

1

This research was supported by the National Institutes of Health, Institute of General Medical Sciences (Grants GM-70004-04 and GM-00038-18) and by the National Institute of Child Health and Human Development (Grant HD-03441-04).

AN INCOMPLETE CONTINGENCY TABLE APPROACH
TO PAIRED-COMPARISON EXPERIMENTS

by

Peter B. Imrey and Gary G. Koch
Department of Biostatistics
University of North Carolina at Chapel Hill

and

William D. Johnson
Louisiana State University Medical Center

Institute of Statistics Mimeo Series No.1004

May 1975

11

AN INCOMPLETE CONTINGENCY TABLE APPROACH
TO PAIRED-COMPARISON EXPERIMENTS

by

Peter B. Imrey
William D. Johnson
Gary G. Koch

ABSTRACT

A wide variety of paired-comparison experiments and surveys may be viewed within the framework of "incomplete" contingency table analysis. Bradley-Terry models for such data can be chosen and further examined by a non-iterative logit analysis. Multivariate paired-comparisons with factor structure can be analyzed through reparameterization of the underlying models for each combination of factor levels. The method is subject to sample size restrictions, but allows the experimenter to design his survey without complex symmetry conditions or independence assumptions.

AN INCOMPLETE CONTINGENCY TABLE APPROACH
TO PAIRED-COMPARISON EXPERIMENTS

by

Peter B. Imrey, William D. Johnson, and Gary G. Koch

1. INTRODUCTION

Paired-comparison techniques have been widely utilized in consumer preference surveys, psychological testing, the evaluation of tournament results, quality control, and a variety of other situations which involve variables difficult to quantify or scale in an objective manner. A paired-comparison is simply a choice between two items; it is frequently easy for individuals to make such choices, even when they cannot characterize grounds for their preferences. Aggregated preferences of many individuals may permit the quantification of a concept prior to discovery of a fully adequate operational definition.

In all paired-comparison data, the unit of analysis is the choice. The data set consists of many choices made by individual "judges" within particular pairs of items, or "objects." The structure of the data and subsequent analysis may be complicated by the method of assignment of pairs to judges, the presence of separate judgements corresponding to multiple criteria of comparison, and by certain types of heterogeneity in the collection of judges.

As illustration, consider the following examples:

1) Each of a sample of women in the child-bearing ages is asked to choose between two "ideal family sizes" selected from the universe $U = \{0,1,2,3,4,5,6+\}$. Each woman chooses within only one pair; approx-

imately equal numbers are confronted with each possible pair (Section 3);

ii) Each member of a group of judges is asked to choose between two chocolate puddings, separately on the basis of a) taste, b) flavor and c) overall quality. Three puddings are under study; each judge compares only two (Section 4);

iii) Each member of a group of judges is asked to choose between two soft drinks, separately on the basis of a) taste and b) fizziness. Four soft drinks are under study; each judge deals with two pairs sharing one member (Section 5);

iv) Two opinion questions are appended to a driver license examination, the first relating to alternative law enforcement policies, the second to alternative auto safety features. Each question is a paired-choice, with two of three alternatives appearing on each examination paper. Each pair of choices on the law enforcement question appears with only one pair of auto safety alternatives. One wishes to analyze the responses with particular attention to differentials by age and sex (Section 6).

We proceed here to demonstrate a general approach as applied to these successively more complex situations. Our analyses are within the framework of the Bradley-Terry [6] approach. The literature on paired-comparisons is extensive, the work of Luce [21], Davidson and Bradley [8,9,10] David [7], Kupper and Rao [20], and Atkinson [2] being among the more significant contributions. The line we follow can avoid problems relating to balance in design, independence of choices, or complex multivariate parametrizations, that have required much attention from other workers. However, limitations with respect to sample size can hinder applications of the method here, which supplements rather than replaces existing techniques.

2. INCOMPLETE CONTINGENCY TABLES AND BRADLEY-TERRY MODELS

A general paired comparison experiment may be described as follows:

Structure of the Experiment

Subject (or judge) set $S = \{s_i\}$, $i = 1, \dots, n$

Object set $O = \{o_j\}$, $j = 1, \dots, v$

$\Sigma = \{S_k\}$, $k = 1, \dots, s$, a partitioning of S

Criterion set $C = \{c_\alpha\}$, $\alpha = 1, \dots, m$

$T =$ comparison set $= \{L_k\}$, $k = 1, \dots, s$, where each L_k is a list, possibly with repetitions, of elements $(o_x, o_y | c_\alpha)$ of $O \times O \times T$, $x \neq y$.

Data from the experiment consists, for all k , of the preferences of each $s_i \in S_k$ for each comparison in the list L_k . We assume from now on that order of presentation within pairs has no effect, so that the triple $(o_x, o_y | c_\alpha)$ may be identified with the triple $(o_y, o_x | c_\alpha)$.

A contingency table is "incomplete" if the probability structure underlying the table is known to specify that certain cells are empty with probability one. Such tables have been studied by Goodman [13], Bishop and Fienberg [5], Fienberg [11], Mantel [22], and Williams and Grizzle [26], among others. Data from the set of subjects S_k may be expressed as a v^{ℓ_k} incomplete contingency table, where ℓ_k is the length of the list L_k , and each one-dimensional margin of the table refers to the choice corresponding to a particular element of L_k . Sometimes the s incomplete tables corresponding to S_1, \dots, S_s may be combined into one incomplete table of simpler form. In any case, the analysis of paired-comparison data as described above may be viewed as a problem in analyzing one or a set of related incomplete contingency tables. A number of approaches to such problems are available; asymp-

totic methods are appropriate if the various S_k are sufficiently large.

Bradley and Terry [6] and Davidson and Bradley [8,9,10] have discussed univariate and multivariate models for paired-comparisons which have been particularly useful. Suppose that judges are selected randomly from some population and that the probability that a random judge chooses o_x in the choice determined by $(o_x, o_y | c_\alpha)$ is denoted by $\pi_{\alpha;xy}$. The Bradley-Terry model and its multivariate generalizations assume that

$$\pi_{\alpha;xy} = \frac{\pi_{\alpha;x}}{\pi_{\alpha;x} + \pi_{\alpha;y}}$$

for sets of $\pi_{\alpha;z}$, $z = 1, \dots, v$, such that $\pi_{\alpha;z} > 0$, $\sum_{z=1}^v \pi_{\alpha;z} = 1$. The parameters $\pi_{\alpha;z}$ provide an indication of the strengths of the different objects with respect to criterion c_α , as viewed by the subjects or judges. Luce [1959] has discussed situations in which the $\pi_{\alpha;z}$ have additional significance, in that $\pi_{\alpha;z}$ may be interpreted as the probability that a random judge, choosing according to c_α , would rank o_z first in a ranking of the entire object set O .

Bradley-Terry models can equivalently be expressed as models of quasi-independence (complete or conditional) within appropriately defined incomplete contingency tables, under the usual multinomial assumptions. We can therefore apply contingency table techniques to estimation and testing associated with Bradley-Terry models. We apply the large-sample weighted least-squares approach of Grizzle, Starmer, and Koch [14], used previously on incomplete data by Imrey and Koch [15], Koch, Imrey, and Reinfurt [18], and Koch, Abernathy, and Imrey [16]. Atkinson [2] uses a similar approach to such data. The examples given are presented to illustrate the techniques; no claims are made as to adequacy of sample sizes. In Section 7 we remark on these and other considerations.

3. SIMPLE PAIRED-CHOICE

"Let's suppose, for a moment, that you have just been married and that you are given a choice of having, during your entire lifetime, either x or y children. Which would you choose, x or y ?"

The data in Table 1 represent response of 447 white North Carolina women under thirty, married to their first husband, to the above question when pairs of the numbers zero through six were substituted for x and y . Each woman was queried only with respect to one pair, and pairs were randomly assigned to women. The data were gathered during the 1968 North Carolina Abortion Survey. We assume for simplicity that these women constitute a simple random sample of all such women throughout the state.

The data are arranged as an incomplete contingency table, with rows corresponding to pairs and columns to choices. The criterion set here has only one element; hence in our notation for this example we will neglect the subscript α . To fit a Bradley-Terry model to this data set one need only note the equivalence of that model to the model of quasi-independence of rows and columns of Table 1. Thus, maximum-likelihood estimates of the parameters $\pi_0, \pi_1, \dots, \pi_6$ of the Bradley-Terry model may be found by applying the iterative proportional fitting technique in e.g., Goodman [13], or Fienberg [11]. An alternative route is that of logistic modeling by weighted least-squares.

In particular, let

n_{xy} = number of women queried about pair (x,y)

$n_{xy,j}$ = number of women preferring j , among those queried about (x,y)

$p_{xy} = n_{xy,x}/n_{xy}$

$p_{yx} = n_{xy,y}/n_{xy}$.

TABLE 1

DESIRED FAMILY SIZE CHOICES OF WHITE NORTH CAROLINA WOMEN
UNDER 30, MARRIED TO THEIR FIRST HUSBANDS

Pair		Choice							
x	y	0	1	2	3	4	5	6	n _{xy}
0	1	2	17	--	--	--	--	--	19
0	2	1	--	22	--	--	--	--	23
0	3	3	--	--	22	--	--	--	24
0	4	1	--	--	--	15	--	--	16
0	5	1	--	--	--	--	26	--	27
0	6	2	--	--	--	--	--	25	27
1	2	--	0	19	--	--	--	--	19
1	3	--	1	--	13	--	--	--	14
1	4	--	10	--	--	10	--	--	20
1	5	--	11	--	--	--	9	--	20
1	6	--	13	--	--	--	--	11	24
2	3	--	--	7	11	--	--	--	18
2	4	--	--	12	--	11	--	--	23
2	5	--	--	18	--	--	6	--	24
2	6	--	--	20	--	--	--	6	26
3	4	--	--	--	13	6	--	--	19
3	5	--	--	--	15	--	0	--	17
3	6	--	--	--	22	--	--	6	28
4	5	--	--	--	--	17	4	--	21
4	6	--	--	--	--	14	--	--	14
5	6	--	--	--	--	--	12	11	23
Total		10	52	98	96	73	59	59	447

Following Berkson [3], we define the functions

$$u_{xy} = \begin{cases} \text{logit } p_{xy} = \log_e(p_{xy}/p_{yx}), & 0 < p_{xy} < 1 \\ \log_e(1/(2n_{xy}-1)), & p_{xy} = 0 \\ \log_e(2n_{xy}-1), & p_{xy} = 1. \end{cases}$$

Then u_{xy} and u_{vw} are statistically independent unless $w = x$ and $v = y$, as they are otherwise derived from different groups of subjects making their choices separately.

Under the Bradley-Terry model

$$E_A\{u_{xy}\} = \log_e(\pi_{xy}/\pi_{yx}) = \log_e(\pi_x/\pi_y) = \beta_{xy}, \text{ say,}$$

and

$$\text{Var}_A\{u_{xy}\} = (n\pi_{xy}\pi_{yx})^{-1},$$

where the subscript "A" indicates the asymptotic value of the appropriate moment operator. Further, a consistent estimator of $\text{Var}_A\{u_{xy}\}$ is given

$$\text{by } \hat{v}_{xy} = (n_{xy}p_{xy}p_{yx})^{-1}.$$

Let

$$\begin{matrix} \tilde{u}' & = & (u_{01}, u_{02}, \dots, u_{56}) \\ 1 \times 21 \end{matrix}$$

$$\begin{matrix} \tilde{\beta}' & = & (\beta_{01}, \beta_{02}, \dots, \beta_{06}) \\ 1 \times 6 \end{matrix}$$

$$\begin{matrix} \hat{\tilde{v}}' & = & (\hat{v}_{01}, \hat{v}_{02}, \dots, \hat{v}_{56}) \\ 1 \times 21 \end{matrix}$$

$$\begin{matrix} \tilde{v}_u & = & D^{-1} \\ \tilde{v}_v & & \\ 21 \times 21 \end{matrix}$$

(where $D_{\tilde{z}}$ indicates a diagonal matrix with diagonal \tilde{z}), and

$$\tilde{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

Then the linear model $E\{\tilde{u}\} = \tilde{X} \tilde{\beta}$, $Cov \tilde{u} = \tilde{V}_{\tilde{u}}$ is appropriate in an asymptotic sense. The application of formal weighted least-squares algorithms to this model leads to the BAN estimate of $\tilde{\beta}$ which minimizes Neyman's [23] modified chi-square statistic, and to various minimum modified chi-square test statistics. These also belong to a general class of asymptotically optimal procedures defined by Wald (see Bhapkar [4], Wald [25]).

In particular, the weighted least-squares solution to this model is

$$\hat{\tilde{\beta}} = (\tilde{X}' \tilde{V}_{\tilde{u}}^{-1} \tilde{X})^{-1} \tilde{X}' \tilde{V}_{\tilde{u}}^{-1} \tilde{u}$$

$$\hat{V}_{\tilde{\beta}} = (\tilde{X}' \tilde{V}_{\tilde{u}}^{-1} \tilde{X})^{-1}$$

where $\hat{V}_{\tilde{\beta}}$ is the estimated covariance matrix of $\hat{\tilde{\beta}}$. Estimates $\hat{\pi}_z$ of the π_z may be recovered from $\hat{\tilde{\beta}}$ through $\hat{\pi}_0 = \{1 + \sum_{y=1}^6 e^{-\hat{\beta}_y}\}^{-1}$, $\hat{\pi}_z = \hat{\pi}_0 e^{-\hat{\beta}_0 z}$, $z = 1, \dots, 6$. The covariance matrix of $\hat{\pi}' = (\hat{\pi}_0, \dots, \hat{\pi}_6)$ may be estimated

as $\underline{\underline{H}}\underline{\underline{\hat{\beta}}}'$, where

$$\underline{\underline{H}} = \begin{bmatrix} \underline{\underline{h}}' \\ \underline{\underline{H}}_1 \end{bmatrix}, \quad \underline{\underline{h}} = \hat{\pi}_0 \hat{\pi}_1', \quad \underline{\underline{H}}_1 = \hat{\pi}_1 \hat{\pi}_1' - D_{\hat{\pi}_1},$$

7×6 6×1 6×6

and $\hat{\pi}_1' = (\hat{\pi}_1, \dots, \hat{\pi}_6)$. Finally, a test statistic for fit of the Bradley-Terry model is obtained as

$$\chi_R^2 = \underline{\underline{u}}' \underline{\underline{V}}_u^{-1} \underline{\underline{u}} - \hat{\beta}' (\underline{\underline{X}}' \underline{\underline{V}}_u^{-1} \underline{\underline{X}})^{-1} \hat{\beta},$$

which is asymptotically distributed as χ^2 with D.F. = {dim($\underline{\underline{u}}$)-dim($\underline{\underline{\beta}}$)} = 21-6 = 15. Tests of linear hypotheses about one or several $\underline{\underline{\beta}}$ -vectors may be generated using the estimated covariances of the β 's; similarly, tests about the π_z 's may be constructed.

Table 2 gives weighted least-squares estimates of the β_{0z} 's, along

TABLE 2
ESTIMATED $\underline{\underline{\beta}}$ AND COVARIANCE MATRIX (SYMMETRIC)
FROM WEIGHTED LEAST-SQUARES ANALYSIS OF TABLE 1

$\hat{\beta}'_{\omega}$	= (-2.026, -2.914, -3.214, -2.654, -1.757, -1.795)					
$\underline{\underline{V}}_{\hat{\beta}_{\omega}}$.148	----	----	----	----	----
	.100	.151	----	----	----	----
	.092	.103	.140	----	----	----
	.107	.112	.102	.154	----	----
	.108	.110	.097	.108	.150	----
	.105	.105	.098	.100	.108	.139

with estimated covariance matrix. Table 3 presents these results trans-

TABLE 3
ESTIMATED PREFERENCE PARAMETERS AND (STANDARD ERRORS)
FROM TABLE 1

Choice							χ^2 (fit)
0	1	2	3	4	5	6	D.F. = 15
.013	.097	.237	.319	.182	.074	.077	19.76
(.004)	(.022)	(.041)	(.052)	(.035)	(.016)	(.016)	

formed to the π -scale. The lack of fit (Neyman) chi-square for these data is $\chi^2 = 19.76$ with 15 degrees of freedom.

4. MULTIVARIATE PAIRED-COMPARISON

The data in Table 4 come from Davidson and Bradley [10], who

TABLE 4
A PAIRED COMPARISON OF THREE CHOCOLATE PUDDINGS

Pair		Pattern of Choice (Overall Quality, Taste, Color) Criteria								n_{xy}
x	y	xxx	yxx	xyx	yyx	xxxy	yxy	xyyy	yyyy	
A	C	6	1	1	0	2	1	1	8	20
A	D	3	1	1	1	2	2	1	11	22
C	D	12	0	0	0	2	1	0	8	23

use a multivariate generalization of the Bradley-Terry approach to study the relationship of choices according to the three criteria involved.

Their multivariate model implies that univariate Bradley-Terry models apply to choices according to any single criterion. If one's primary interest is in the parameters of these "marginal" univariate models, it is desirable to be able to avoid the complexities of their multivariate model, while fitting univariate models in a way which still recognizes the multivariate structure of the data. Weighted least-squares methods similar to those of Section 3 may be used for this purpose.

First, consider the vectors

$$p'_{AC} = \frac{1}{20} (6, 1, 1, 0, 2, 1, 1, 8)$$

$$p'_{AD} = \frac{1}{22} (3, 1, 1, 1, 2, 2, 1, 11)$$

$$p'_{CD} = \frac{1}{23} (12, 0, 0, 0, 2, 1, 0, 8)$$

generated from the rows of Table 4. The vector p_{AC} is multinomially distributed with

$$E\{p_{AC}\} = \pi_{AC}, \quad \text{say}$$

and

$$\text{Cov}\{p_{AC}\} = \frac{1}{20} (D_{\pi_{AC}} - \pi_{AC}\pi'_{AC}) = V_{\pi_{AC}}.$$

Further, $E_A \{D_{p_{AC}} - p_{AC}p'_{AC}\} = \{D_{\pi_{AC}} - \pi_{AC}\pi'_{AC}\}$. Similar results hold for

p_{AD} and p_{CD} . Thus, the vector $p' = (p'_{AC}, p'_{AD}, p'_{CD})$ has expectation

$\pi' = (\pi'_{AC}, \pi'_{AD}, \pi'_{CD})$, and a covariance matrix V_{π} which may be consistently estimated by

$$V_{\pi} = \begin{bmatrix} V_{\pi_{AC}} & 0 & 0 \\ 0 & V_{\pi_{AD}} & 0 \\ 0 & 0 & V_{\pi_{CD}} \end{bmatrix}$$

where $V_{\sim P_{AC}}$, for instance, $= \frac{1}{n_{AC}} (D_{\sim P_{AC}} - P_{AC} P_{AC}')$.

Now, let $K_{9 \times 18} = [1 \ -1] \otimes I_9$ and

$$A_{18 \times 24} = A^*_{6 \times 8} \otimes I_3, \quad A^*_{6 \times 8} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix},$$

where \otimes denotes the Kronecker matrix product. Then the vector

$$u = K \log A p$$

consists of functions of the data identical to those used in Section 3, the first, fourth, and seventh relating to choice on color, the second, fifth and eighth to choice on taste, and the remainder to choice on overall quality. The covariance matrix of u may be estimated consistently by

$$V_{\sim u} = K D_{\sim A p}^{-1} (A V_{\sim p} A') D_{\sim A p}^{-1} K'.$$

Hence by an asymptotic argument identical to that in Section 3, an appropriate procedure for fitting the marginal Bradley-Terry models is to look at the linear model

$$\begin{aligned} E u &= X \beta \\ \text{Cov } u &= V_{\sim u} \end{aligned}$$

where

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

$$\underset{\sim}{\beta}' = (\beta_{1;AC}, \beta_{2;AC}, \beta_{3;AC}, \beta_{1;AD}, \beta_{2;AD}, \beta_{3;AD}),$$

$$\beta_{\alpha;xy} = \log \left(\frac{\pi_{\alpha;x}}{\pi_{\alpha;y}} \right),$$

and the values of $\alpha = 1, 2, 3$ correspond to the criteria color, taste, and overall quality, respectively. Weighted least-squares applied to this model gives estimates of the $\hat{\beta}_{\alpha;xy}$ which may be used to solve as in Section 3 for the $\hat{\pi}_{\alpha;z}$. The lack of fit statistic generated by this data is $X^2 = 3.05$ with D.F. = 3, indicating that the Bradley-Terry formulation is realistic.

Table 5 gives the estimates $\hat{\pi}_{\alpha;z}$, with their estimated standard errors. For comparison we have calculated the weighted least-squares and maximum likelihood estimates of these parameters using the univariate margins obtained from Table 4. We also reproduce (from [10]) maximum likelihood estimates, under Davidson and Bradley's [8,9,10] multivariate model. Note that these latter depart considerably from the other sets of estimates, which are very similar. This is possibly due to lack of fit of the multivariate model [8].

5. CHOICES AMONG MORE THAN ONE PAIR

In the previous two examples, each individual chose only between one pair of elements of the object set O . In many paired-comparison sit-

TABLE 5

ESTIMATES OF PREFERENCE PARAMETERS (STANDARD ERRORS)
FOR CHOCOLATE PUDDINGS, BY VARIOUS METHODS

Criterion	Method	Pudding		
		A	C	D
Overall Quality	WLS Multivariate	.255 (.061)	.377 (.071)	.368 (.070)
	WLS Univariate	.256 (.061)	.385 (.073)	.359 (.072)
	ML Univariate	.253	.387	.360
	Davidson-Bradley	.297	.362	.341
Taste	WLS Multivariate	.273 (.062)	.399 (.074)	.329 (.067)
	WLS Univariate	.272 (.063)	.408 (.076)	.320 (.068)
	ML Univariate	.270	.410	.320
	Davidson-Bradley	.302	.401	.297
Color	WLS Multivariate	.190 (.051)	.360 (.071)	.450 (.076)
	WLS Univariate	.202 (.053)	.379 (.073)	.420 (.077)
	ML Univariate	* .200	.379	.421
	Davidson-Bradley	.238	.372	.390

uations, individuals are asked to consider more than one pair. To obtain a practical analysis, it has usually been necessary to assume that choices within different pairs, made by the same individual, were statistically independent random variables. This assumption is often unwarranted or questionable at the very least. However, in many instances this assumption is unnecessary, as one can estimate and adjust for an unknown dependence structure relating choices between different pairs in the same manner that was used to adjust for dependence of choices according to different criteria in the last example.

Table 6 gives data from a paired-choice study conducted by the

TABLE 6
A PAIRED-COMPARISON OF FOUR SOFT DRINKS

Pairs	Patterns of Choice (Taste in Pair 1, Fizz in Pair 1, T in 2, F in 2)															
	ii kk	ii kl	ii lk	ii ll	ij kk	ij kl	ij lk	ij ll	jk kk	jk kl	jk lk	jk ll	jl kk	jl kl	jl lk	jl ll
(i,j),(k,l)																
(C,D),(D,B)	3	3	1	5	1	3	2	1	0	0	0	6	1	1	0	8
(C,A),(A,D)	0	0	0	1	2	0	2	1	1	1	0	0	3	1	2	0
(C,B),(A,B)	1	0	0	0	1	0	0	0	1	0	0	0	5	2	3	5

authors to compare two cola soft drinks with two "uncolas." Each person in the study was asked to choose within each of two pairs of unidentified samples of different drinks, on the basis of a) taste and b) fizziness. As in Section 4, the three rows of the table came from independent multinomial distributions. The strung-out vector of proportions

$$p' = \left(\frac{3}{35}, \dots, \frac{8}{35}, 0, 0, 0, \frac{1}{14}, \dots, 0, \frac{1}{18}, \dots, \frac{5}{18} \right)$$

has covariance matrix V_p which may be consistently estimated exactly as in Example 2.

Let $\underline{u} = K[\log A_p]$, where

$$K_{12 \times 24} = [1 \ -1] \otimes I_{12}, \quad A_{24 \times 48} = A^* \otimes I_3,$$

and

$$A^*_{8 \times 16} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

Again, a consistent estimator for $\text{Cov } \underline{u}$ is $\underline{V}_u = K D_{Ap}^{-1} (A V A') D_{Ap}^{-1} K'$. The Bradley-Terry model implies that

$$E_A \{ \underline{u} \} = X \beta$$

where

$$\beta' = (\beta_{T;AB}, \beta_{T;AC}, \beta_{T;AD}, \beta_{F;AB}, \beta_{F;AC}, \beta_{F;AD}),$$

and

$$X_{12 \times 6} = \begin{bmatrix} 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

The first three parameters of β correspond to choice with respect to taste, the second three to choice according to fizziness. Table 7 gives $\hat{\beta}$ and

TABLE 7
ESTIMATED β AND COVARIANCE MATRIX (SYMMETRIC)
FOR PAIRED-COMPARISON OF SOFT DRINKS

$\hat{\beta}' =$	(-.007,	.767,	.750,	.249,	1.54,	1.47)
$V_{\hat{\beta}}$.137	---	---	---	---	---
	.059	.141	---	---	---	---
	.089	.098	.118	---	---	---
	.051	.003	.026	.149	---	---
	.006	.017	.004	.085	.168	---
	.018	.019	.020	.095	.120	.151

its estimated covariance matrix $V_{\hat{\beta}} = (X'V^{-1}X)^{-1}$ obtained from this model by weighted least-squares. Table 8 displays the estimates of the pref-

TABLE 8
ESTIMATED PREFERENCE PARAMETERS (STANDARD ERRORS)
FROM PAIRED-COMPARISON OF SOFT DRINKS

Criteria	Soft Drink			
	A	B	C	D
Taste	.340 (.071)	.342 (.068)	.158 (.041)	.160 (.029)
Fizziness	.450 (.086)	.350 (.075)	.096 (.029)	.103 (.027)

erence parameters $\pi_{T;ij}$ and $\pi_{F;ij}$ associated with the marginal Bradley-Terry models for taste and fizziness comparisons. The lack of fit statistic is $X^2 = 7.16$ with D.F. = 6.

At this stage of the analysis, comparisons of various soft drinks may be made using test statistics associated with linear hypotheses about β . If C is a matrix of rank d , a test of $H_0: C\beta = 0$ is the statistic

$(\hat{C}\hat{\beta})' (CV_{\hat{\beta}}C')^{-1} (\hat{C}\hat{\beta})$, which under H_0 is asymptotically distributed as χ^2 with D.F.= d .

For example, to test $\pi_{T;C} = \pi_{T;D}$ we may use $C = (0 \ 1 \ -1 \ 0 \ 0 \ 0)$. To test $\pi_{T;A} = \pi_{T;B}$, $C = (1 \ 0 \ 0 \ 0 \ 0 \ 0)$. Likewise, $C = (1 \ -1 \ -1 \ 0 \ 0 \ 0)$ tests whether the geometric mean of $\pi_{T;A}$, $\pi_{T;B}$ equals that of $\pi_{T;C}$, $\pi_{T;D}$. Various test statistics generated in this way are given in Table 9. The

TABLE 9

CHI-SQUARE STATISTICS FOR HYPOTHESES ABOUT SOFT DRINK PREFERENCE PARAMETERS

Hypothesis	Test Statistics	D.F.
$\pi_{T;A} = \pi_{T;B} = \pi_{T;C} = \pi_{T;D} =$ $\pi_{F;A} = \pi_{F;B} = \pi_{F;C} = \pi_{F;D} = .25$	29.49**	6
$\pi_{T;A} = \pi_{T;B} = \pi_{T;C} = \pi_{T;D} = .25$	9.50*	3
$\pi_{F;A} = \pi_{F;B} = \pi_{F;C} = \pi_{F;D} = .25$	22.31**	3
$\pi_{T;A} = \pi_{T;B}, \pi_{F;A} = \pi_{F;B}$.49	2
$\pi_{T;C} = \pi_{T;D}, \pi_{F;C} = \pi_{F;D}$.06	2
$\pi_{T;AB} = \pi_{T;CD}, \pi_{F;AB} = \pi_{F;CD}$	25.76**	2
$\pi_{T;AB} = \pi_{T;CD}$	7.88**	1
$\pi_{F;AB} = \pi_{F;CD}$	21.85**	1

first three test statistics show significant differences between soft-drinks on both dimensions of taste and fizziness. The next two indicate that neither the two colas nor the two uncolas were significantly different from each other. The bottom three lines give pooled tests for differences between types of soft drink (cola vs. uncola); here $\pi_{T;AB} = (\pi_{T;A} \pi_{T;B})^{1/2}$, for example. In this (non-controlled) survey (on a self-selected sample), uncolas were preferred on both dimensions of comparison (on the day the experiment was performed).

6. SUB-GROUP COMPARISONS OF PREFERENCE PARAMETERS AND INTERACTIONS

As part of a driver's license examination, examinees were asked the following questions

1. "If you were stopped for running a traffic light would you prefer to _____ or _____?" where the blanks were filled (in random order) with two of the following choices:
 - a) Receive a ticket with a \$35.00 fine. (A)
 - b) Attend a 3-day seminar (free of charge) on highway safety. (B)
 - c) Take a three-week (1 1/2 hours per day) drivers education course (free of charge) where you must attend lectures and drive under the supervision of an instructor. (C)

2. "If you were provided with free installation, would you prefer to have your car installed with _____ or _____ in order to provide resistance to crash injuries?" where the blanks were filled (in random order) with two of the following choices:
 - a) Seat belts (lap and shoulder). (X)
 - b) Crash bags. (Y)
 - c) Ejection seats. (Z)

The choices were given in pairs in an effort to reduce the tendency of respondents to give "take the easy way out" or give the response with which they are most familiar. As part of the examination a short (10 min.) film was shown illustrating the effects of seat belts, crash bags, and ejection seats on dummies on impact at 45 mph. The resulting artificial data are given in Table 10.

TABLE 10
PREFERENCES OF DRIVER LICENSE EXAMINEES, BY AGE AND SEX

Age Group	Sex	Paired Choices	Responses									
			XA	XB	XC	YA	YB	YC	ZA	ZB	ZC	TOTAL
Young (<u><</u> 25 Years)	Male	XY,AB	18	15	--	25	23	--	--	--	--	81
		XZ,AC	12	--	9	--	--	--	26	--	19	66
		YZ,BC	--	--	--	--	22	17	--	25	16	80
Young	Female	XY,AB	29	22	--	20	18	--	--	--	--	89
		XZ,AC	32	--	35	--	--	--	19	--	12	98
		YZ,BC	--	--	--	--	19	22	--	6	10	57
Mature (<u>></u> 25 Years)	Male	XY,AB	17	20	--	--	12	--	--	--	--	61
		XZ,AC	19	--	13	--	--	--	7	--	5	44
		YZ,BC	--	--	--	--	11	6	--	7	7	31
Mature	Female	XY,AB	18	14	--	4	1	--	--	--	--	37
		XZ,AC	15	--	9	--	--	--	8	--	4	36
		YZ,BC	--	--	--	--	14	10	--	10	5	39

Each row of Table 10 displays a strung-out incomplete 3x3 contingency table, with five of nine cells constrained to be zero. Thus, each row represents bivariate responses (Question 1, Question 2) of a group of individuals of particular levels of the cross-classified factors "Age" (2 levels), "Sex" (2 levels), "Pairs Presented" (3 levels). Our interest chiefly relates to the individual questions (univariate margins of the bivariate response),

and how these vary with Age and Sex. Note that this data is structurally identical to a "split-plot categorical data model" as defined by Koch and Reinfurt [19], where Age, Sex and Pairs Presented are whole-plot factors, Question (1 or 2) is a split-plot factor, and where each of the component mixed models (rows of Table 10) is incomplete. However, our interest is not in comparisons between Questions 1 and 2, and hence a somewhat different analysis is necessary. We will use preference ratios from marginal Bradley-Terry models to summarize data on a single question over different levels of the Pairs Presented factor. We then analyze these preference ratios in terms of Age and Sex. Though our primary interest is in the single questions rather than in their relationship, the analysis is done jointly for both questions to account for such covariance structure as may be present in the data.

Initially, we proceed by jointly describing marginal models for each demographic subgroup, using essentially the same method as in previous sections. Let

$$P'_{YM} = \left(\frac{18}{81}, \frac{15}{81}, \frac{25}{81}, \frac{23}{81}, \frac{12}{66}, \dots, \frac{17}{80} \right)$$

be the vector of observed proportions for Young Males (from Table 10), and

$$\beta'_{YM} = (\beta_{YM;XY}, \beta_{YM;XZ}, \beta_{YM;AB}, \beta_{YM;AC})$$

be the "log-ratio" vector derived from preference parameters of young males on the two questions, taken with numerator choices X and A respectively. Define

$$P_{YF}, P_{MM}, P_{MF}, \beta_{YF}, \beta_{MM}, \beta_{MF}$$

analogously for the three succeeding layers in Table 10. Then let

$$\underset{1 \times 16}{\beta'} = (\underset{4}{\beta'_{YM}}, \underset{4}{\beta'_{YF}}, \underset{4}{\beta'_{MM}}, \underset{4}{\beta'_{MF}})$$

and

$$\underset{1 \times 48}{P'} = (\underset{12}{P'_{YM}}, \underset{12}{P'_{YF}}, \underset{12}{P'_{MM}}, \underset{12}{P'_{MF}}).$$

Taking

$$\underset{4 \times 4}{A^*} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix},$$

$$\underset{6 \times 4}{X^*} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad \underset{1 \times 2}{K^*} = [1 \ -1],$$

we have $E_{\underset{\sim}{A}}\{\underset{\sim}{u}\} = \underset{\sim}{X} \underset{\sim}{\beta}$, with $\underset{24 \times 1}{\underset{\sim}{u}} = \underset{\sim}{K} [\log \underset{48 \times 48}{\underset{\sim}{A}} \underset{\sim}{p}]$, $\underset{48 \times 48}{\underset{\sim}{A}} = \underset{\sim}{A}^* \otimes \underset{12}{I_{12}}$, $\underset{24 \times 48}{\underset{\sim}{K}} = \underset{\sim}{K}^* \otimes \underset{24}{I_{24}}$,

and $\underset{24 \times 16}{\underset{\sim}{X}} = \underset{\sim}{X}^* \otimes \underset{4}{I_4}$.

Now, reparametrize these models to $\underset{16 \times 1}{\underset{\sim}{\eta}_1}$, where $\underset{\sim}{\beta} = \underset{16 \times 1}{Z} \underset{\sim}{\eta}_1$, and

$$\underset{16 \times 16}{Z_1} = \begin{bmatrix} (1 \ 1 \ 1 \ 1) \otimes I_4 \\ (1 \ 1 \ -1 \ -1) \otimes I_4 \\ (1 \ -1 \ 1 \ -1) \otimes I_4 \\ (1 \ -1 \ -1 \ 1) \otimes I_4 \end{bmatrix}.$$

Z expresses each log preference-ratio in the usual two-way factorial model.

Thus,

$$\underset{16}{\underset{\sim}{\eta}_1} = (\underset{4}{\eta'_{XY}}, \underset{4}{\eta'_{XZ}}, \underset{4}{\eta'_{AB}}, \underset{4}{\eta'_{AC}}).$$

where $\eta'_{XY} = (\mu_{XY}, \alpha_{XY}, \delta_{XY}, \gamma_{XY})$, η_{XZ} , η_{AB} , η_{AC} are written similarly, and α 's, δ 's, γ 's represent, respectively, effects "due to" Age, Sex, and their interaction on the subscripted log preference-ratio. The usual generalized least-squares expressions for $\text{Cov } u$, $\hat{\eta}_1$, and $\text{Cov } \hat{\eta}_1$ apply with design matrix $Y = XZ_1$.

This formulation is useful because we may now express many hypotheses of interest in the form $C\eta_1 = 0$, if C is chosen appropriately. The method of Section 5 may then be used to generate test statistics for these hypotheses. For example, the matrix

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

2x16

can be used to represent the "no interaction" hypothesis as applied to log preference-ratios for question #2. If this hypothesis is true, we have a simple multiplicative model on the scale of preference-ratios. We can then make a statement of the form: "The odds of an individual preferring a ticket to a seminar are k_{XY} times greater if the person is 25 or younger than if older, regardless of sex." In such a case, we may drop γ terms from the model and test for the significance of $k_{XY} = \exp(\alpha_{XY})$.

Table 11 gives the values of test statistics for several hypotheses

TABLE 11

DRIVER'S LICENSE EXAMINEES:
CHI-SQUARE STATISTICS FOR MODEL I

Source Hypothesis	D.F.	X ²	P-Level
Marginal Bradley-Terry Fit	8	8.89	.35
No Demographic Differences	12	53.16***	<.001
Question #1	6	45.50***	<.001
Question #2	6	7.60	.27
No Age by Sex Interaction	4	10.16	.38
Question #1	2	2.46	.29
AB Comparison	1	0.31	.58
AC Comparison	1	2.43	.12
Question #2	2	7.81*	.02
XY Comparison	1	0.05	.82
XZ Comparison	1	6.37*	<.02

of interest for Table 10. Examination of the $\hat{\beta}_{XZ}$ by eye, or by Scheffe's multiple comparison technique (as described e.g. in Goodman [12]), or examination of Table 10 indicate clearly the nature of the Age x Sex interaction on question #2. The data reveal a striking difference between young men and the other three demographic groups, which themselves appear similar. Young men are much more likely to prefer ejection seats to seat belts than other demographic groups. We incorporate this observations and the other results of Table 11 into a new model with no interaction on log preference ratios β_{XY} , β_{AB} , β_{AC} and a simple two-group model of Young Males and Others for the β_{XZ} . For this Model II $\beta = Z_2 \eta_2$, where Z_2 is obtained by deleting from Z_1 columns 4, 12, and 16 (corresponding to γ_{XY} , γ_{AB} , γ_{AC}), and replacing columns 5 - 8 (η_{XZ}) with \bar{Z}_2 , where

$$\bar{Z}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

Here $\eta_2' = (\mu_{XY}, \alpha_{XY}, \delta_{XY}, \beta_{YM;XZ}, \beta_{OTHER;XZ}, \mu_{AB}, \alpha_{AB}, \delta_{AB}, \mu_{AC}, \alpha_{AC}, \delta_{AC})$.

Lack of fit for Model II is $\chi^2 = 12.46$ with D.F. = 13, so the fit is adequate.

Within the context of Model II, we may test for various "main effects." Results of such tests are given in Table 12. These allow us

TABLE 12

DRIVER'S LICENSE EXAMINEES:
CHI-SQUARE STATISTICS FOR MODEL II

Source Hypothesis	D.F.	χ^2	P-Level
Fit of Model II	13	12.46	.49
No Age Effect			
Question #1	2	2.70	.26
AB Comparison	1	0.03	.87
AC Comparison	1	1.92	.17
Question #2			
XY Comparison	1	16.42***	<.001
No Sex Effect			
Question #1	2	2.74	.25
AB Comparison	1	1.40	.24
AC Comparison	1	0.24	.62
Question #2			
XY Comparison	1	7.25**	.01

a further reduction to a seven-parameter model by eliminating α_{AB} , δ_{AB} , α_{AC} , δ_{AC} from Model II. This is accomplished simply by deleting columns 7, 8, 10 and 11 from Z_2 . Lack of fit for this final Model III is $\chi^2 = 17.74$ (D.F. = 17). Estimated parameters and standard errors for Model III are shown in Table 13. Estimated preference parameters (π 's) are shown in Table 14.

TABLE 13

DRIVERS LICENSE EXAMINEES:
ESTIMATED PARAMETERS AND STANDARD ERRORS FOR MODEL III

Parameter	μ_{XY}	α_{XY}	δ_{XY}	$\beta_{YM;XZ}$	$\beta_{OTHER;XZ}$	μ_{AB}	μ_{AC}
Estimate	.300	.449	.276	.928	1.052	.113	.287
Standard Error	.115	.109	.105	.136	.240	.103	.106

TABLE 14

DRIVERS LICENSE EXAMINEES:
ESTIMATED PREFERENCE PARAMETERS (STANDARD ERRORS),
FROM MODEL III, BY DEMOGRAPHIC SUBGROUP

Age	Sex	Question 1			Question 2		
		A	B	C	X	Y	Z
Young	Male	.378 (.021)	.338 (.021)	.284 (.019)	.231 (.028)	.353 (.035)	.416 (.042)
Young	Female	.378 (.021)	.338 (.021)	.284 (.019)	.439 (.033)	.387 (.038)	.174 (.020)
Mature	Male	.378 (.021)	.338 (.021)	.284 (.019)	.495 (.035)	.309 (.039)	.196 (.022)
Mature	Female	.378 (.021)	.338 (.021)	.284 (.019)	.570 (.037)	.205 (.035)	.225 (.022)

Transforming back to the preference-ratio scale, results of our analysis may be summarized as follows:

1. A \$35 ticket is the least unpopular of the three alternatives for question #1. The odds are estimated at 1.33:1 that a randomly chosen driver candidate will prefer such a ticket to a three week course, and only slightly less (1.12:1) that it will be preferred to a three-day seminar. There is no evidence that these odds change with Age or Sex of the individual driver candidate.
2. Young male candidates prefer crash bags to seat belts (lap and shoulder) by a margin estimated at 1.53:1. Mature male candidates prefer seat belts by 1.60:1. Regardless of age, the odds that a female candidate prefers seat belts are 1.74 times the (Age-specific) odds for males. Young males also prefer ejections seats to seat belts, these in the ratio 1.80:1. All other groups prefer seat belts by a 2.53:1 margin.

In Table 11, note that the chi-square statistics for Questions 1 and 2 add almost exactly to the joint statistics for both questions. This situation is precisely analogous to a MANOVA in which the univariate test statistics sum to the multivariate test, and reflects the virtual lack of association in this data between responses to Question 1 and Question 2. In fact, the test of association formed by summing the twelve uncorrected Pearson chi-square statistics for the 2x2 subtables extracted from each row of Table 10 yields $X^2 = 4.42$, D.F. = 12. The very low value of this statistic excludes association. This absence of interaction allows us to regard the marginal Bradley-Terry models described above, acting independently, as totally sufficient to characterize the data in Table 10.

However, it is important to note that the methods of this section are applicable when interaction is present. The marginal Bradley-Terry models will still be of greatest interest, and joint tests of models and hypotheses for both questions fulfill the same function in the paired comparison situation as standard MANOVA in the conventional setting of bivariate continuous data. When interaction is present, the association between questions itself may be the subject of study. A measure of association (e.g., the log cross-product ratio) may be extracted from each row of Table 10, using \underline{A}^* and \underline{K} matrices defined appropriately. The measure of association for each combination of comparisons $((A,B),(X,Y))$ may then be modeled for effects of Age and Sex, say, in the same way that observed marginal preference ratios were modeled above. Such a procedure can give considerable insight into the pattern of relationship between questions, and sometimes can lead to more powerful procedures for the marginal analyses (in this regard, see Ameniya [1]).

Finally, observe that the factor Pairs Presented is a design component of our paired comparison survey. For simplicity of exposition, we have presented data for which levels of this design factor are identical for the various combinations of Age and Sex. This symmetry is in no way essential to the analysis. All that is necessary is for the preference ratios (β 's) to be estimable parameters in the usual least-squares sense. We may vary the choice presented, both within and across subgroups, in any manner that satisfies this estimability condition.

7. SAMPLE SIZE CONSIDERATIONS AND DISCUSSION

As noted in Section 2, the examples in this paper were chosen for purposes of illustration without much regard for sample size, which

must effect the validity of the asymptotic analysis used. The analysis depends upon both convergence to normality (and zero bias) of the functions $K[\log(A_p)]$ (log ratios) derived from the data, and consistency of the estimates used for covariance of these functions. In either case, convergence rates depend upon the precise nature of the functions $K[\log(A_p)]$, and the structures of the populations from which they derive. This makes it very difficult to give rules of thumb for sample size requirements.

Nevertheless, some remarks are possible. It should be noted that the asymptotics of importance are determined by individual cell probabilities only insofar as they contribute to the margins or other functions $K[\log A_p]$. Thus, we may begin with contingency tables quite sparsely populated, provided our analysis is directed at margins or other functions obtained by collapsing the table or by pooling many cells. In Table 4, the four zeroes in row (C,D) are thus unimportant, as the three margins computed from this row are each well-distributed with respect to the appropriate two choices. The nine zeroes in the last row of Table 6 are more worrisome, as the observed marginal frequency of individuals choosing C over B for taste is only 2. Several of the rows of Table 1 are also disturbing. (Note that the definition of u_{XY} in Section 3 incorporates a technique of smoothing zero cell counts.) However, provided that sample sizes are large enough so that central limit theory may reasonably be applied to most of the elements of $K[\log(A_p)]$ (e.g., each marginal category occurring with frequency ≥ 5 for marginal logits), modeling such as we have done should be fairly robust to sample sizes of the other elements except in truly pathological situations. The experience of obtaining reasonable analyses from many data sets with "marginally small" samples is reassuring on this score.

For a more elaborate discussion of sample size requirements in relation to the general methodology of Grizzle, Starmer and Koch [14], the reader is referred to Koch, Freeman, Freeman, and Lehnen [17]. An extension of the general approach by Tolley and Koch [24] somewhat alleviates the problem, at cost in terms of conceptual simplicity and computational convenience.

In summary, we have suggested a method of treating paired comparison data by use of a non-iterative logit analysis. Paired comparison data are viewed as forming multi-dimensional incomplete contingency tables. Categorizations of respondent populations are viewed as dimensions of these tables; multivariate paired comparisons for several such populations are formally similar to incomplete "split-plot categorical data," but the questions of interest dictate a different analysis. The primary advantage of the method we present is its flexibility. Thus, the experimenter may take a more relaxed attitude towards symmetry concerns in planning his experiment and still achieve an informative analysis of reasonable parsimony and power.

ACKNOWLEDGMENTS

The authors would like to thank Sam A. Caltagirone and Daniel H. Freeman, Jr. for their diligent computer work and Rebecca L. Wesson for her spectacular typing.

REFERENCES

- [1] Ameniya, T., "Bivariate Probit Analysis: Minimum Chi-Square Methods," Journal of the American Statistical Association, 69 (December 1974), 940-44.
- [2] Atkinson, A. C., "A Test of the Linear-Logistic and Bradley-Terry Models," Biometrika, 59 (April 1972), 37-42.
- [3] Berkson, J., "Maximum Likelihood and Minimum Chi-Square Estimates of the Logistic Functions," Journal of the American Statistical Association, 50 (March 1955), 130-62.
- [4] Bhapkar, V. P., "A Note on the Equivalence of Two Test Criteria for Hypotheses in Categorical Data," Journal of the American Statistical Association, 61 (March 1966), 228-35.
- [5] Bishop, Y. M. M. and Fienberg, S. E., "Incomplete Two-Dimensional Contingency Tables," Biometrics, 25 (March 1969), 119-28.
- [6] Bradley, R. A. and Terry, M. B., "Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons," Biometrika, 39 (December 1952), 324-45.
- [7] David, H. A., The Method of Paired Comparisons, New York: Hafner Publishing Company, Inc., 1963.
- [8] Davidson, R. R. and Bradley, R. A., "Multivariate Paired Comparisons: The Extension of a Univariate Model and Associated Estimation and Test Procedures," Biometrika, 56 (April 1969), 81-95.
- [9] Davidson, R. R. and Bradley, R. A., "Multivariate Paired Comparisons: Some Large-Sample Results on Estimation and Tests of Equality of Preference," in M. L. Puri, ed., Nonparametric Techniques in Statistical Inference, Cambridge: Cambridge University Press, 1970, 111-25.
- [10] Davidson, R. R. and Bradley, R. A., "A Regression Relationship for Multivariate Paired Comparisons," Biometrika, 58 (December 1971), 555-560.
- [11] Fienberg, S. E., "An Iterative Procedure for Estimation in Contingency Tables," Annals of Mathematical Statistics, 41 (June 1970), 901-17.
- [12] Goodman, L. A., "Simultaneous Confidence Intervals for Contrasts Among Multinomial Populations," Annals of Mathematical Statistics, 35 (June 1964), 716-25.
- [13] Goodman, L. A., "The Analysis of Cross-Classified Data: Independence, Quasi-Independence, and Interactions in Contingency Tables With or Without Missing Entries," Journal of the American Statistical Association, 63 (December 1968), 1091-131.

- [14] Grizzle, J. E., Starmer, C. F. and Koch, G. G., "Analysis of Categorical Data by Linear Models," Biometrics, 25 (September 1969), 489-504.
- [15] Imrey, P. B. and Koch, G. G., "Linear Models Analysis of Incomplete Multivariate Categorical Data," Mimeo Series No. 820, Institute of Statistics, University of North Carolina, 1972.
- [16] Koch, G. G., Abernathy, J. A. and Imrey, P. B., "On a Method for Studying Family Size Preferences," Demography, 12 (February 1975), 57-66.
- [17] Koch, G. G., Freeman, J. L., Freeman, D. H., and Lehnen, R. G., "A General Methodology for the Analysis of Experiments with Repeated Measurement of Categorical Data," Mimeo Series No. 961, Institute of Statistics, University of North Carolina, 1974.
- [18] Koch, G. G., Imrey, P. B. and Reinfurt, D. W., "Linear Model Analysis of Categorical Data with Incomplete Response Vectors," Biometrics, 28 (September 1972), 663-92.
- [19] Koch, G. G. and Reinfurt, D. W., "The Analysis of Categorical Data from Mixed Models," Biometrics, 27 (March 1971), 157-73.
- [20] Rao, P. V. and Kupper, L. L., "Ties in Paired-Comparison Experiments: A Generalization of the Bradley-Terry Model," Journal of the American Statistical Association, 62 (March 1967), 194-204.
- [21] Luce, R. D., Individual Choice Behavior: A Theoretical Analysis, New York: John Wiley and Sons, Inc., 1959.
- [22] Mantel, N., "Incomplete Contingency Tables," Biometrics, 26 (June 1970), 291-304.
- [23] Neyman, J., "Contribution to the Theory of the Chi-Square Test," in J. Neyman, ed., Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles: University of California Press, 1949, 239-72.
- [24] Tolley, H. D. and Koch, G. G., "A Two-Stage Approach to the Analysis of Longitudinal Type Categorical Data," Mimeo Series No. 962, Institute of Statistics, University of North Carolina, 1974.
- [25] Wald, A., "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations Is Large," Transactions of the American Mathematical Society, 54 (November 1943), 426-82.
- [26] Williams, O. D. and Grizzle, J. E., "Analysis of Categorical Data With More than One Response Variable by Linear Models," Mimeo Series No. 715, Institute of Statistics, University of North Carolina.