The Annals of Statistics
1991, Vol. 19, No. 4, 2090–2101

THE ORDER-RESTRICTED RC MODEL FOR ORDERED CONTINGENCY TABLES: ESTIMATION AND TESTING FOR FIT

By Ya'acov Ritov and Zvi Gilula

Hebrew University of Jerusalem

The RC model has been proposed as a model for ordered contingency tables. It involves parametric scores that are assigned to the rows and columns of the table so that these scores reflect the ordinality of the row and column categories. Efficient estimation of these parameters subject to order constraints remained an open problem mainly due to severe difficulties in computing these estimates and difficulties in deriving an appropriate asymptotic goodness-of-fit test. A nonstandard yet very simple algorithm is derived which produces the desired order-restricted maximum likelihood estimates with probability converging to 1. Testing the order-restricted RC model for fit is also discussed.

1. Introduction. Goodman (1979, 1981, 1985, 1986) introduced and discussed a family of association models suitable for ordered two-way contingency tables. Some of these models—the column-effect model (the C model), the row-effect model (the R model) and the row–column-effect model (the RC model)—involve parametric scores assigned to the rows and the columns of the table. These scores should reflect the order of the underlying categories. In the above-mentioned references an unrestricted maximum likelihood estimation procedure for the scores is presented together with a testing procedure for testing equality of scores (to be used when score estimates "violate" a presumed order). Agresti, Chuang and Kezouh (1987) justifiably claim that in many cases the order for the parametric scores should be predetermined since a researcher frequently has in mind—before observing any data—some notion about the *direction* of the association. A positive direction should be expressed by row and column scores that are both monotone increasing or monotone decreasing. A negative direction should be expressed by monotone scores that vary in opposite direction in the rows and in the columns. Agresti, Chuang and Kezouh (1987) developed a maximum likelihood estimation and a testing procedure for the ordered C model and for the ordered R model. As for the ordered RC model, they [and also Goodman (1985)] mentioned some estimation difficulties mainly due to the fact that the RC model is *not* a log-linear model. Consequently, the estimation problem under the ordered RC model was left open. Agresti, Chuang and Kezouh (1987) recognized that "the discovery of sufficient conditions for the solution of the ordinary or order-restricted row

Received May 1988; revised January 1991.
AMS 1980 *subject classifications*. 62H17, 62H20.
Key words and phrases. Collapsed tables, maximum likelihood, order restriction, parametric scores.

and column effects model, plus the development of an algorithm for fitting these models, are important problems... ."

In this paper solutions to these problems are provided. We derive estimates for the parametric scores of the RC model subject to order restrictions. These estimates are shown, by some lemmas proven below, to be the maximum likelihood estimates for the ordered RC model with probability converging to 1. The problem of estimating the parameters of the ordered RC model requires a nonstandard solution because the underlying parameters functionally depend upon each other. It is argued that the common method of amalgamating (unrestricted) estimates does not always produce the right estimates. To solve this problem, particular functions of the parameters are defined. Amalgamation of these functions is shown to produce the desired maximum likelihood estimates with probability converging to 1, and it is easy to tell at each stage of this procedure whether the solutions to the likelihood equations have been obtained or not. The standard maximum likelihood estimation procedure for our problem will usually require a long and tedious series of collapsed tables obtained from the original table even for tables with moderate number of cells. Our procedure, apart from being rapidly convergent, is quite parsimonious and only a single collapsing operation will usually be needed to obtain the optimal estimates. We also discuss the issue of testing the order-restricted RC model for fit.

2. Preliminaries and notation. Let X and Y be two categorical random variables with ranges consisting of the integers 1 to I and 1 to J, respectively. Let

$$\begin{aligned} P_{ij} &= \Pr(X=i, Y=j), && 1 \le i \le I, 1 \le j \le J,\\ P_{i\cdot} &= \Pr(X=i), && 1 \le i \le I,\\ P_{\cdot j} &= \Pr(Y=j), && 1 \le j \le J. \end{aligned}$$

The RC model states

$$P_{ij} = \alpha_i \beta_j \exp(\phi \mu_i \nu_j), \qquad 1 \le i \le I, 1 \le j \le J, \tag{2.1}$$

where

$$\sum_i P_{i\cdot}\mu_i^b = \sum_j P_{\cdot j}\nu_j^b = b-1, \qquad b = 1,2. \tag{2.2}$$

Condition (2.2) was suggested by Goodman (1981) for purposes of identifiability and comparison with canonical correlation models.

For the order-restricted case we assume in addition to (2.1) and (2.2) the following:

$$\begin{gathered} \mu_1 \le \mu_2 \le \cdots \le \mu_I, \qquad \mu_1 < \mu_I,\\ \nu_1 \le \nu_2 \le \cdots \le \nu_J, \qquad \nu_1 < \nu_J,\\ \phi > 0,\\ P_{ij} > 0 \quad \text{for all } i \text{ and } j. \end{gathered} \tag{2.3}$$

Although we consider here only the positive association case, the results below are easily extendable to the negative association case as well.

We use the following notation for all i and j: $\tilde{P}_{ij}$ is the empirical distribution (given in an observed contingency table with sample size n); $\hat{P}_{ij}$ is the unrestricted maximum likelihood estimates for P_{ij} in (2.1); $\hat{\mu} = (\hat{\mu}_1, \dots, \hat{\mu}_I)$ is the vector of unrestricted maximum likelihood estimates for row scores; $\hat{\nu} = (\hat{\nu}_1, \dots, \hat{\nu}_J)$ is the vector of unrestricted maximum likelihood estimates for column scores; μ^* and ν^* are the vectors of order-restricted maximum likelihood estimates for row and column scores, respectively. Let

$$
\begin{aligned}
E_i(\nu) &= \sum_j P_{ij}\nu_j / P_{i\cdot}, \\
F_j(\mu) &= \sum_i P_{ij}\mu_i / P_{\cdot j}, \qquad (2.4)\\
\tilde{E}_i(\nu) &= \sum_j \tilde{P}_{ij}\nu_j / \tilde{P}_{i\cdot}
\end{aligned}
$$

and

$$\tilde{F}_j(\mu) = \sum_i \tilde{P}_{ij}\mu_i / \tilde{P}_{\cdot j}.$$

Using the above notation, it is easily verified [see also Goodman (1981)] that the solution of the likelihood equations for the unrestricted RC model satisfies

$$
\begin{aligned}
\hat{P}_{i\cdot} &= \tilde{P}_{i\cdot}, && 1 \le i \le I, \\
\hat{P}_{\cdot j} &= \tilde{P}_{\cdot j}, && 1 \le j \le J, \\
\sum_{i,j} \hat{P}_{i,j}\hat{\mu}_i\hat{\nu}_j &= \sum_{i,j} \tilde{P}_{ij}\hat{\mu}_i\hat{\nu}_j, && \qquad (2.5)\\
\sum_j \hat{P}_{ij}\hat{\nu}_j &= \sum_j \tilde{P}_{ij}\hat{\nu}_j, && 1 \le i \le I, \\
\sum_i \hat{P}_{ij}\hat{\mu}_i &= \sum_i \tilde{P}_{ij}\hat{\mu}_i, && 1 \le j \le J.
\end{aligned}
$$

These equations will be used later.

3. Main results. A well-known idea in maximizing the likelihood under order restrictions is to first apply the unrestricted maximum likelihood procedure [e.g., Goodman (1981) for our context]. If the resulting estimates follow (2.3), then we are done. Otherwise some inequalities between adjacent μ's and some inequalities between adjacent ν's must be turned into equalities and the unrestricted estimation procedure then reapplied. The problem is the lack of a parsimonious search strategy that will tell how many equalities should be imposed and where. The common practice is to base the search on the actual estimates. Here, for example, if a violation of the order occurs with $\hat{\mu}_e > \hat{\mu}_{e+1}$ and with $\hat{\mu}_k > \hat{\mu}_{k+1}$, $e \neq k$, one would be tempted to impose $\mu_k = \mu_{k+1}$ and

$\mu_e = \mu_{e+1}$. This algorithm, known also as "pooling adjacent violators" [e.g., Barlow, Bartholomew, Bremmer and Brunk (1972), page 13], is quite common in ordered inference under normal theory but is problematic here. The main problem lies in the fact that (unlike the common case in standard theory) the μ's and ν's are functions of each other as is evident from the (unrestricted) likelihood-solution equations (2.5) for model (2.1). This means that imposing an equality between a pair of adjacent μ's in order to correct for order violation may result in a (new) violation of order between some ν's and between some μ's that followed the desired order prior to the amalgamation process. This problem leads to severe doubt as to whether pooling adjacent violators on the score parameters is useful [see also Neusch (1966) for similar arguments]. As is proven next, all the above-mentioned problems can be overcome if amalgamation is done on the quantities $\tilde{E}_j(\nu)$ and $\tilde{F}_j(\mu)$ defined in (2.4) instead of the μ's and the ν's directly.

For expository purposes in the next lemma, we abbreviate some notation in (2.4) as follows:

$$E_i \equiv E_i(\nu), \qquad \tilde{E}_i \equiv \tilde{E}_i(\nu), \qquad F_j \equiv F_j(\mu), \qquad \tilde{F}_j \equiv \tilde{F}_j(\mu).$$

LEMMA 1. *The random variables* $\sqrt{n}(\tilde{E}_i - E_i)$ *and* $\sqrt{n}(\tilde{F}_j - F_j)$, $1 \le i \le I$, $1 \le j \le J$, *are asymptotically jointly normal with zero mean and covariance matrix having entries*

$$n\,\mathrm{Cov}\left[\tilde{E}_i, \tilde{E}_k\right] = \begin{cases} \mathrm{Var}(\nu|i)/P_{i\cdot}, & i = k, \\ 0, & i \neq k, \end{cases}$$

$$n\,\mathrm{Cov}\left[\tilde{F}_j, \tilde{F}_m\right] = \begin{cases} \mathrm{Var}(\mu|j)/P_{\cdot j}, & j = m, \\ 0, & j \neq m, \end{cases}$$

$$n\,\mathrm{Cov}\left[\tilde{F}_j, \tilde{E}_i\right] = (P_{ij}/P_{i\cdot}P_{\cdot j})(\nu_j - E_i)(\nu_i - F_j),$$

where $\mathrm{Var}(\nu|i) = \sum_j P_{ij}(\nu_j - E_i)^2/P_{i\cdot}$ *and* $\mathrm{Var}(\mu|j) = \sum_i P_{ij}(\mu_i - F_j)^2/P_{\cdot j}$.

PROOF. Using straightforward algebraic manipulations and a Taylor expansion yields

$$\begin{aligned}
\tilde{E}_i &= \sum_j \tilde{P}_{ij}\nu_j/\tilde{P}_{i\cdot} = \sum_j \tilde{P}_{ij}\nu_j/P_{i\cdot} + \sum_j \tilde{P}_{ij}\nu_j\left(1/\tilde{P}_{i\cdot} - 1/P_{i\cdot}\right) \\
&= E_i + \sum_j \left(\tilde{P}_{ij} - P_{ij}\right)\nu_j/P_{i\cdot} - \left(\sum_j \tilde{P}_{ij}\nu_j/\tilde{P}_{i\cdot}P_{i\cdot}\right)\sum_j \left(\tilde{P}_{ij} - P_{ij}\right) \\
&= E_i + \sum_j \left(\tilde{P}_{ij} - P_{ij}\right)\nu_j/P_{i\cdot} - \left(\sum_j P_{ij}\nu_j/P_{i\cdot}^2\right)\sum_j \left(\tilde{P}_{ij} - P_{ij}\right) + o_p(n^{-1/2}) \\
&= E_i + \sum_j \left(\tilde{P}_{ij} - P_{ij}\right)(\nu_j - E_i)/P_{i\cdot} + o_p(n^{-1/2}).
\end{aligned}$$

Normality follows from the normality of $\tilde{P}_{ij}$. As for the covariance, note that for $i \neq k$ we have

$$
\begin{aligned}
n \operatorname{Cov}\left(\tilde{E}_i, \tilde{E}_k\right) &= n \sum_j \sum_e \operatorname{Cov}\left(\tilde{P}_{ie}, \tilde{P}_{kj}\right)(\nu_j - E_k)(\nu_e - E_i)/P_{i\cdot}P_{k\cdot} + o(1) \\
&= -\sum_j \sum_e P_{ie}P_{kj}(\nu_j - E_k)(\nu_e - E_i)/P_{i\cdot}P_{k\cdot}] + o(1) \\
&= \left[\sum_e P_{ie}(\nu_e - E_i)\right]\left[\sum_j P_{kj}(\nu_j - E_k)/P_{i\cdot}P_{k\cdot}\right] + o(1) = o(1).
\end{aligned}
$$

For $i = k$ the above terms become simply

$$
\begin{aligned}
n \operatorname{Var}\left(\tilde{E}_i\right) &= \sum_j P_{ij}(\nu_j - E_j)^2/P_{i\cdot}^2 + o(1) \\
&= \operatorname{Var}(\nu|i)/P_{i\cdot} + o(1).
\end{aligned}
$$

Finally, by similar arguments we obtain

$$
\begin{aligned}
n \operatorname{Cov}\left[\tilde{E}_i, \tilde{F}_j\right] &= \sum_{m,e} n \operatorname{Cov}\left(\tilde{P}_{im}, \tilde{P}_{ej}\right)(\nu_m - E_i)(\mu_e - F_j)P_{i\cdot}P_{\cdot j} + o(1) \\
&= (P_{ij}/P_{i\cdot}P_{\cdot j})(\nu_j - E_i)(\mu_i - F_j) \\
&\quad - \sum_{m,e} P_{im}P_{ej}(\nu_m - E_i)(\mu_e - F_j)/P_{i\cdot}P_{\cdot j} + o(1) \\
&= (P_{ij}/P_{i\cdot}P_{\cdot j})(\mu_j - E_i)(\mu_i - F_j) + o(1). \qquad \square
\end{aligned}
$$

COROLLARY. *If $\mu_i = \mu_{i+1}$ and $\nu_i = \nu_{j+1}$ for some i and j, then*

$$
n \operatorname{Cov}\left\{[E_i(\nu) - E_{i-1}(\nu)], [F_j(\nu) - F_{j-1}(\mu)]\right\} \to 0.
$$

LEMMA 2. *Let $C_n \to 0$ and suppose that $\bar{\mu}$ and $\bar{\nu}$ are consistent vector estimates for parameter vectors μ and ν, respectively, such that*

$$
P\left\{\max_{1 \le i \le I}[|\bar{\mu}_i - \mu_i|] + \max_{1 \le j \le J}[|\bar{\nu}_j - \nu_j|] > C_n\right\} \to 0.
$$

Then for $2 \le i \le I$, $2 \le j \le J$,

(i) $$P\left\{\left[\tilde{E}_i(\nu) - \tilde{E}_{i-1}(\nu)\right]\left[\tilde{E}_i(\bar{\nu}) - \tilde{E}_{i-1}(\bar{\nu})\right] < 0\right\} \to 0,$$

(ii) $$P\left\{\left[\tilde{F}_j(\mu) - \tilde{F}_{j-1}(\mu)\right]\left[\tilde{F}_j(\bar{\mu}) - \tilde{F}_{j-1}(\bar{\mu})\right] < 0\right\} \to 0,$$

where $\tilde{E}_i(\bar{\nu})$ and $\tilde{F}_j(\bar{\mu})$ are defined as $\tilde{E}_i(\nu)$ and $\tilde{F}_j(\mu)$ in (2.4) with $\bar{\nu}$ and $\bar{\mu}$ replacing ν and μ, respectively.

PROOF. To facilitate the proof, two distinct cases are considered. Suppose that $\mu_i > \mu_{i-1}$. This implies $E_i > E_{i-1}$. We have

$$(3.1)\quad \left|\tilde{E}_k(\nu) - \tilde{E}_k(\bar{\nu})\right| = \left| \sum_j \tilde{P}_{kj}(\nu_j - \bar{\nu}_j) \sum_j \Big/ \tilde{P}_{kj} \right| \to 0, \qquad k = i, i-1.$$

By Lemma 1 we have

$$\tilde{E}_i(\nu) - \tilde{E}_{i-1}(\nu) \to E_i(\nu) - E_{i-1}(\nu) > 0,$$

and by (3.1)

$$\tilde{E}_i(\bar{\nu}) - \tilde{E}_{i-1}(\bar{\nu}) \to E_i(\nu) - E_{i-1}(\nu),$$

which implies (i) above.

Suppose now that $\mu_i = \mu_{i-1}$ for some i. Let $P_{j|i} = P_{ij}/P_{i\cdot}$. The equality $\mu_i = \mu_{i-1}$ implies $P_{j|i} - P_{j|i-1} = 0$ and hence $\tilde{P}_{j|i} - \tilde{P}_{j|i-1} = O_p(\sqrt{n})$ and

$$(3.2)\quad \begin{aligned}\left\{\tilde{E}_i(\bar{\nu}) - \tilde{E}_{i-1}(\bar{\nu})\right\} - \left\{\tilde{E}_i(\nu) - \tilde{E}_{i-1}(\nu)\right\} &= \sum_j \left(\tilde{P}_{j|i} - \tilde{P}_{j|i-1}\right)(\bar{\nu}_j - \nu_j) \\ &= O_p\left(C_n n^{-1/2}\right).\end{aligned}$$

From Lemma 1 we obtain

$$(3.3)\quad P\left\{\left|\tilde{E}_i(\nu) - \tilde{E}_{i-1}(\nu)\right| \le (C_n/n)^{1/2}\right\} \to 0.$$

Now by (3.2) and (3.3)

$$\begin{aligned}&P\left\{\left[\tilde{E}_i(\nu) - \tilde{E}_{i-1}(\nu)\right]\left[\tilde{E}_i(\bar{\nu}) - \tilde{E}_{i-1}(\bar{\nu})\right] > 0\right\} \\ &\qquad > P\left\{\left|\tilde{E}_i(\nu) - \tilde{E}_{i-1}(\nu)\right| \Rightarrow (C_n/n)^{1/2}\right\}\end{aligned}$$

and

$$P\left\{\left|E_i(\bar{\nu}) - E_{i-1}(\bar{\nu}) - \tilde{E}_i(\nu) + \tilde{E}_{i-1}(\nu)\right| < (C_n/n)^{1/2}\right\} \to 1,$$

so (i) above follows. Result (ii) is obtained similarly. □

By similar arguments the following lemma can be proven.

LEMMA 3. *Let S and S' be two mutually exclusive subsets of* $\{1, \dots, I\}$ *and let $\bar{\nu}$ be a consistent estimate of ν. Then*

$$P\left\{\left[\sum_{i\in S} P_{i\cdot}\tilde{E}_i(\nu) - \sum_{i\in S'} P_{i\cdot}\tilde{E}_i(\nu)\right]\left[\sum_{i\in S} P_{i\cdot}\tilde{E}_i(\bar{\nu}) - \sum_{i\in S'} P_{i\cdot}\tilde{E}_i(\bar{\nu})\right] < 0\right\} \to 0.$$

A similar result holds for $\tilde{F}_j(\mu)$.

We now prove that the order-restricted maximum likelihood estimates μ^* and ν^* are indeed the unrestricted maximum likelihood estimates of μ and ν

in a properly collapsed table. This table is obtained by combining particular rows and particular columns. To accomplish this task, we need some notation.

Let $T = \{S_1, \dots, S_K\} \times \{R_1, \dots, R_M\}$, where $S_1, \dots, S_K$ and $R_1, \dots, R_M$ are partitions of $\{1, \dots, I\}$ and $\{1, \dots, J\}$, respectively, such that $i < i'$ and $j < j'$ for all $i \in S_k$, $i' \in S_{k+1}$, $j \in R_m$, $j' \in R_{m+1}$, $1 \le k \le K-1$ and $1 \le m \le M-1$. With respect to such partitions we define the collapsed table

$$P^T_{km} = \sum_{i \in S_k} \sum_{j \in R_m} P_{ij}, \qquad 1 \le k \le K, 1 \le m \le M.$$

We will use superscript T to denote that the given quantity is with respect to this table. Under this notation, for example, μ^T_k denotes the unrestricted mle for μ_k in the collapsed table.

LEMMA 4. *There exist partitions $S_1, \dots, S_K$ and $R_1, \dots, R_M$ as above such that the order-restricted maximum likelihood estimates (μ^*, ν^*) satisfy $\mu^*_i = \hat{\mu}^T_k$ and $\nu^*_j = \hat{\nu}^T_m$, $i \in S_k$, $j \in R_m$.*

PROOF. The standard way to find the maximum of the log likelihood subject to the order restrictions (2.1) is the following: Consider all partitions as above. For a given pair of partitions $S_1, \dots, S_K$, and $R_1, \dots, R_M$, find the global maximum of the log likelihood when $\mu_i = \hat{\mu}^T_k$ and $\nu_j = \hat{\nu}^T_m$, $i \in S_k$, $j \in R_m$, is imposed. The maximum value of the order-restricted log likelihood is the largest of the global maxima for which $\hat{\mu}^T_1 \le \cdots \le \hat{\mu}^T_k$ and $\hat{\nu}^T_1 \le \cdots \le \hat{\nu}^T_m$ is satisfied. Consider now such a partition. Then the log likelihood is

$$\begin{aligned}
&\sum_{ij} P_{ij}\left[\log \alpha_i + \log \beta_j + \phi \mu_i \nu_j\right] \\
&\qquad = \sum_i P_{i\cdot} \log \alpha_i + \sum_j P_{\cdot j} \log \beta_j + \phi \sum_{km} P^T_{km} \mu^T_k \nu^T_m,
\end{aligned}$$

which should be maximized subject to

$$\sum_i \sum_j \alpha_i \beta_j \exp(\phi \mu_i \nu_j) = \sum_k \sum_m \exp(\phi \mu^T_k \nu^T_m)\left(\sum_{i \in S_k} \sum_{j \in R_m} \alpha_i \beta_j\right),$$

$$l - 1 = \sum_i P_{i\cdot} \mu^l_i = \sum_k P^T_{k\cdot} (\mu^T_k)^l, \qquad l = 1, 2,$$

and

$$l - 1 = \sum_j P_{\cdot j} \nu^l_j = \sum_m P^T_{\cdot m} (\nu^T_m)^l, \qquad l = 1, 2. \qquad \square$$

The maximum can be found using Lagrange multipliers. It is easily verified then that $(\hat{\mu}^T, \hat{\nu}^T)$ are the unrestricted mle's for the collapsed table and satisfy the likelihood-solution equations (2.5).

Let $\hat{T}$ be the collapsed table such that (μ^*, ν^*) correspond to the (restricted) mle $(\hat{\mu}^T, \hat{\nu}^T)$ with respect to $\hat{T}$. The notation $\hat{T}$ is used to indicate the specific ("properly") collapsed table yielding the desired estimations, while the notation T used earlier indicates a collapsed table in general.

Finally, we need the following lemma.

LEMMA 5. *The order-restricted estimates μ^* and ν^* satisfy, in the properly collapsed $K \times M$ table $\hat{T}$,*

$$\tilde{E}_{i-1}^{\hat{T}}(\hat{\nu}^{\hat{T}}) \le \tilde{E}_i^{\hat{T}}(\hat{\nu})^{\hat{T}}, \qquad 2 \le i \le K,$$

$$\tilde{F}_{j-1}^{\hat{T}}(\hat{\mu}^{\hat{T}}) \le \tilde{F}_j^{\hat{T}}(\hat{\mu}^{\hat{T}}), \qquad 2 \le j \le M.$$

PROOF. From the likelihood equations in (2.5) for the unrestricted model (2.1), we have the following equations for $1 \le i \le K$,

$$\sum_j \tilde{P}_{ij}^{\hat{T}} \hat{\nu}_j^{\hat{T}} = \alpha_i^{\hat{T}} \sum_j \beta_j^{\hat{T}} \hat{\nu}_j^{\hat{T}} \exp\left(\hat{\phi} \hat{\mu}_i^{\hat{T}} \hat{\nu}_j^{\hat{T}}\right), \tag{3.4}$$

$$\tilde{P}_{i\cdot} = \hat{\alpha}_i^{\hat{T}} \sum_j \beta_j^{\hat{T}} \exp\left(\hat{\phi} \hat{\mu}_i^{\hat{T}} \hat{\nu}_j^{\hat{T}}\right). \tag{3.5}$$

Upon dividing (3.4) by (3.5), we obtain $\tilde{E}_i^{\hat{T}}(\nu)$ in (2.4) expressed as a function of $\hat{\mu}^{\hat{T}}$. Now

$$\begin{aligned} \frac{\partial \tilde{E}_i(\hat{\nu}^{\hat{T}})}{\partial \hat{\mu}_i^{\hat{T}}} &= \hat{\phi}\left\{\left[\frac{\sum_j \hat{\beta}_j^{\hat{T}} \hat{\nu}_j^{\hat{T}} \exp\left(\hat{\phi} \hat{\mu}_i^{\hat{T}} \hat{\nu}_j^{\hat{T}}\right)}{\sum_j \hat{\beta}_j \exp\left(\hat{\phi} \hat{\mu}_i^{\hat{T}} \hat{\nu}_j^{\hat{T}}\right)}\right] - E_i^2(\nu)\right\} \\ &= \phi \operatorname{Var}(\nu | i) \ge 0. \end{aligned} \tag{3.6}$$

Result (3.6) indicates that $E_i^{\hat{T}}(\hat{\nu}^{\hat{T}})$ is an increasing function in $\hat{\mu}_i^{\hat{T}}$. Arguments similar to those mentioned above yield that $\tilde{E}_i^{\hat{T}}(\hat{\nu}^{\hat{T}})$ is also an increasing function of $\hat{\mu}_i^{\hat{T}}$. Now since, by definition, $\hat{\mu}_{i+1} \ge \hat{\mu}_i$, Lemma 5 follows at once. □

The results proven so far can be summarized as follows: The restricted estimates are given by the unrestricted estimates pertaining to some collapsed table (Lemma 4). If this table is *properly* collapsed, then the order is preserved among the quantities $\tilde{E}(\cdot)$ and among the quantities $\tilde{F}(\cdot)$ (Lemma 5). With high probability, the order between the $\tilde{E}(\cdot)$'s and the order between the $\tilde{F}(\cdot)$'s does *not* depend on the consistent estimate used whether it is $\hat{\mu}$ or μ^* (Lemmas 2 and 3).

All these results together provide the way of deriving the "properly collapsed table": Collapse the original table over rows and over columns for which the quantities $\tilde{E}(\hat{\nu})$ and $\tilde{F}(\hat{\mu})$ violate the desired order, respectively. Amalgamation should be done with weights that are the relevant relative frequencies. Here, for instance, amalgamation with respect to rows $i-1$ and i is

$$\left\{\tilde{P}_{i\cdot} \tilde{E}_i(\hat{\nu}) + \tilde{P}_{(i-1)} \tilde{E}_{i-1}(\hat{\nu})\right\} \Big/ \left\{\tilde{P}_{i\cdot} + \tilde{P}_{(i-1)}\right\}.$$

Amalgamation is, as stated above, equivalent to collapsing the table over rows and column for which the quantities $\tilde{E}_i(\hat{\nu})$ and $\tilde{F}_j(\hat{\mu})$ violate the desired order.

We have therefore derived and proven the following estimation procedure.

1. Compute the unrestricted maximum likelihood estimates $\hat{\mu}$ and $\hat{\nu}$. If these estimates follow the desired order, then stop. Otherwise, proceed to step 2.
2. Calculate $\tilde{E}_i(\hat{\nu})$ and $\tilde{F}_j(\hat{\mu})$ and amalgamate them (by pooling adjacent violators). Lemma 1 and its corollary ensure that with high probability, amalgamation can be done separately (independently) for rows and columns.
3. Repeat steps 1 and 2 until the quantities $\tilde{F}_j^T(\hat{\mu}^T)$ and $\tilde{E}_i^T(\hat{\nu}^T)$ follow the desired order.

The unrestricted maximum likelihood estimates pertaining to the (first) collapsed table for which $\tilde{E}_i(\hat{\nu})$ and $\tilde{F}_j(\hat{\mu})$ follow the desired order are the final estimates. The above lemmas prove that if the maximum likelihood estimates exist (and are unique), then with probability approaching 1 (as $n \to \infty$) the estimates produced by our procedure are the order-restricted maximum likelihood estimates. Moreover, it follows from the above lemmas that if n is large, then (with high probability) no more than one-step amalgamation is needed for obtaining the desired estimates.

4. Testing for fit. The lemmas proven in Section 3 allow us to use some well-known results in order-restricted inference under normal theory. Let m_{ij} denote the expected frequencies in cell (i, j) under the unrestricted RC model and let m_{ij}^* denote the expected frequencies in cell (i, j) under the order-restricted RC model. The statistic

$$\bar{\chi}^2 = \sum_i \sum_j (m_{ij} - m_{ij}^*)^2 / m_{ij}^* \tag{4.1}$$

is an appropriate statistic for testing the nested hypothesis H_0: Order-restricted RC model against H_1: Unrestricted RC model. Under H_0 (within H_1) this statistic has an asymptotic distribution which is a mixture of central chi-squared distributions similar (in principle) to the distribution derived in Barlow, Bartholomew, Bremner and Brunk (1972), page 126.

To be able to derive a test stated in terms of $E_i(\nu)$ and $F_j(\nu)$ which is asymptotically equivalent to the test in (4.1), assume that the following order restrictions apply for some $I' \leq I$, $J' \leq J$,

$$\begin{aligned} \mu_1 = \mu_2 = \cdots = \mu_{I'} < \mu_{I'+1} < \cdots < \mu_I, \\ \nu_1 = \nu_2 = \cdots = \nu_{J'} < \nu_{J'+1} < \cdots < \nu_J. \end{aligned} \tag{4.2}$$

LEMMA 6. *Let $U_1, \dots, U_I, V_1, \dots, V_J$ be independent r.v.'s, where $U_i \sim N(\mu_U^i, \operatorname{Var}(\tilde{E}_i(\nu))$, $V_j \sim N(\nu_V^j, \operatorname{Var}(\tilde{F}_j(\mu))$. Let $T_U(U_1, \dots, U_I)$ be the standard χ^2-test for testing $\mu_U^1 \leq \cdots \leq \mu_U^I$ against "at least one inequality is not satisfied." Let $T_V(V_1, \dots, V_J)$ be defined similarly.*

Then the $\bar{\chi}^2$-test (4.1) *is asymptotically distributed as*

$$T = T_U\big(\tilde{E}_1(\hat{\nu}), \dots, \tilde{E}_I(\hat{\nu})\big) + T_V\big(\tilde{F}_1(\hat{\mu}), \dots, F_J(\hat{\mu})\big).$$

PROOF. Note first that by Lemma 5, H_0 and H_1 can be expressed by:

H_0: $E_1(\nu) \le E_2(\nu) \le \cdots \le E_I(\nu)$, $F_1(\mu) \le F_2(\mu) \le \cdots \le F_J(\mu)$.
H_1: At least one inequality does not hold.

Now by Lemma 2,

$$\tilde{E}_{I'}(\nu) < \tilde{E}_{I'+1}(\nu) < \cdots < \tilde{E}_I(\nu)$$

and

$$\tilde{F}_{J'}(\mu) < \tilde{F}_{J'+1}(\mu) < \cdots < \tilde{F}_J(\mu)$$

with probability converging to 1. By Lemma 2 these strict inequalities hold with probability converging to 1 even if we replace ν and μ by consistent estimators, for example, $\hat{\mu}$ and $\hat{\nu}$. Thus we need consider only

$$\tilde{E}_2(\hat{\nu}) - \tilde{E}_1(\hat{\nu}), \dots, \tilde{E}_{I'}(\hat{\nu}) - \tilde{E}_{I'-1}(\hat{\nu})$$

and

$$\tilde{F}_2(\hat{\mu}) - \tilde{F}_1(\hat{\mu}), \dots, \tilde{F}_{J'}(\hat{\mu}) - \tilde{F}_{J'-1}(\hat{\mu}).$$

By Lemma 1 and its corollary, these differences are asymptotically distributed like the differences between the U's and the V's, respectively. Moreover, by Lemma 2

$$\tilde{E}_{i+1}(\hat{\nu}) - \tilde{E}_i(\hat{\nu}) = \tilde{E}_{i+1}(\nu) - \tilde{E}_i(\nu) + o_p(n^{-1/2}), \qquad i = 1, \dots, I',$$

$$\tilde{F}_{j+1}(\hat{\mu}) - \tilde{F}_j(\hat{\mu}) = \tilde{F}_{j+1}(\mu) - \tilde{F}_j(\mu) + o_p(n^{-1/2}), \qquad j = 1, \dots, J'.$$

So, the test appropriate for the vectors $U_1, \dots, U_I, V_1, \dots, V_J$ can equivalently be used on $\tilde{E}_1(\hat{\nu}), \dots, \tilde{E}_I(\hat{\nu}), \tilde{F}_1(\hat{\mu}), \dots, \tilde{F}(\hat{\mu})$.

Let e denote the total number of equalities among the μ's and among the ν's. Let e_1 denote the number of equalities among the μ's alone. Let $P_X(e_1|I')$ denote the conditional probability of having e_1 equalities among the μ's given I' μ's are equal. Let $P_Y(e_2|J')$ be similarly defined with respect to the ν's where $e_2 = e - e_1$. By the corollary to Lemma 1 regarding the asymptotic independence between $(\tilde{E}_2 - \tilde{E}_1), \dots, (\tilde{E}_{I'} - \tilde{E}_{I'-1})$ and $(\tilde{F}_2 - \tilde{F}_1), \dots, (\tilde{F}_{I'} - \tilde{F}_{I'-1})$, we can write β_e (the mixture probability of having a total e equalities among the same parameters given I' equal μ's and J' equal ν's) as

$$\beta_e = \sum_{e_1=0}^{e} P_x(e_1|I')P_y[(e - e_1)|J'].$$

Now the statistic (4.1) has asymptotically and under the RC model with restrictions (4.2) the distribution given by the mixture

$$P(\bar{\chi}^2 > c) = \sum_{e=1}^{I'+J'-2} \beta_e P\left(x^2_{(e)} > c\right), \tag{4.3}$$

where $\chi^2_{(e)}$ is the central chi-squared distribution with e degrees of freedom. $\square$

The mixture probabilities β_e depend on P_{ij} as is evident from the variance structure in Lemma 1. Existing tables for mixture probabilities [i.e., Barlow, Bartholomew, Bremner and Brunk (1972), page 363] have no use in our context since they are based on equal weights (equal variances of E_i, $1 \leq i \leq I$, and equal variances of F_j, $1 \leq j \leq J$, in our case). In order to be able to explicitly calculate the probabilities β_e, one must perform a Monte Carlo study or a bootstrap estimation which is beyond the scope of this paper. In some special cases, however (as is shown in the next section), the probabilities β_e can be readily computed.

If the ordered RC model is to be tested against the *general alternative*, then an appropriate test statistic can be derived for such a case by replacing m_{ij} in (4.1) by the cell counts n_{ij}. By Goodman (1986) the resulting statistic will have (under the ordered RC model) an asymptotic distribution which is a mixture of central chi-squares with $(I - 1)(J - 2) + e$ degrees of freedom.

5. Discussion. In this paper we have developed an estimation procedure for estimating the score parameters of Goodman's RC model under order restrictions. This procedure is basically quite similar to common order-restricted estimation techniques under normal theory except for one important aspect. Direct use of normal theory results is not allowed in our context as the vector $\hat{\mu}$ is not independent (even asymptotically) of the vector $\hat{\nu}$. We have bypassed this difficulty by working with the quantities $\tilde{E}_i$ and $\tilde{F}_j$ rather than with $\hat{\mu}_i$ and $\hat{\nu}_j$, and have shown that $\tilde{E}_i$ and $\tilde{F}_j$ are jointly asymptotically normal with a specified covariance matrix. Properly amalgamating the $\tilde{E}_i$ and the $\tilde{F}_j$ results in ordered estimates (for the score parameters) which are asymptotically equivalent to the restricted maximum likelihood estimates (if they exist). All results proven in this paper are asymptotic. The nonasymptotic case involves a possible violation of the Kuhn–Tucker conditions as the log likelihood with respect to the RC model may not be concave. Therefore, we have no guarantee, in the nonasymptotic case, that the order-restricted estimates are unique. This is in contrast to the situation with the R model and the C model, where the Kuhn–Tucker conditions are always satisfied and there are no problems with existence and uniqueness [Agresti, Chuang and Kezouh (1987)]. Testing the restricted RC model for fit is relatively straightforward and is based on the same principles as in the common normal theory case. The only exception is that, while mixture probabilities in the regular case are assumed to be based on equal weights, such probabilities in our context are based on possibly unequal weights. We know of no general tables for such mixture probabilities. It is not the intent of this paper to develop such tables and we believe that relatively simple Monte Carlo studies or bootstrapping techniques can be used to explicitly calculate such probabilities.

There are situations, however, in which testing procedure (4.1) can be used with mixture probabilities that are readily obtained. To exemplify such situations, assume that only *two* column scores are assumed to be equal while all other scores are strictly monotone [compare with Goodman (1981), Section 6]. In this case $I' = 1$, $J' = 2$ and $\beta_1 = 1/2$ by (4.3). Testing the above proposed model for fit at the (upper) α percentile point leads to

$$\alpha = \tfrac{1}{2}P\left(x^2_{(1)} > c\right),$$

implying

$$c = x^2_{(1)}(2\alpha). \tag{5.1}$$

Goodman (1981) presents a test for equality of scores. If equality of scores is tested under the *unrestricted* RC model, then result (5.1) suggests that in such cases the test procedure (4.1) together with (4.3) is *more conservative* than the test mentioned by Goodman (1981). Professor Goodman drew our attention, in a personal communication, to the fact that a similar result was obtained by him earlier [i.e., Goodman (1985), page 62].

Acknowledgments. The authors wish to thank Alan Agresti whose earlier work on the subject inspired this work. Thanks are also due to Leo Goodman, Shelby Haberman, Peter McCullagh, two referees and an Associate Editor for useful comments and suggestions.

REFERENCES

AGRESTI, A., CHUANG, C. and KEZOUH, A. (1987). Order-restricted source parameters in association models for contingency tables. *J. Amer. Statist. Assoc.* **82** 619–623.

BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference under Order Restrictions*. Wiley, New York.

GOODMAN, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *J. Amer. Statist. Asoc.* **74** 537–552.

GOODMAN, L. A. (1981). Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **76** 320–334.

GOODMAN, L. A. (1985). The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing values. *Ann. Statist.* **13** 10–69.

GOODMAN, L. A. (1986). Some useful extensions of the usual correspondence analysis approach and the usual log-linear approach in the analysis of contingency tables. *Internat. Statist. Rev.* **54** 254–309.

NUESCH, P. E. (1966). On the problem of testing locations in multivariate populations for restricted alternatives. *Ann. Math. Statist.* **37** 113–119.

DEPARTMENT OF STATISTICS
FACULTY OF SOCIAL SCIENCES
HEBREW UNIVERSITY OF JERUSALEM
JERUSALEM 91905
ISRAEL