Data," to be submitted to Public Opinion Quarterly, (Winter 1974).

[27] Morrison, Donald F., Multivariate Statistical Methods, New York: McGraw-Hill Co. (Chapters 4 and 5), 1967.

[28] Neyman, J., "Contribution to the Theory of the χ^2 Test," Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles: University of California Press (239-72), 1949.

[29] Steel, Robert G. D. and Torrie, James H., Principles and Procedures of Statistics, New York: McGraw-Hill Co. (Chapter 12), 1960.

[30] Wald, A., "Tests of Statistical Hypotheses Concerning General Parameters When the Number of Observations is Large," Transactions of the American Mathematical Society, 54 (November 1943), 426-82.