Loglinear Representation for Paired and Multiple Comparisons Models

by

Stephen E. Fienberg*
Kinley Larntz

Technical Report # 252a

Department of Applied Statistics
School of Statistics
University of Minnesota
St. Paul, Minnesota 55108

*Currently on leave in the Center for Analysis of Health Practices and the Department of Statistics, Harvard University.

August 19, 1975
(revised November 12, 1975)

Loglinear Representation for Paired and Multiple Comparisons Models

by

Stephen E. Fienberg*
Kinley Larntz

Department of Applied Statistics
School of Statistics
University of Minnesota
St. Paul, Minnesota 55108

*Currently on leave in the Center for Analysis of Health Practices and the Department of Statistics, Harvard University.

August 19, 1975
(Revised November 12, 1975)

SUMMARY

This paper develops loglinear model representations for the Bradley-Terry paired comparisons and Luce multiple comparisons models. Various multivariate extensions of the Bradley-Terry model are then considered. A model is presented, similar to one of Davidson and Bradley (1969), but which is based on the log-odds ratio as a measure of association rather than the ϕ-coefficient of correlation. It is shown that, by the adaptive use of loglinear model theory, complex data sets involving paired and multiple comparisons as well as rankings may be easily analyzed. In all cases estimates of cell expectations may be computed using an iterative proportional scaling algorithm that is computationally easy to carry out.

<u>Some key words</u>: Iterative proportional scaling; Loglinear models; Maximum likelihood estimation; Multiple comparisons; Multivariate attributes; Paired comparisons; Quasi-independence; Quasi-symmetry; Ranked preferences.

1. INTRODUCTION

Methods of paired comparisons have undergone extensive development in the statistical literature, and these methods have been applied in a variety of experimental contexts where judges or subjects are asked to make subjective appraisals of treatments, items, or individuals presented in pairs. The two methods of paired comparisons most commonly used in practice are the ones due to Thurstone and Mosteller (see e.g., Mosteller, 1951) and to Bradley and Terry (1952).

In this paper we reexamine the Bradley-Terry model as a loglinear model, and we show that, depending on the form chosen to display the paired comparisons data, the Bradley-Terry model corresponds either to the model of quasi-independence (see Bishop, Fienberg, and Holland, 1975; Goodman, 1968) or to the model of quasi-symmetry (see Bishop, Fienberg and Holland, 1975; Caussinus, 1966). The method of iterative proportional fitting used for the loglinear model problems can be shown to be equivalent to the iterative method of Bradley and Terry.

The major advantage of the loglinear model representation is the ease with which it can be generalized to handle various multivariate extensions to the Bradley-Terry model. Three extensions examined in this paper are:

(1) rankings in multiple comparisons (Pendergrass and Bradley, 1960);

(2) preference choices in paired and multiple comparisons (Luce, 1959);

(3) multivariate paired comparisons (Sen and David, 1968; Davidson and Bradley, 1969).

In the latter case, we propose an alternative to the Davidson-Bradley model which includes as a special case the null model of Sen and David. Since all the models we consider are log-linear in form, we can use the general theory for loglinear models (see Andersen, 1974; Haberman, 1974) to ensure the existence of unique maximum likelihood estimates of parameters, the specification of minimal sufficient statistics and likelihood equations, and finally the iterative solution of the likelihood equations.

The analysis of paired and multiple comparisons problems in a contingency table format provides an important link between two active areas of statistical research and application.

2. THE BRADLEY-TERRY MODEL

Suppose t items are to be compared in pairs by sets of judges. The Bradley-Terry model postulates that the probability that treatment T_i is preferred to treatment T_j is

$$Pr(T_i \succ T_j) = \pi_i/(\pi_i + \pi_j) \qquad i \neq j, i, j=1,2,\ldots,t, \quad (2.1)$$

where the parameters $\pi_i \geq 0$ for $i=1,2,\ldots t$, and $\pi_1+\ldots+\pi_t = 1$. The model assumes the independence of the ratings of the same pair by different judges and different pairs by the same judge. This model does not permit a decision of "no preference," but various extensions discussed by Bradley (1976) do.

Let T_i be compared with T_j $n_{ij} > 0$ times, for $i < j=1,2,...,t$. Then we have a total of

$$N= \sum_{i<j} n_{ij} \tag{2.2}$$

paired comparisons. Let x_{ij} be the observed number of times T_i is preferred to T_j in the n_{ij} comparisons. Then

$$x_{ij} + x_{ji} = n_{ij}, \tag{2.3}$$

and the overall likelihood function for the N paired comparisons is

$$\prod_{i<j} \binom{n_{ij}}{x_{ij}} \left(\frac{\pi_i}{\pi_i+\pi_j}\right)^{x_{ij}} \left(\frac{\pi_j}{\pi_i+\pi_j}\right)^{n_{ij}-x_{ij}} . \tag{2.4}$$

The kernel of the loglikelihood can then be written as

$$\sum_{i=1}^{t} x_{i+}\log\pi_i - \sum_{i<j} n_{ij}\log(\pi_i+\pi_j) \tag{2.5}$$

where $x_{i+} = \sum_{j=1}^{t} x_{ij}$ (adopting the convention that $x_{ii} = 0$).

Note that the minimal sufficient statistics for this model are the $\{x_{i+}\}$.

Suppose we denote the expected number of times T_i is preferred to T_j by m_{ij}, i.e.

$$m_{ij} = E(x_{ij}) = n_{ij}\pi_i/(\pi_i+\pi_j). \tag{2.6}$$

Then general results of Haberman (1974) can be used to show that the maximum likelihood estimates of the $\{m_{ij}\}$ found by maximizing expression (2.4) are the same as those resulting from maximizing the kernel of the multinomial loglikelihood function

$$\sum_{i\neq j} x_{ij} m_{ij} = \sum_i x_{i+} \log e_i + \sum_{i<j} n_{ij} \log c_{ij} \qquad (2.7)$$

where

$$m_{ij} = e_i c_{ij},$$

$$(2.8)$$

and $c_{ij} = c_{ji}$ for $i\neq j$.

The parameters $\{e_i\}$ are simply scaled versions of the $\{\pi_i\}$. The natural paired comparisons sampling constraints given by expression (2.3) are automatically satisfied in the maximization of expression (2.7). Note that the problem of maximizing the original loglikelihood function has been transformed into one of maximizing a loglikelihood that is linear in the parameters of interest.

The model given in (2.8) is simply the loglinear model of quasi-symmetry for a t x t contingency table with entries $\{x_{ij}\}$. The likelihood equations for this model are

$$\hat{m}_{i+} = x_{i+}, \quad i = 1,2,\ldots t,$$

$$(2.9)$$

$$\hat{m}_{ij} + \hat{m}_{ji} = x_{ij} + x_{ji} = n_{ij}, \qquad i \neq j,$$

and these equations can be solved using a version of the general iterative scaling technique described by Darroch and Ratcliff (1972). We take as initial values $\hat{m}_{ij}^{(0)}=1$ for $i\neq j$, and $\hat{m}_{ii}^{(0)}=0$. The iterative method consists of successively adjusting the estimated expected values so that the two sets of equations in (2.9) are satisfied, i.e. for $\upsilon=0,1,...$ we let

$$\hat{m}_{ij}^{(2\upsilon+1)} = \hat{m}_{ij}^{(2\upsilon)} x_{i+} / \hat{m}_{i+}^{(2\upsilon)},$$

$$\hat{m}_{ij}^{(2\upsilon+2)} = \hat{m}_{ij}^{(2\upsilon+1)} n_{ij} / (\hat{m}_{ij}^{(2\upsilon+1)} + \hat{m}_{ji}^{(2\upsilon+1)}). \tag{2.10}$$

Such an iterative procedure is equivalent to the one used by Bradley and Terry. This procedure converges to the unique maximum likelihood estimates provided that the conditions described at the end of this section are satisfied. The likelihood ratio goodness-of-fit test statistic for the model of quasi-symmetry in the t x t table is identical to the likelihood ratio statistics used by Bradley and Terry.

Bradley (1953), David (1963), Cox (1970), and Maxwell (1974) all note that the Bradley-Terry model corresponds to a linear preference model in the logit scale, i.e. if

$$\theta_{ij} = \log\{\Pr(T_i \succ T_j)/\Pr(T_j \succ T_i)\}, \tag{2.11}$$

then we have that

$$\theta_{ij} + \theta_{jk} = \theta_{ik}, \tag{2.12}$$

for all i,j and k. Substituting for the θ's in (2.12) using expressions (2.11), (2.6) and (2.1) we get that

$$\frac{m_{ij}m_{jk}m_{ki}}{m_{ji}m_{kj}m_{ik}} = 1 \qquad \forall\ i,j,k. \tag{2.13}$$

Expression (2.13) gives the constraint equations typically used to describe the model of quasi-symmetry in a square table (see Bishop, Fienberg, and Holland, 1975).

To illustrate a second loglinear representation for the Bradley-Terry model let us look at the case of $t = 4$. The quasi-symmetry formulation of the model suggests laying out the data in a 4 x 4 table of counts as in Table 1. The dashes down the diagonal represent structural zeros.

Table 1 goes about here.

An alternative to this layout is suggested by the binomial nature of the sampling scheme used for the comparisons. When $t = 4$ we have $\binom{4}{2} = 6$ binomial samples comparing 2 of the 4 categories, so we can lay the data out in a 4 x 6 table of counts as in Table 2. Here we have 12 structural zeros. In general the format would be that of a $t \times \binom{t}{2}$ table. Note that the row totals in Table 2 are the sufficient statistics $\{x_{i+}\}$, and the column totals are the sample sizes $\{n_{ij}\}$. The Bradley-Terry model corresponds to the model of quasi-independence or quasi-homogeneity of proportions in the $t \times \binom{t}{2}$ table of expected counts, $\{m_{ij}\}$, laid out as in Table 2.

Table 2 goes about here.

Savage (1973) gives necessary and sufficient conditions for the existence of unique positive maximum likelihood estimates for the expected values in a rectangular array under the model of quasi-independence. In the special case under consideration here Savage's conditions reduce to the conditions given earlier by Ford (1957):

> "In every possible partition of the objects into two non-empty subsets, some object in the second set has been preferred at least once to some object in the first set."

In particular, these conditions require that $x_{i+} > 0$ for all i.

Koch, Abernathy, and Imrey (1975) have used the quasi-independence model applied to the $t \times \binom{t}{2}$ table layout for testing the suitability of the Bradley-Terry model. This particular representation allows for simple generalizations to the multiple comparisons problem as we illustrate in the next section.

3. MULTIPLE COMPARISONS MODELS

There are two extreme versions of the multiple comparisons problem which we briefly reexamine here. In both of these the judge is given a subset M of the t treatments. In the first version the judge is asked to choose the one treatment he prefers (the preference choice problem), while in the second he is asked to rank all of the treatments in the subset M.

Suppose M is any subset of the t treatments which contains T_i. Then for the preference choice problem Luce (1959) proposed the following model:

$$\Pr\{T_i \text{ preferred} | M\} = \pi_i / \sum_M \pi_j, \qquad (3.1)$$

where $\pi_i \geq 0$ for $i = 1,2,\ldots,t$, and $\pi_1+\ldots+\pi_t = 1$ as in Section 2. Now if s different subsets of the treatments are selected for presentation to sets of judges, we can array the resulting data in the form of a $t \times s$ table, where for the j th column we get observed counts in the rows corresponding to the treatments in the subset under examination. In the paired comparisons problem $s = \binom{t}{2}$ and each subset consists of exactly two treatments.

By pursuing an argument similar to that of the previous section we can see that the Luce model in expression (3.1) is in fact the model of quasi-independence or quasi-homogeneity of proportions in the $t \times s$ table. Indeed the standard interpretation of the quasi-independence model is the "independence from irrelevant alternatives," i.e. if the subgroup M contains T_i and T_j, the ratio of the preference probabilities of the two treatments is π_i/π_j, no matter what other objects are in M. The Luce model clearly implies the "logit additivity" property of expression (2.11), both for triples of treatments within any one subgroup or pairs within different groups. The conditions of Savage (1973) can be used here to ensure the existence of unique positive maximum likelihood estimates for the expected values under the Luce model.

Vidmar (1972) conducted an experiment to examine restricted decision choices for verdicts by jurors in a murder trial situation. There are four possible decision alternatives corresponding to "treatments": not guilty, guilty of first degree murder, guilty of second degree murder, guilty of manslaughter. Seven different groups of 24 subjects each were presented with various subsets of the 4 verdicts and were required to make a choice based on the evidence for a particular trial. All seven subsets include "not guilty." Three involve paired comparisons, three involve triple comparisons, and the last includes all four alternatives. Table 3a gives the observed data for this experiment, and Table 3b the expected values under the Luce model, (3.1), as presented by Larntz (1975). The expected values were computed using the standard iterative scaling algorithm for fitting the quasi-independence model to a t x s table (see Bishop, Fienberg, and Holland, 1975). The fit between observed and expected values appears to be reasonably good. The likelihood ratio statistic for testing the goodness-of-fit of the model has a value of G^2 = 18.8 with 9 d.f., corresponding to a descriptive level of significance between .10 and .05 when referred to the χ^2_9 distribution. The estimated values of the parameters in the Luce model are: $\hat{\pi}_F = .071$, $\hat{\pi}_S = .532$, $\hat{\pi}_M = .347$, and $\hat{\pi}_N = .050$.

Table 3 goes about here.

The second extreme version of the multiple comparisons model involves the complete rankings of the treatments. In the case where each comparison involves three treatments Pendergrass and Bradley (1960) proposed the model:

$$\Pr(T_i \succ T_j \succ T_k) = \frac{\pi_i^2\pi_j}{\pi_i^2(\pi_j+\pi_k)+\pi_j^2(\pi_i+\pi_k)+\pi_k^2(\pi_i+\pi_j)} = \frac{\pi_i^2\pi_j}{\Delta_{ijk}}, \quad (3.2)$$

where as usual $\pi_i \geq 0$ for all i and $\pi_1+\ldots+\pi_t = 1$. The numerator of (3.2) is the product of the numerators of the three pairwise Bradley-Terry preference probabilities implied by the ranking.

If the triplet (T_i,T_j,T_k) is ranked by n_{ijk} judges, let x_{ijk} be the observed number of times the preference rankings is given by $T_i \succ T_j \succ T_k$. Suppose now that all possible triplets of treatments are so ranked. For convenience we define $x_{iji}=x_{ikk}=x_{iii}=0$ for all i,j,k. Then the kernel of the loglikelihood under model (3.2) is given by

$$\sum_{i=1}^{t} a_i \log\pi_i - \sum_{i<j<k} n_{ijk} \log \Delta_{ijk}, \quad (3.3)$$

where

$$a_i = 2 \sum_{j,k} x_{ijk} + \sum_{j,k} x_{jik} = 2x_{i++} + x_{+i+}. \quad (3.4)$$

As in Section 2, loglinear model theory tells us that the maximum likelihood estimates of the expected counts, $m_{ijk} = E(x_{ijk})$, obtained by maximizing expression (3.3), are equal to those obtained by maximizing the multinomial loglikelihood

$$\sum_{i,j,k} x_{ijk} \log m_{ijk}, \quad (3.5)$$

where

$$m_{ijk} = e_i^2 e_j c_{ijk}, \quad (3.6)$$

and c_{ijk} is invariant under all possible permutations of subscripts. The minimal sufficient statistics for this multinomial problem turn out to be the $\{a_i\}$ and $\{x_{ijk}+x_{jik}+x_{ikj}+x_{jki}+x_{kij}+x_{kji}\}$, and the likelihood equations are given by setting the minimal sufficient statistics equal to their expectations. The solution of these likelihood equations can be found using a version of the generalized iterative scaling procedure of Darroch and Ratcliff (1972), which is similar to that suggested by Pendergrass and Bradley (1960).

As in the paired comparison situation the fit of the model can be checked using the usual chi-square statistics based on observed and expected values. If the model fits, various tests regarding the structure of the $\{\pi_i\}$ can be carried out accordingly. Extensions of these techniques to ranking problems involving m-comparisons, where $m > 3$, are immediate.

4. MULTIVARIATE PAIRED COMPARISON MODELS

Suppose now that t items are to be compared on p attributes, $A_1, A_2, \ldots, A_p$, instead of just a single characteristic. An extension of the Bradley-Terry model to this situation has been given by Davidson and Bradley (1969). Their basic idea is to fit a Bradley-Terry model for each attribute with additional parameters to allow for the possible association of the comparisons involving items T_i and T_j on attributes A_α and A_β. In the case of two attributes the measure of association incorporated into their model is the

ϕ-coefficient of correlation for the single 2 x 2 table. The loglinear model we propose also allows for Bradley-Terry parameters for each attribute, but the corresponding measure of association is the cross-product (or log-odds) ratio (Bishop, Fienberg, and Holland, 1975; Cox, 1970; Haberman, 1974).

As in Section 2 we have T_i compared with T_j $n_{ij} > 0$ times for $i < j = 1,2,\ldots t$. Let $\underset{\sim}{s} = (s_1,\ldots,s_p)$ be a preference vector such that, when treatments T_i and T_j $(i<j)$ are compared, $s_\alpha = i$ if $T_i \succ T_j$ on attribute α and $s_\alpha = j$ if $T_j \succ T_i$ on attribute α. Let $x_{ij\cdot\underset{\sim}{s}}$ be the number of times the preference vector $\underset{\sim}{s}$ is observed in the n_{ij} comparisons of T_i and T_j. Then

$$n_{ij} = \sum_{\underset{\sim}{s}} x_{ij\cdot\underset{\sim}{s}},$$

where the summation is over the 2^p possible preference vectors $\underset{\sim}{s}$.

The model for multivariate paired comparisons that we propose postulates that the probability of observing preference vector $\underset{\sim}{s}$ when comparing T_i and T_j $(i<j)$ on p attributes is

$$\Pr(\underset{\sim}{s}|(i,j)) = C_{ij}\gamma(\underset{\sim}{s})\left[\prod_{\alpha=1}^{p} \frac{\pi_{\alpha s_\alpha}}{\pi_{\alpha i}+\pi_{\alpha j}}\right] \qquad (4.1)$$

where $\pi_{\alpha 1}+\cdots+\pi_{\alpha t} = 1$ as before, C_{ij} is a normalizing constant,

$$\gamma(\underset{\sim}{s}) = \prod_{\alpha<\beta} \exp[\delta(s_\alpha, s_\beta)\gamma_{\alpha\beta}] \ , \tag{4.2}$$

$\delta(s_\alpha, s_\beta) = +1$ if $s_\alpha = s_\beta$, and $\delta(s_\alpha, s_\beta) = -1$ otherwise. The parameter $\gamma_{\alpha\beta}$ corresponds to the log of the ratio of similar to dissimilar preferences on attributes α and β. The kernel of the loglikelihood under model (4.1) is

$$\sum_{\alpha=1}^{p} \sum_{i=1}^{t} v_{\alpha i} \log\pi_{\alpha i} - \sum_{\alpha=1}^{p} \sum_{i<j} n_{ij} \log(\pi_{\alpha i} + \pi_{\alpha j}) + \sum_{\alpha<\beta} w_{\alpha\beta}\gamma_{\alpha\beta}, \tag{4.3}$$

where, if we set $x_{ii\cdot \underset{\sim}{s}} = 0$,

$$v_{\alpha i} = \sum_{j=1}^{t} \sum_{\{\underset{\sim}{s}: s_\alpha = i\}} x_{ij\cdot \underset{\sim}{s}} \ , \tag{4.4}$$

and

$$w_{\alpha\beta} = \sum_{i<j} \sum_{\underset{\sim}{s}} \delta(s_\alpha, s_\beta) x_{ij\cdot \underset{\sim}{s}} . \tag{4.5}$$

Note that $v_{\alpha i}$ is simply the total number of times T_i is preferred to any other treatment on attribute α, and $w_{\alpha\beta}$ is the difference between the number of similar $(s_\alpha = s_\beta)$ and dissimilar $(s_\alpha \neq s_\beta)$ preferences on the pair of attributes (α, β). The minimal sufficient statistics for this model are the $\{v_{\alpha i}\}$ and the $\{w_{\alpha\beta}\}$. As before, the maximum likelihood estimates of the expected cell counts, $m_{ij\cdot \underset{\sim}{s}} = E(x_{ij\cdot \underset{\sim}{s}})$, obtained by maximizing (4.3), are equal to those obtained by maximizing the multinomial likelihood

$$\sum_{i<j} \sum_{\underset{\sim}{s}} x_{ij\cdot \underset{\sim}{s}} \log m_{ij\cdot \underset{\sim}{s}} \ , \tag{4.6}$$

where $m_{ij\cdot\underset{\sim}{s}}$ is of the form

$$m_{ij\cdot\underset{\sim}{s}} = h_{ij}\left[\prod_{\alpha=1}^{p} e_{\alpha s_\alpha}\right]\left[\prod_{\alpha<\beta} g_{\alpha\beta}^{\delta(s_\alpha,s_\beta)}\right] . \qquad (4.7)$$

The likelihood equations are constructed by setting the minimal sufficient statistics equal to their expectations :

$$\text{(i)} \quad \sum_{\underset{\sim}{s}} m_{ij\cdot\underset{\sim}{s}} = n_{ij} , \qquad i < j=1,\ldots,t ,$$

$$\text{(ii)} \quad \sum_{j=1}^{t} \sum_{\{\underset{\sim}{s}:s_\alpha=i\}} m_{ij\cdot\underset{\sim}{s}} = v_{\alpha i} , \quad \alpha=1,\ldots,p;\ i=1,\ldots,t , \qquad (4.8)$$

$$\text{(iii)} \quad \sum_{i<j} \sum_{\underset{\sim}{s}} \delta(s_\alpha,s_\beta) m_{ij\cdot\underset{\sim}{s}} = w_{\alpha\beta} , \quad \alpha<\beta=1,\ldots,p .$$

Again the generalized iterative scaling procedure of Darroch and Ratcliff (1972) can be used to find the solution of these equations. The general results of Haberman (1974) can be used here to determine necessary and sufficient conditions for the existence of unique positive maximum likelihood estimates.

An attractive feature of model (4.1) and the approach just described is the ease with which they can be generalized to handle higher order dependencies among the response attributes. These generalizations are suggested by analogy with models for multi-dimensional contingency tables.

In the special case of equal treatment preferences on both of two attributes, the factor in square brackets in expression (4.1) becomes $\frac{1}{4}$, and our model reduces to the null model examined by

Sen and David (1968). To test the adequacy of this null model they use the statistic (in our notation)

$$D_N = t^{-1} \sum_{i=1}^{t} \left[\frac{2N}{N+w_{12}} (v_{1i}+v_{2i}-n_{i+})^2 + \frac{2N}{N-w_{12}} (v_{1i}-v_{2i})^2 \right] \quad (4.9)$$

where n_{i+} is the number of comparisons involving item i. The statistic D_N is asymptotically distributed as $\chi^2_{2(t-1)}$, and is an alternative to one of the statistics discussed below.

We consider now the two examples of Davidson and Bradley (1969). Their vanilla pudding data is given in Table 4. Five times were compared on two attributes (1) taste and (2) appearance. In this example constraint (iii) of the expression (4.8) implies that the expected number of similar and dissimilar preferences on attributes A_1 and A_2 must equal the observed number of similar and dissimilar preferences. The degrees of freedom for testing the goodness-of-fit of the model are the number of cells minus the number of linearly independent constraints imposed by (4.8), i.e. 40-10-8-1=21. The expected cell frequencies are given in parentheses in Table 4. The likelihood ratio chi-square statistic has a value of $G^2 = 23.6$, a value indicating no significant lack-of-fit when compared to the χ^2_{21} distribution.

Table 4 goes about here

The loglinear model approach can be used (assuming an appropriate model has been found) to test various hypotheses about the model parameters. For example, to test if

$\gamma_{12} = 0$ (i.e., independent Bradley-Terry models for each attribute) the corresponding reduced loglinear model may be formed. In this case the cell expectations can be found by using the constraints (4.8) but omitting constraint (iii). When this is done the likelihood-ratio statistic can then be calculated and the difference between this G^2 and the previous goodness-of-fit G^2 will give a one degree-of-freedom test of $\gamma_{12} = 0$. In this example $\Delta G^2 = 34.0$ and thus it seems unreasonable to assume $\gamma_{12} = 0$. Similarly, tests for equal preferences on attribute A_1, or on A_2, or on both attributes simultaneously (the Sen-David situation) can be constructed. Table 5 summarizes some of the possible tests for the vanilla pudding data. The basic conclusions here are essentially the same as those reached by Davidson and Bradley (1969): there is little difference among the puddings in terms of preference but responses on the two attributes are related. Our model provides a slightly better fit than the Davidson-Bradley one.

Table 5 goes about here

Table 6 gives a second data set from Davidson and Bradley. In this case three brands of chocolate pudding were compared in paired comparisons on three attributes, (1) taste, (2) color, and (3) texture. There are 24-3-6-3=12 degrees of freedom here for model (4.1), and the likelihood ratio goodness-of-fit statistic takes the value $G^2 = 8.4$. Again we have a good fit when G^2 is compared to the χ^2_{12} distribution. Table 7 gives several tests relating to the model parameters. In all cases the fitting of the reduced loglinear model was accomplished by removing the appropriate constraints from the set (4.8). Taking all of these tests into account we reach the same conclusion as in the vanilla pudding example. There is little or no difference in the preferences for the three puddings but the attributes are slightly related to one another. As before, our model provides only a slight improvement in fit over the Davidson-Bradley model.

The first author's research was partially supported by grants from the Robert Wood Johnson Foundation and the Commonwealth Fund to the Center for the Analysis of Health Practices and from the National Science Foundation to the Department of Statistics, Harvard University.

Tables 6 and 7 go about here

REFERENCES

Andersen, A.H. (1974). Multidimensional contingency tables. Scand. J. Statist. 1, 115-27.

Bishop, Y.M., Fienberg, S.E., and Holland, P.W. (1975). Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: M.I.T. Press.

Bradley, R.A. (1953). Some statistical methods in taste testing and quality evaluation. Biometrics 9, 22-38.

Bradley, R. A. (1976). Science, statistics and paired comparisons. Biometrics 32 (in press).

Bradley, R.A. and Terry, M.E.(1952). The rank analysis of incomplete block designs. I. The method of paired comparisons. Biometrika 39, 324-45.

Caussinus, H. (1966). Contribution à l'analyse statistique des tableaux de corrélation. Ann. Fac. Sci. Univ. Toulouse 29, 77-182.

Cox, D.R. (1970). The Analysis of Binary Data. London: Methuen.

Darroch, J.N. and Ratcliff, D. (1972). Generalized iterative scaling for loglinear models. Ann. Math. Statist. 43, 1470-80.

David, H.A. (1962). The Method of Paired Comparisons. London: Griffin.

Davidson, R.R. and Bradley, R.A. (1969). Multivariate paired comparisons: The extension of a univariate model and associated estimation and test procedures. Biometrika 56, 81-95.

Ford, L.R. Jr. (1957). Solution of a ranking problem from binary comparisons. Ann. Math. Monthly 64, 28-33.

Goodman, L.A. (1968). The analysis of cross-classified data: independence, quasi-independence, and interaction in contingency tables with or without mi-sing cells. J. Amer. Statist. Assoc. 63, 1091-131.

Haberman, S.J. (1974). The Analysis of Frequency Data. Chicago: Univ. of Chicago Press.

Koch, G.G., Abernathy, J.R., and Imrey, P.B. (1975). On a method for studying family size preferences. Demography (in press).

Larntz, K. (1975). Reanalysis of Vidmar's data on the effects of decision alternatives on verdicts of simulated jurors. J. Personality and Social Psychol. 31, 123-5.

Luce, R.D. (1959). Individual Choice Behavior. New York: Wiley.

Maxwell, A.E. (1974). The logistic transformation in the analysis of paired-comparison data. Br. J. Math. Statist. Psychol. 27, 62-71.

Mosteller, F. (1951). Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. Psychometrika 16, 3-9.

Pendergrass, R.N. and Bradley, R.A. (1960). Rankings in triple comparisons. Contributions to Probability and Statistics, ed. I. Olkin et al., Stanford: Stanford University Press, 331-51.

Savage, I.R. (1973). Incomplete contingency tables: condition for the existence of unique MLE. Mathematics and Statistics. Essays in Honor of Harold Bergström, ed. P. Jagars and L. Råde, Göteborg, Sweden: Chalmers Institute of Technology, 87-99.

Sen, P.K. and David, H.A. (1968). Paired comparisons for paired characteristics. Ann. Math. Statist. 39, 200-8.

Vidmar, N. (1972). Effects of decision alternatives on the verdicts and social perceptions of simulated jurors. J. Personality and Social Psychol. 22, 211-8.

Table 1

Typical Layout for Data in Paired-Comparisons with t=4.

		Against			
		T_1	T_2	T_3	T_4
For	T_1	--	x_{12}	x_{13}	x_{14}
	T_2	x_{21}	--	x_{23}	x_{24}
	T_3	x_{31}	x_{32}	--	x_{34}
	T_4	x_{41}	x_{42}	x_{43}	--

Table 2

Alternative Layout for Data in Paired Comparisons with t=4.

Paired Comparison Number

		1	2	3	4	5	6
	T_1	x_{12}	x_{13}	x_{14}	--	--	--
Preferred	T_2	x_{21}	--	--	x_{23}	x_{24}	--
Treatment	T_3	--	x_{31}	--	x_{32}	--	x_{34}
	T_4	--	--	x_{41}	--	x_{42}	x_{43}

Table 3

Decision Alternatives for Murder Trial
(Vidmar, 1972; Larntz, 1975)

(a) Observed values.

Alternative	Condition						
	1	2	3	4	5	6	7
First degree	11	--	--	2	7	--	2
Second degree	--	20	--	22	--	11	15
Manslaughter	--	--	22	--	16	13	5
Not guilty	13	4	2	0	1	0	2

(b) Expected values under Luce model, (3.1).

Alternative	Condition						
	1	2	3	4	5	6	7
First degree	14.0	--	--	2.6	3.6	--	1.7
Second degree	--	21.9	--	19.5	--	13.8	12.8
Manslaughter	--	--	21.0	--	17.8	9.0	8.3
Not guilty	10.0	2.1	3.0	1.8	2.6	1.3	1.2

Table 4

Observed and Expected Values for
Vanilla Pudding Data (Davidson and Bradley [1969])

Treatment Pair (i,j)	Preference Configuration (i,i)	(i,j)	(j,i)	(j,j)
(1,2)	1 (3.80)	1 (1.22)	3 (1.20)	5 (3.78)
(1,3)	8 (6.14)	1 (1.55)	1 (1.81)	4 (4.50)
(1,4)	7 (6.87)	4 (1.64)	2 (1.65)	1 (3.85)
(1,5)	6 (5.61)	1 (2.17)	1 (1.92)	9 (7.29)
(2,3)	2 (3.50)	0 (0.88)	1 (1.04)	5 (2.57)
(2,4)	5 (6.37)	2 (1.51)	3 (1.54)	3 (3.58)
(2,5)	7 (5.60)	1 (2.16)	1 (1.94)	8 (7.30)
(3,4)	3 (4.73)	1 (1.42)	2 (1.23)	5 (3.62)
(3,5)	4 (3.31)	2 (1.61)	1 (1.23)	5 (5.86)
(4,5)	2 (2.33)	1 (1.21)	1 (1.06)	6 (5.40)

Table 5

Goodness-of-fit Statistics of Models
for Vanilla Pudding Data

Model	G^2	d.f.	ΔG^2	df
Expression (4.1)	23.6	21	--	--
$\gamma_{12} = 0$	57.7	22	34.1	1
$\pi_{1j}=1/5, i=1,\ldots,5$	24.4	25	0.8	4
$\pi_{2i}=1/5, i=1,\ldots,5$	25.3	25	1.7	4
$\pi_{\alpha i}=1/5, i=1,\ldots,5$, $\alpha=1,2$	28.4	29	4.8	8

Table 6

Observed and Expected Values for
Chocolate Pudding Data
(Davidson and Bradley [1969])

Treatment Pair (i,j)	Preference Configuration (i,i,i)	(i,i,j)	(i,j,i)	(i,j,j)	(j,i,i)	(j,i,j)	(j,j,i)	(j,j,j)
(1,2)	8 (6.88)	0 (0.69)	1 (0.99)	0 (0.41)	1 (0.92)	2 (1.12)	1 (1.82)	9 (9.17)
(1,3)	6 (6.10)	1 (1.11)	1 (1.05)	1 (0.78)	0 (0.37)	0 (0.82)	1 (0.87)	9 (7.91)
(2,3)	7 (7.58)	3 (1.94)	1 (1.20)	1 (1.26)	1 (0.27)	1 (0.83)	1 (0.57)	6 (7.37)

Table 7

Goodness-of-fit Statistics of Models
for Chocolate Pudding Data

Model	G^2	df	ΔG^2	df
Expression (4.1)	8.4	12	--	--
$\gamma_{\alpha\beta}=0, \alpha,\beta=1,2,3.$	71.8	15	63.4	3
$\pi_{1i}=1/3, i=1,2,3.$	10.5	14	2.1	2
$\pi_{2i}=1/3, i=1,2,3.$	8.6	14	0.2	2
$\pi_{3i}=1/3, i=1,2,3.$	9.8	14	1.4	2
$\pi_{\alpha i}=1/3, i=1,2,3. \alpha=1,2,3.$	11.5	8	3.1	6