

# Testing Stochastic Orders in Tails of Contingency Tables

Chi Tim Ng, Johan Lim, Kyu S. Hahn \*

## Abstract

Testing for the difference in the strength of bivariate association in two independent contingency tables is an important issue that finds applications in various disciplines. Currently, many of the commonly used tests are based on single index measures of association. More specifically, one obtains single index measurements of association from two tables and compares them based on asymptotic theory. Although they are usually easy to understand and use, often much of the information contained in the data is lost with single index measures. Accordingly, they fail to fully capture the association in the data. To remedy this shortcoming, we introduce a new summary statistic measuring various types of association in a contingency table. Based on this new summary statistic, we propose a likelihood ratio test comparing the strength of association in two independent contingency tables. The proposed test exams the stochastic order between summary statistics. We derive its asymptotic null distribution and demonstrate that the least favorable distributions are chi-bar distributions. We numerically compare the power of the proposed test to that of the tests based on single index measures. Finally, we provide two examples illustrating the new summary statistics and the related tests.

**Key words:** Association, contingency table, likelihood ratio test, ordered categorical variable, stochastic order, tail probabilities.

## 1 Introduction

Testing for the difference in the strength of association between two ordered categorical variables in two contingency tables is an important issue that finds applications in reliability theory, sociology, education, and many other fields. Just to list a few examples, Hand et al. (1994) assess the degree of association in socioeconomic status of fathers and sons. In education research, Loughin and Scherer (1998) examine the association between farmers'

---

\*Chi Tim Ng is at the Department of Applied Mathematics, The Hong Kong Polytechnic, Hong Kong. Johan Lim is at the Department of Statistics, Seoul National University, Seoul, Korea. Kyu S. Hahn is at Underwood International College, Yonsei University, Seoul, Korea and at the Department of Communication Studies, University of California-Los Angeles, Los Angeles, CA, USA. Johan Lim's research is supported by New Faculty Research Fund at College of Natural Sciences, Seoul National University. All correspondence should be sent to Johan Lim at [johanlim@snu.ac.kr](mailto:johanlim@snu.ac.kr).

education levels and their sources of veterinary information. In economics, Umesh (1995) examines the association between country origin and user gender in care selection. In all these examples, the researchers' primary interest lies in assessing the positive association between two (ordered) categorical variables and requires determining the level of statistical significance.

Many existing methods for testing the association between categorical variables rely on single index measures. The existing work on association measures dates back to Karl Pearson's  $\chi^2$ -test in early 1900s. Stuart (1953) also considers a measure of association based on concordant and discordant pairs and applies it to testing for the difference in the bivariate association in two contingency tables. Cohen (1960) introduces a single index measure, named the  $\kappa$  or kappa measure, of association that makes use of the number of perfectly matched pairs. Modifying the  $\kappa$  measure, Fleiss and Cohen (1973) proposed to use a weighted  $\kappa$  measure, where the weights are subjectively determined. Additionally, the notion of  $\kappa$  is generalized to the cases with two or more ratings or with varying sets of ratings (Fleiss, 1971; Light, 1971; Landis and Koch, 1977a,b; Fleiss and Cuzick, 1979; Lau, 1993). Using a logistic regression model, Barlow (1996) further suggested a technique for adjusting  $\kappa$  for covariate effects. These single index association measures offer convenient tools for assessing the association between two categorical variables.

A popular alternative to index based methods are the approaches based on log-linear models as pointed out by Tomizawa and Hatanaka (2001), Tomizawa and Tokunaga (2006) and Liu and Agresti (2005). One most general model in previous literature is given by Goodman (1985):

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \sum_{k=1}^M \beta_k u_{ik} u_{jk}, \quad (1)$$

where  $M \leq \min(r-1, c-1)$ . The scores of  $u_{ik}$  and  $u_{jk}$  define different models. Unfortunately, however,  $u_{ik}$  and  $u_{jk}$  are defined subjectively and the results vary accordingly.

Despite their convenience, single index measures often lose much of the information contained in the data, failing to fully reflect the association between two categorical variables

in a table. For example, let us consider Cohen's kappa measure defined as

$$\kappa^{(k)} = \frac{\sum_i \pi_{ii}^{(k)} - \sum_i \pi_{i+}^{(k)} \pi_{+i}^{(k)}}{1 - \sum_i \pi_{i+}^{(k)} \pi_{+i}^{(k)}}.$$

The kappa measure only counts the concordant and dis-concordant pairs while ignoring the magnitude of disagreement. To be more specific, given Table  $\{\pi_{ij}, i, j = 1, 2, \dots, m\}$ , the kappa measure uses the summary statistic on  $\sum_i \pi_{ii}$  to obtain an index score. By doing so, the kappa measure clearly disregards much of the information contained in the raw data. To illustrate this point, for example, the kappa measure cannot tell the difference between

$$\alpha \mathbf{I}_5 + (1 - \alpha) \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \alpha \mathbf{I}_5 + (1 - \alpha) \frac{1}{8} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

where  $\mathbf{I}_5$  is a  $5 \times 5$  identity matrix. On the other hand, the latter clearly shows a higher level of association than the former.

In the current paper, we introduce a new multivariate summary measure of association between two ordered categorical variables. Our method allows us to define local association measures as well as a single index global association measure. More specifically, consider a  $m$ -square contingency table  $\mathbf{C}_k$ ,  $k = 1, 2$ , where cell probabilities are  $\{\pi_{i,j}^{(k)}, i = 1, \dots, m, j = 1, \dots, m, k = 1, 2\}$ . We summarize the cell probabilities based on  $d = i - j$  so that the resulting marginal probabilities would be

$$\mathbf{q}_k = (q_{0k}, q_{-1,k}, q_{-2,k}, \dots, q_{-(m-1),k}, q_{1k}, \dots, q_{(m-2)k}, q_{(m-1)k}),$$

where

$$q_{dk} = \sum_{\{(i,j):i-j=d\}} \pi_{ij}^{(k)}.$$

We let  $q_{-1} = \sum_{k=0}^{-(m-1)} q_{1k}$ ,  $q_{+1} = \sum_{k=0}^{(m-1)} q_{1k}$ ,  $q_{-2} = \sum_{k=0}^{-(m-1)} q_{2k}$ , and  $q_{+2} = \sum_{k=0}^{(m-1)} q_{2k}$ . Although we only consider square contingency tables, our results can easily be extended to any two-way contingency tables by redefining  $ds$ . We analyze non-square contingency tables in our examples presented in Section 4.

The multivariate summary statistic allows us to examine several types of association between  $\mathbf{C}_1$  and  $\mathbf{C}_2$ , which we refer them as "high-low" association (SO-HL)", "low-high" association (SO-LH), and "global" association (GO). In this paper, we study the likelihood ratio test (LRT) for testing stochastic orders in  $\mathbf{q}_k$ s to test the degree of aforementioned associations between  $\mathbf{C}_1$  and  $\mathbf{C}_2$ . We show that the asymptotic null distribution of LRTs is a chi-bar squared distribution and find the least favorable distribution for the SO-HL, the SO-LH, and the GO.

The remainder of this paper is organized as follows. In Section 2, we introduce three types of association based on the aforementioned multivariate summary statistics. In Section 3, we develop LRTs for testing stochastic orderings in  $q_{lk}$ s. In Section 4, we numerically compare the power of the proposed stochastic order test to that of Stuart's test and Cohen's  $\kappa$  based test, by assessing the global stochastic order between  $\mathbf{C}_1$  and  $\mathbf{C}_2$ . In Section 5, we provide two examples. In the first example, we test for the magnitude of gender and ethnicity-based inequality in returns to education. In the second example, we reanalyze the eye-sight power data used in Stuart (1953).

## 2 Types of associations

The multivariate summary statistic introduced earlier allows us to examine several types of association between  $\mathbf{C}_1$  and  $\mathbf{C}_2$ .

- the "high-low" association (SO-HL):

$\mathbf{C}_2$  has a higher level of "high-low" association than  $\mathbf{C}_1$  ( $\mathbf{C}_2 \succeq_{\text{S-HL}} \mathbf{C}_1$ ), if

$$\frac{q_{01}/q_{-1}}{(1/q_{-1})(q_{01} + \sum_{i=-1}^l q_{i1})} \leq \frac{q_{02}/q_{-2}}{(1/q_{-2})(q_{02} + \sum_{i=-1}^l q_{i2})}, \quad \text{for } l = -1, -2, \dots, -(m-1). \quad (2)$$

- the "low-high" association (SO-LH):

$\mathbf{C}_2$  has a higher level of low-high association than  $\mathbf{C}_1$  ( $\mathbf{C}_2 \succeq_{\text{S-LH}} \mathbf{C}_1$ ), if

$$\begin{aligned} q_{01}/q_{+1} &\leq q_{02}/q_{+2} \\ (1/q_{+1})(q_{01} + \sum_{i=1}^l q_{i1}) &\leq (1/q_{+2})(q_{02} + \sum_{i=1}^l q_{i2}), \quad \text{for } l = 1, 2, \dots, (m-1). \end{aligned} \quad (3)$$

- the "global" association (SO):

$\mathbf{C}_2$  has a higher level of global association than  $\mathbf{C}_1$  ( $\mathbf{C}_2 \succeq_{\text{S}} \mathbf{C}_1$ ), if

$$\begin{aligned} q_{01} &\leq q_{02} \\ q_{01} + \sum_{i=-1}^l q_{i1} &\leq q_{02} + \sum_{i=-1}^l q_{i2}, \quad \text{for } l = -1, -2, \dots, -(m-2), \\ q_{01} + \sum_{i=1}^l q_{i1} &\leq q_{02} + \sum_{i=1}^l q_{i2}, \quad \text{for } l = 1, 2, \dots, (m-2). \end{aligned} \quad (4)$$

All three types of association presume that there would be clusters of data points around the diagonal if two categorical variables are highly associated with each other.

Global association (SO) is closely connected to the order based on the kappa measure and the D-symmetry in Goodman (1979). Suppose  $\mathbf{C}_1 \succeq_{\text{K}} \mathbf{C}_2$  denotes  $\mathbf{C}_1$  has a higher  $\kappa$  index than  $\mathbf{C}_2$ . Simple algebra yields the following relationships:

- (1)  $\{\mathbf{C}_1 \succeq_{\text{S}} \mathbf{C}_2\} \subset \{\mathbf{C}_1 \succeq_{\text{K}} \mathbf{C}_2\}$ .
- (2)  $\{\mathbf{C}_1 \succ_{\text{S}} \mathbf{C}_2\} \cap \{\mathbf{C}_1 =_{\text{K}} \mathbf{C}_2\} \neq \emptyset$ .
- (3)  $\{\mathbf{C}_1 =_{\text{S}} \mathbf{C}_2\} \subset \{\mathbf{C}_1 =_{\text{K}} \mathbf{C}_2\}$ .
- (4)  $\{\mathbf{C}_1 \succeq_{\text{S}} \mathbf{C}_2\}^{\text{C}} \cap \{\mathbf{C}_1 \succ_{\text{K}} \mathbf{C}_2\} \neq \emptyset$ .
- (5)  $\{\mathbf{C}_1 \succeq_{\text{S}} \mathbf{C}_2\}^{\text{C}} \cap \{\mathbf{C}_1 =_{\text{K}} \mathbf{C}_2\} \neq \emptyset$ .

### 3 LRTs

In this section, we study the LRT for testing three types of associations (*i.e.*, the SO-HL, the SO-LH, and the GO) introduced in the previous section. First, for SO-HL and SO-LH, we use the results obtained from testing the simple stochastic order in multinomial probabilities. By doing so, we show that the asymptotic null distribution of LRTs is a chi-bar squared distribution and its common least favorable distribution is

$$\frac{1}{2}\chi_{m-2}^2 + \frac{1}{2}\chi_{m-1}^2, \quad (5)$$

where  $\chi_{m-2}^2$  and  $\chi_{m-1}^2$  are independent chi-square distributions with  $m-2$  and  $m-1$  degrees of freedom respectively. Second, for SO, we derive the asymptotic null distribution of the LRT, which is a chi-bar squared distribution. We further demonstrate that any distribution with  $p_{0k} = 1/2$  and  $p_{ik} = 0$ , for  $i = \pm 1, \pm 2, \dots, \pm(m-2)$ ,  $k = 1, 2$ , is least favorable. Thus,

$$\sup_{\mathcal{H}_0} \text{P}(\text{LRT} \geq t | \mathcal{H}_0) = \frac{1}{2} \text{P}(\chi_{2m-3}^2 \geq t) + \frac{1}{2} \text{P}(\chi_{2m-4}^2 \geq t). \quad (6)$$

### 3.1 Testing the orders in HL or LH association

Since SO-HL and SO-LH are parallel to each other, we only discuss the likelihood ratio test (LRT) for SO-LH. All our discussion can be extended to the SO-HL case. Consider two contingency tables  $\mathbf{C}_1$  and  $\mathbf{C}_2$  introduced earlier. We test the hypothesis  $\mathbf{C}_2 \succeq_{\text{S-LH}} \mathbf{C}_1$ . That is, we test if

$$\begin{aligned} q_{01}/q_{+1} &\leq q_{02}/q_{+2} \\ (1/q_{+1})(q_{01} + \sum_{i=1}^l q_{i1}) &\leq (1/q_{+2})(q_{02} + \sum_{i=1}^l q_{i2}) \quad \text{for } l = 1, 2, \dots, (m-1). \end{aligned} \quad (7)$$

Let us consider a new set of parameters, for  $k = 1$  and  $2$ ,

$$\mathbf{p}_k = (p_{0k}, p_{1k}, \dots, p_{(m-2)k}, p_{(m-1)k}),$$

where  $p_{dk} = q_{dk}/q_{k+}$  for  $d = 0, 1, 2, \dots, m-1$  and  $k = 1, 2$ . Also let

$$\mathbf{f}_k = (f_{0k}, f_{1k}, \dots, f_{(m-2)k}, f_{(m-1)k})$$

be the observations corresponding to  $\mathbf{p}_k$ . Then,  $\mathbf{f}_k$  has a multinomial distribution with probability  $\mathbf{p}_k$ , and testing (7) is equivalent to testing for the simple stochastic ordering between  $\mathbf{p}_1$  and  $\mathbf{p}_2$ . Therefore, using the results from Chapter 6 in Silvapulle and Sen (2005), asymptotically the LRT has a chi-bar squared distribution, and its least favorable distribution is the chi-bar squared distribution in (5).

### 3.2 Testing the order in global association

As with SO-HL and SO-LH, we introduce a new set of parameters

$$\mathbf{p}_k = (p_{0k}, p_{1k}, p_{2k}, \dots, p_{(2m-4)k}, p_{(2m-3)k})$$

where

$$\begin{aligned} p_{0k} &= q_{0k}, \\ p_{ik} &= q_{-i,k} && \text{for } i = 1, 2, \dots, m-2, \\ p_{(m-2+i)k} &= q_{ik} && \text{for } i = 1, 2, \dots, m-2, \\ p_{(2m-3)k} &= q_{-(m-1),k} + q_{(m-2)k} \end{aligned}$$

Using this notation, the stochastic order  $\succ_S$  in two contingency tables  $\mathbf{C}_1 = \{\pi_{ij}^{(1)}\}$  and  $\mathbf{C}_2 = \{\pi_{ij}^{(2)}\}$  is defined as

$$\begin{aligned} p_{01} &\leq p_{02} \\ p_{01} + \sum_{i=1}^l p_{i1} &\leq p_{02} + \sum_{i=1}^l p_{i2}, \quad \text{for } l = 1, 2, \dots, m-2, \end{aligned} \quad (8)$$

$$p_{01} + \sum_{i=1}^l p_{(m-2+i)1} \leq p_{02} + \sum_{i=1}^l p_{(m-2+i)2}, \quad \text{for } l = 1, \dots, m-2. \quad (9)$$

In other words, the order shows the simple partial stochastic ordering between  $\{p_{i1}, i = 1, 2, \dots, (2m-3)\}$  and  $\{p_{i2}, i = 1, 2, \dots, (2m-3)\}$ . We denote this as  $\mathbf{C}_2 \succ_S \mathbf{C}_1$ .

We now develop a LRT to test  $\mathcal{H}_1 : \mathbf{C}_1 \succ_S \mathbf{C}_2$  against  $\mathcal{H}_0 : \mathbf{C}_1 =_S \mathbf{C}_2$ . Let  $\{f_{i1}, i = 0, \dots, (2m-3)\}$  (resp.  $\{f_{i2}, i = 0, \dots, (2m-3)\}$ ) be the observed frequencies of multinomial distributions with size  $n$  and probabilities  $\{p_{i1}, i = 0, \dots, (2m-3)\}$  (resp.  $\{p_{i2}, i = 0, \dots, (2m-3)\}$ ) as defined earlier. The log-likelihood function is

$$l(\mathbf{p}_1, \mathbf{p}_2) = \sum_{k=1}^2 \left\{ \sum_{i=0}^{2m-4} f_{ik} \log p_{ik} + \left( n - \sum_{i=0}^{2m-4} f_{ik} \right) \log \left( 1 - \sum_{i=0}^{2m-4} p_{ik} \right) \right\}, \quad (10)$$

where  $\mathbf{p}_1 = (p_{01}, \dots, p_{(2m-4)1})^T$  and  $\mathbf{p}_2 = (p_{02}, \dots, p_{(2m-4)2})^T$ .

Suppose we let the set of parameters associated with the hypothesis  $\mathcal{H}_1$  (resp.  $\mathcal{H}_0$ ) as  $\mathcal{S}_1$  (resp.  $\mathcal{S}_0$ ).

$$\begin{aligned} \mathcal{S}_1 &= \left\{ \mathbf{p} : [B, -B] \mathbf{p} \geq 0_{(2m-3) \times 1} \right\} \\ \mathcal{S}_0 &= \left\{ \mathbf{p} : [B, -B] \mathbf{p} = 0_{(2m-3) \times 1} \right\}, \end{aligned}$$

where  $\mathbf{p} = (\mathbf{p}_1^T, \mathbf{p}_2^T)$ , and

$$B = \begin{pmatrix} 1 & 0_{1 \times (m-2)} & 0_{1 \times (m-2)} \\ 1_{(m-2) \times 1} & L_{m-2} & 0_{(m-2) \times (m-2)} \\ 1_{(m-2) \times 1} & 0_{(m-2) \times (m-2)} & L_{m-2} \end{pmatrix}.$$

Here,  $b_{l \times k}$  is a  $l \times k$  matrix whose elements are all  $bs$ , and  $L_k$  is a  $k \times k$  lower triangle matrix whose  $(i, j)$ th element  $l_{ij} = 1$  if  $i \geq j$  and 0 otherwise. We further let  $A = [B, -B]$  for notational simplicity.

The log-LRT is

$$T = 2 \left\{ \sup_{\mathbf{p} \in \mathcal{H}_0} l(\mathbf{p}) - \sup_{\mathbf{p} \in \mathcal{H}_1} l(\mathbf{p}) \right\} = 2 \left\{ l(\tilde{\mathbf{p}}) - l(\hat{\mathbf{p}}) \right\}, \quad (11)$$

where  $\tilde{\mathbf{p}}$  is the maximum likelihood estimate (MLE) of  $\mathbf{p}$  under  $\mathcal{H}_0$  and  $\hat{\mathbf{p}}$  is that under  $\mathcal{H}_1$ .

Now we examine the asymptotic null distribution of the log-LRT. From Proposition 4.3.1 in Silvapulle and Sen (2005) (also see Feng and Wang (2007) for details), it can be shown that, when testing  $\mathcal{H}_1$  against  $\mathcal{H}_0$ , the log-LRT (11) converges in distribution to

$$T_0 = \min_{\beta \in \mathcal{T}_0} (\mathbf{Z} - \beta)^T Q (\mathbf{Z} - \beta) - \min_{\beta \in \mathcal{T}_1} (\mathbf{Z} - \beta)^T Q (\mathbf{Z} - \beta), \quad (12)$$

where

$$\begin{aligned} \mathcal{T}_1 &= \left\{ \beta \in \mathcal{R}^{2(2m-3)} : [B, -B]\beta \geq 0_{(2m-3) \times 1} \right\} \\ \mathcal{T}_0 &= \left\{ \beta \in \mathcal{R}^{2(2m-3)} : [B, -B]\beta = 0_{(2m-3) \times 1} \right\}, \end{aligned}$$

and  $\mathbf{Z} \sim N\left(0_{2(2m-3) \times 1}, Q^{-1}\right)$  with

$$Q = \begin{pmatrix} (1/2)M & 0_{(2m-3) \times (2m-3)} \\ 0_{(2m-3) \times (2m-3)} & (1/2)M \end{pmatrix}$$

and

$$M = \text{diag}(1/p_0, 1/p_1, \dots, 1/p_{2m-4}) + (1/p_{2m-3})1_{(2m-3) \times (2m-3)}.$$

Thus, the asymptotic null distribution is the chi-bar square distribution

$$\lim_{n \rightarrow \infty} P(\text{LRT} \geq t | \mathcal{H}_0) = \sum_{i=0}^{2m-3} w_i(2m-3, A Q^{-1} A^T, \mathcal{R}_{(2m-3)}^+) P(\chi_i^2 \geq t), \quad (13)$$

where  $\mathcal{R}_{(2m-3)}^+$  is the positive quadrant of  $\mathcal{R}^{(2m-3)}$  and  $w_i(p, W, \mathcal{C})$ s are non-negative numbers whose sum equals 1. We refer readers to Chapter 3 in Silvapulle and Sen (2005) for more details on  $w_i(p, W, \mathcal{C})$ .



From Theorem 1, it can be shown that the least favorable null hypothesis is  $\mathbf{p}_1 = \mathbf{p}_2 = (1/2, 0_{1 \times (2m-4)})^T$ , and the least favorable asymptotic null distribution is a simple two component chi-bar square distribution. The proof is included in Appendix.

**Theorem 1.** In testing  $\mathcal{H}_1$  against  $\mathcal{H}_0$ , the least favorable null value for the asymptotic distribution of LRT is  $\mathbf{p}_1 = \mathbf{p}_2 = (1/2, 0_{1 \times (2m-4)})^T$ , and

$$\sup_{\mathcal{H}_0} \lim_{n \rightarrow \infty} P(\text{LRT} \geq t | \mathcal{H}_0) = \frac{1}{2} \left\{ P(\chi_{2m-4}^2 \geq t) + P(\chi_{2m-3}^2 \geq t) \right\}. \quad (14)$$

## 4 The power study

We numerically compare the power of the proposed LRT with that of the kappa-based test when testing the order in global association (SO) between two contingency tables. We use the Plackett copula to generate a contingency table. In the Plackett copula, a bivariate random variable  $(U, V)$  has the joint distribution function

$$f_{\theta}(u, v) = \frac{\{1 + (\theta - 1)(u + v)\} - \{(1 + (\theta - 1)(u + v))^2 - 4uv\theta(\theta - 1)\}^{1/2}}{2(\theta - 1)} \quad (15)$$

whose marginal distributions are uniform distributions on  $(0, 1)$ . The Plackett copula is specified by a parameter  $\theta > 0$  called the odd-ratio. The parameter  $\theta$  controls the sum of diagonal probabilities in the contingency table. When  $\theta = 1$ , the copula corresponds to the bivariate random variable  $(U, V)$  consisting of two independent components with uniform marginal distributions on  $(0, 1)$ . In the current study, we generate  $10 \times 10$  contingency tables with a fixed sum of diagonal probabilities  $p$  and size  $n$ . To do so, we first numerically find the Plackett parameter  $\theta$  which determines the given sum of diagonal probability  $p$ . Subsequently, we generate  $n$  copies of bivariate random variables  $(U, V)$  from (15) and distribute them uniformly. Interested readers should refer to Joe (1997) and Nelson (1999) for more details on the Plackett copula.

The power study is conducted at two levels. First, in Section 4.1, we compare the power of the proposed test (SO) to that of a kappa-based test (CO) for the case in which  $\mathbf{C}_1 \succ_S \mathbf{C}_2$

but  $\mathbf{C}_1 \succ_K \mathbf{C}_2$ . The second simulation study illustrates the difference between SO and CO. We generate a pair of tables  $\mathbf{C}_1$  and  $\mathbf{C}_2$  where  $\mathbf{C}_1 \succ_S \mathbf{C}_2$  but  $\mathbf{C}_1 =_K \mathbf{C}_2$ .

#### 4.1 The first simulation study

For  $k = 1, 2$ , let  $p_k$  be the probability that two discrete random variables  $X_k$  and  $Y_k$  in a contingency table  $\mathbf{C}_k$  are equal to each other. Also, let  $\Delta = p_2 - p_1$ . Suppose  $p_1 = 0.1, 0.3$ , and  $0.7$  and  $\Delta = 0, 0.05, 0.1$ , and  $0.2$ . We generate 1000 data sets with  $n = 30, 100, 300$ , and  $500$  for each  $(p_1, p_2)$ . We apply the three methods to compare the strength of bivariate association in  $\mathbf{C}_1$  and  $\mathbf{C}_2$ : (1) Stuart's test (ST), (2) Cohen's method (CO), and (3) the stochastic order test (SO) (see Table 1). In applying SO, we compute the power using the results in Theorem 1.

As can be seen from Table 1, our method is nearly always more powerful than ST. Indeed, ST performs poorer than both of its competitors in all cases considered in our analysis although its performance improves somewhat as the bivariate association in two tables becomes weaker.

Simulation results show that, when compared with CO, our method is generally superior when the bivariate association in two tables is weak. For example, our method is nearly always more powerful than CO when  $p_1 = 0.1$ . However, a closer examination of our results reveals that, as the strength of bivariate association in two tables increases, the power of our method becomes relatively weaker when compared with that of CO. As noted earlier, CO only considers the information on concordant and discordant pairs. In other words, it only compares the diagonal probabilities in two tables when testing for the difference in the strength of bivariate association in two contingency tables. On the other hand, the amount of information contained in diagonal probabilities increases as the strength of bivariate association in two tables increases. As a result, CO's statistical power also increases as  $p_1$  increases. In contrast, our method considers the information contained in both diagonal and off-diagonal probabilities. Accordingly, our test is inevitably more conservative than

CO. In order to reject the null hypothesis, there ought to be a significant difference between two tables not only in terms of their diagonal probabilities but also in terms of off-diagonal probabilities. Therefore, our test becomes especially conservative as the bivariate association in two tables becomes weaker since the amount of information contained in off-diagonal probabilities decreases.

$p_1 = 0.1$									
	$\Delta = 0$			$\Delta = 0.05$			$\Delta = 0.1$		
n	ST	SO	CO	ST	SO	CO	ST	SO	CO.
50	0.000	0.072	0.061	0.054	0.233	0.191	0.270	0.476	0.448
100	0.001	0.065	0.029	0.176	0.356	0.289	0.717	0.788	0.636
300	0.000	0.037	0.043	0.842	0.791	0.573	1.000	1.000	0.976
500	0.000	0.019	0.038	0.987	0.965	0.761	1.000	1.000	0.995
$p_1 = 0.2$									
	$\Delta = 0$			$\Delta = 0.05$			$\Delta = 0.1$		
n	ST	SO	CO	ST	SO	CO	ST	SO	CO
50	0.000	0.039	0.057	0.002	0.057	0.131	0.011	0.126	0.322
100	0.000	0.046	0.049	0.008	0.154	0.214	0.059	0.386	0.513
300	0.000	0.073	0.043	0.057	0.444	0.428	0.582	0.934	0.877
500	0.000	0.062	0.045	0.178	0.638	0.571	0.928	0.998	0.977
$p_1 = 0.3$									
	$\Delta = 0$			$\Delta = 0.05$			$\Delta = 0.1$		
n	ST	SO	CO	ST	SO	CO	ST	SO	CO
50	0.000	0.031	0.062	0.000	0.076	0.140	0.000	0.147	0.308
100	0.000	0.046	0.064	0.000	0.103	0.181	0.005	0.299	0.436
300	0.000	0.060	0.062	0.001	0.251	0.365	0.083	0.712	0.812
500	0.000	0.062	0.049	0.018	0.350	0.534	0.344	0.917	0.946
$p_1 = 0.7$									
	$\Delta = 0$			$\Delta = 0.05$			$\Delta = 0.1$		
n	ST	SO	CO	ST	SO	CO	ST	SO	CO
50	0.000	0.015	0.045	0.000	0.032	0.138	0.000	0.094	0.331
100	0.000	0.011	0.061	0.000	0.061	0.189	0.000	0.201	0.498
300	0.000	0.023	0.042	0.000	0.151	0.371	0.000	0.611	0.887
500	0.000	0.024	0.053	0.000	0.225	0.542	0.000	0.829	0.980

Table 1: First simulation: "ST" implies Stuart's method, "SO" implies the proposed stochastic order test, and "CO" implies the method based on Cohen's  $\kappa$  measure.

## 4.2 The second simulation study

The second simulation study illustrates the difference between SO and CO. More precisely, it compares the order based on the proposed multivariate summary measure and that based on Cohen's kappa statistic.

In doing so, let us first explain how we generate tables  $\mathbf{C}_1$  and  $\mathbf{C}_2$ . Let  $\mathbf{A}_1$  be the tables generated from the Plackett copular where the sum of diagonal probabilities is  $p$ . Also let  $\mathbf{A}_2$  be an independent table and  $\mathbf{A}_3(\eta)$  be a table in which either  $(i, i + \eta)$  or  $(i + \eta, i)$  is 1 for every possible  $i$ . For the given  $\eta$ , we generate tables  $\mathbf{C}_1$  and  $\mathbf{C}_2$  as follows:

$$\mathbf{C}_1 = 0.3\mathbf{A}_1 + 0.7\mathbf{A}_2$$

$$\mathbf{C}_2 = 0.3\mathbf{A}_1 + 0.6\mathbf{A}_2 + 0.1\mathbf{A}_3.$$

We generate 1000 pairs of  $10 \times 10$  tables  $(\mathbf{C}_1, \mathbf{C}_2)$  with size  $n = 50, 100, 300,$  and  $500$ . We again apply the three tests we examined earlier: (1) Stuart's test (ST), (2) the proposed stochastic order test (SO), and (3) the test based on Cohen's  $\kappa$  measure (CO). The results are reported in Table 2.

Table 2 shows that CO has an inadequate level of power when testing for the difference between  $\mathbf{C}_1$  and  $\mathbf{C}_2$ , because their  $\kappa$  measures are equal to one another. In contrast, SO is significantly more powerful than CO in nearly all the cases considered. The power of ST is similar to that in the first simulation.

## 5 Examples

This section illustrates the proposed stochastic order test through two examples. The first example illustrates SO-LH and SO-HL in terms of gender- and race-based inequality in returns to education. The second example tests the order in global association (SO) in left and right eyesight power among men and women. The eyesight power data came from Stuart (1953).

$\eta$	Method	n				n			
		50	100	300	500	50	100	300	500
		$p = 0.1$				$p = 0.275$			
0	ST	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000
	SO	0.057	0.055	0.019	0.024	0.065	0.038	0.030	0.029
	CO	0.055	0.055	0.058	0.051	0.052	0.049	0.051	0.047
1	ST	0.001	0.000	0.000	0.000	0.001	0.000	0.001	0.005
	SO	0.124	0.206	0.575	0.867	0.107	0.198	0.545	0.833
	CO	0.069	0.055	0.056	0.052	0.047	0.056	0.044	0.051
2	ST	0.000	0.001	0.005	0.004	0.002	0.003	0.005	0.008
	SO	0.155	0.217	0.636	0.910	0.132	0.225	0.635	0.889
	CO	0.058	0.048	0.044	0.051	0.059	0.062	0.061	0.042
8	ST	0.032	0.101	0.564	0.907	0.053	0.183	0.730	0.942
	SO	0.310	0.524	0.976	1.000	0.296	0.591	0.980	1.000
	CO	0.056	0.057	0.060	0.077	0.054	0.055	0.080	0.076
		$p = 0.45$				$p = 0.8$			
0	ST	0.000	0.001	0.001	0.000	0.000	0.000	0.003	0.001
	SO	0.051	0.059	0.047	0.033	0.051	0.061	0.036	0.052
	CO	0.064	0.048	0.054	0.049	0.050	0.054	0.052	0.047
1	ST	0.001	0.001	0.000	0.006	0.000	0.001	0.005	0.010
	SO	0.131	0.206	0.561	0.857	0.147	0.309	0.718	0.928
	CO	0.066	0.044	0.052	0.041	0.045	0.052	0.046	0.061
2	ST	0.004	0.004	0.010	0.017	0.003	0.008	0.014	0.024
	SO	0.110	0.252	0.669	0.913	0.174	0.319	0.737	0.956
	CO	0.050	0.058	0.054	0.046	0.062	0.048	0.061	0.046
8	ST	0.069	0.182	0.771	0.961	0.084	0.211	0.787	0.970
	SO	0.292	0.572	0.992	1.000	0.359	0.630	0.994	1.000
	CO	0.069	0.048	0.060	0.082	0.058	0.066	0.058	0.059

Table 2: The power of three tests in the second simulation study. "ST", "CO", and "SO" denote Stuart's method, the method based on Cohen's measure, and the proposed stochastic order test respectively.

## 5.1 Returns to education data

In this section, we examine the inequality in returns to education. Schooling is conceived as a process by which atomized individuals acquire knowledge or credentials. Traditional theories of stratification assume that education contributes to inequality by endowing people with different amounts of human capital (knowledge and skills) or credentials (*e.g.*, Becker (1975); Mincer (1974); Blau and Duncan (1967)). Alternatively, signaling theorists maintain

that education is instrumental in determining one's life chances because it functions as a criteria by which employers use to screen job applicants (Arrow, 1973; Spence, 1974; Blaug, 1985).

The literature on cost-benefit analysis of education however suggests that the monetary returns to education vary greatly across the more and the less privileged strata of the population. For example, many scholars have argued that returns to female education are low, because of the lower earnings, lower labor force participation, and shorter working hours of women, compared with men (Moreh, 1971; Morris and Ziderman, 1971; Psacharopoulos, 1973). Similarly, others have argued that returns to education for members of ethnic minority are low (Freeman, 1976).

This raises an important policy concern. If the rate of return to education varies greatly, it would support the view that the higher education of their children is a waste of money depending upon the profiles of their background. Higher education is an investment for a certain subgroup of the population, whereas it is largely consumption for the others. This suggests that education cannot be a vehicle for accomplishing economic equality between the more and the less privileged strata of the population.

Our data come from the 2004 American National Election Studies (ANES) Survey. The survey was conducted as part of the ANES time series (which dates back to 1948) and asked about Americans' basic political beliefs, attitudes, and behaviors. The 2004 ANES sample consisted of a new cross-section of respondents that yielded 1,212 face-to-face interviews in the pre-election study prior to the 2004 presidential election. Pre-election interviews were conducted September 7 through November 1, 2004.

As described earlier, there are two scenarios under which the association between education and income can be weakened. The presence of highly educated individuals whose income levels are low (the high-low's) weakens the association between education and income. Alternatively, the presence of poorly educated individuals with high income (the low-high's) also weakens the association between education and income. In this section, we test the group

difference between men and women and between whites and members of ethnic minorities in both the "high-low" and the "low-high" dimensions. These results are presented in Table 5.1.

Our results show that the gender difference was significant when considering the "high-low" cases (or the probability of the poorly educated attaining high levels of income). The null hypothesis was clearly rejected ( $p = 0.0188$ ). On other hand, men and women were not vastly different in terms of the "low-high" dimension (or being highly educated and attaining low levels of income) ( $p = 0.1811$ ). In other words, there was a discrepancy between men and women in terms of returns to education; however, this discrepancy was asymmetric. Both men and women had to be well educated to attain high levels of income; but, compared with women, the poorly educated men still had a descent chance of attaining high levels of income.

A similar pattern emerged when examining the group difference across racial lines. The probability of the highly educated attaining high levels of income were not vastly different for whites and others ( $p = 0.3961$ ). On the other hand, whites had a significantly higher chance of attaining high levels of income when poorly educated compared with members of ethnic minorities ( $p = 0.0370$ ).

## 5.2 Eye-sight power Data

As our second example, we apply our method to the eye-testing data previously examined in Stuart (1953). The data consist of eye test records of 3242 male and 7477 female employees in Royal Ordnance factories in 1943-6. From the data set, Stuart (1953) tests the strength of association in the power of left and right eyes. The data set is also analyzed in Bartolucci and Scaccia (2004). Our null and alternative hypotheses are as follows:

$\mathcal{H}_0$  : The strength of association between left and right eyesight powers is equal for men and women.

$\mathcal{H}_1$  : The strength of association between left and right eyesight powers is unequal for men and women.

$\mathcal{H}_1$			LRT	p-value
male	$\succeq_{\text{SO-LH}}$	female	6.9557	0.1811
female	$\succeq_{\text{SO-HL}}$	male	12.8021	0.0188
white	$\succeq_{\text{SO-LH}}$	minority	4.6219	0.3961
minority	$\succeq_{\text{SO-HL}}$	white	11.1200	0.0370

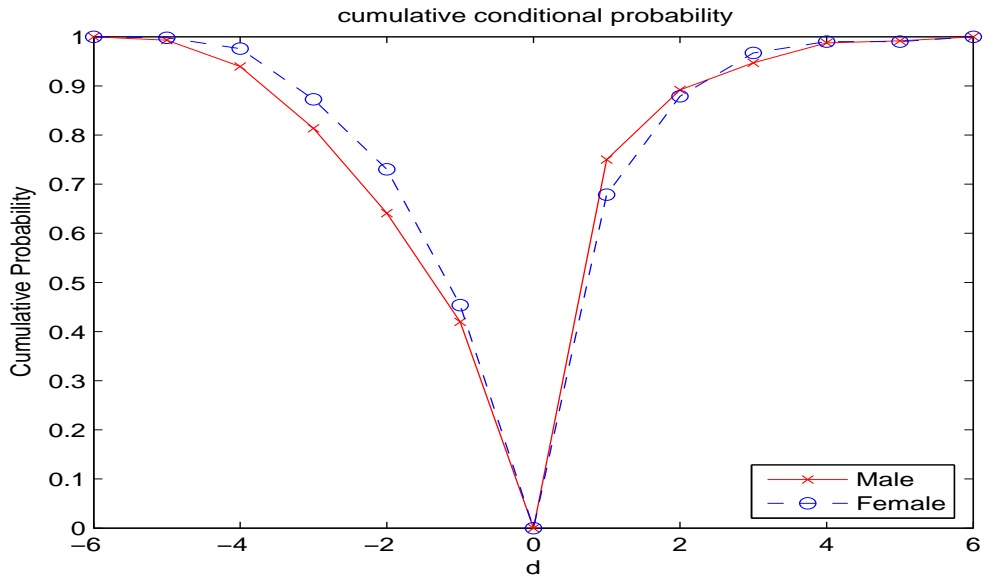
Table 3: If "male  $\succeq_{\text{SO-LH}}$  female", the poorly educated men often obtain higher levels of income when compared with the poorly educated women. If "female  $\succeq_{\text{SO-HL}}$  male", the highly educated women often obtain lower levels of income when compared with the highly educated men. Orders between white and minority are defined similarly.

When applying the three methods we examined earlier (e.g. ST, SO, and CO), p-values are 0.4609,  $< 0.0001$ , and 0.0564 respectively. Thus, we cannot not reject the null hypothesis  $\mathcal{H}_0$  based on ST and CO, where SO allows us to reject the null hypothesis. These results are consistent with those concerning the global log odds ratio in Bartolucci and Scaccia (2004). This is sensible since our method is comparable with the authors' test on global odds ratio (not one on reverse continuation).

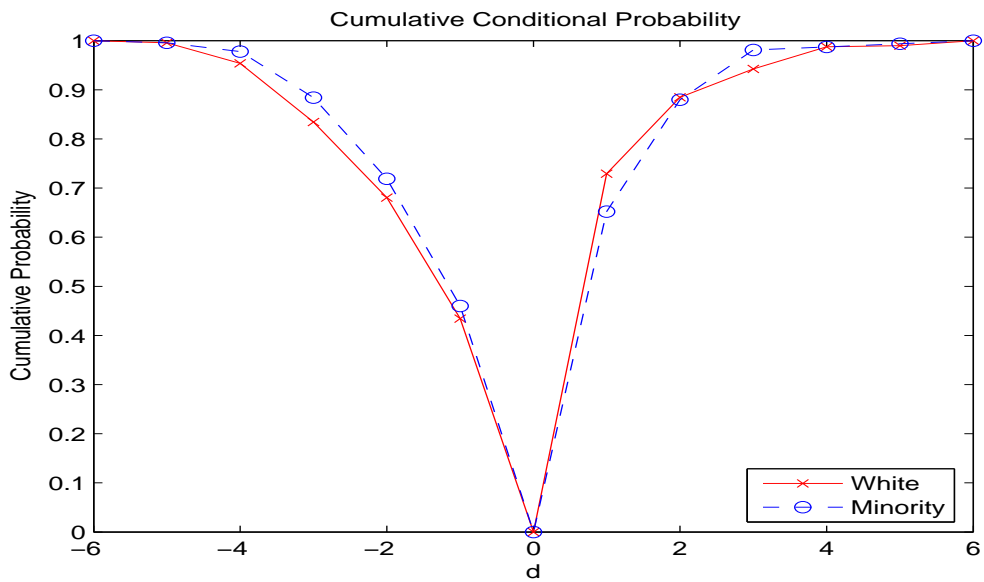
## 6 Conclusion

In this paper, we introduced a new multivariate summary statistic of association between two ordered categorical variables in a contingency table. This summary statistic allows the stochastic ordering of two tables in accordance with the strength of association. We showed that the the proposed order is closely related to the order based on Cohen's  $\kappa$  measure. We developed a likelihood ratio test examining the proposed stochastic order. We derived its asymptotic null distribution and the least favorable distribution to attain the largest  $p$ -values in the composite null space. The power study shows the proposed test has several





(a) men versus women



(b) whites versus ethnic minorities

Figure 1: Conditional cumulative probabilities for SO-LH (left wing) and SO-HL (right wing) defined in (2) and (3)

advantages when comparing the strength of association in two contingency tables. Finally, we provided two examples illustrating our method.

## Appendix: Proof of Theorem 1

For  $\mathbf{Z}$  and  $\beta$  defined in Section 2.1, we consider a linear transformation

$$\begin{pmatrix} \mathbf{Y} \\ \mathbf{Y}^\perp \end{pmatrix} = \begin{pmatrix} A \\ A^\perp \end{pmatrix} \mathbf{Z} = \begin{pmatrix} B & -B \\ A_1^\perp & A_2^\perp \end{pmatrix} \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{pmatrix}. \quad (16)$$

where  $A = (B, -B)$  is the matrix to represent the null hypothesis, and  $A^\perp = (A_1^\perp, A_2^\perp)$  is a  $(2m-3) \times 2(2m-3)$  matrix whose column vectors are orthogonal to those in  $A$ . The random vectors  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  are the first and the second  $(2m-1)$  elements of  $\mathbf{Z}$ , which are independent to each other and are identically distributed from  $N(0_{(2m-3) \times 1}, 2M^{-1})$ . Consider the linear transformation

$$\begin{pmatrix} \Theta \\ \Theta^\perp \end{pmatrix} = \begin{pmatrix} B & -B \\ A_1^\perp & A_2^\perp \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad (17)$$

where again  $\beta_1$  (resp.  $\beta_2$ ) is the first  $(2m-3)$  elements (resp. the second  $(2m-3)$  elements) of  $\beta$ . Then, with transformed variables and parameters, the LRT is

$$\text{LRT} = \mathbf{Y}^T V^{-1} \mathbf{Y} - \min_{\Theta \geq 0} (\mathbf{Y} - \Theta)^T V^{-1} (\mathbf{Y} - \Theta),$$

where  $V = BM^{-1}B^T$  and  $\mathbf{Y} \sim N(0, V)$ .

Now we look for the Cholesky decomposition of  $V$  to get the dual of chi-bar square distribution, which allows us to find a least favorable null hypothesis. Suppose  $\mathcal{L}_{M^{-1}}$  is the lower triangular matrix in the Cholesky decomposition of  $M$ , say  $M = \mathcal{L}_{M^{-1}} \mathcal{L}_{M^{-1}}^T$ . Then, since the matrix  $B$  is lower triangular, thus, the Cholesky factor of the covariance matrix  $V$  is

$$\mathcal{L}_V = B \mathcal{L}_{M^{-1}}. \quad (18)$$

Here,  $M^{-1}$  is the variance covariance matrix of a multinomial distribution with probability  $(p_0, \dots, p_{2m-4})$ , and it is defined as

$$M^{-1} = \text{diag}(p_0, \dots, p_{2m-4}) - (p_0, \dots, p_{(2m-4)})^T (p_0, \dots, p_{(2m-4)}),$$

where  $p_i = p_{i1} = p_{i2}$ , for  $i = 0, \dots, 2m-4$ . Its Cholesky factor is given in Tanabe and Sagae (1992) as

$$\mathcal{L}_{M^{-1}} = L \cdot \text{diag}(d_0^{1/2}, \dots, d_{2m-4}^{1/2}), \quad (19)$$

where, if we let

$$q_i = 1 - (p_0 + p_1 + \dots + p_i), \quad i = 0, 1, 2, \dots, 2m - 4,$$

and  $q_{-1} = 1$ , then  $L$  is a  $(2m - 3) \times (2m - 3)$  lower triangular matrix whose  $(i, j)$ th element is

$$l_{ij} = \begin{cases} p_{i-1}/q_{j-1} & i > j \\ 1 & i = j \\ 0 & i < j \end{cases} \quad (20)$$

and  $d_i = (p_i q_i)/q_{i-1}$  for  $i = 0, \dots, 2m - 4$ . Thus,

$$\begin{aligned} \mathcal{L}_V &= B\mathcal{L}_{M-1} \\ &= \begin{pmatrix} \sqrt{d_0} & 0 & 0 & 0 & 0 & 0 \\ \sqrt{d_0} \left\{ 1 + \frac{p_1}{q_0} \right\} & \sqrt{d_1} & 0 & 0 & 0 & 0 \\ \sqrt{d_0} \left\{ 1 + \frac{p_1+p_2}{q_0} \right\} & \sqrt{d_1} \left\{ 1 + \frac{p_2}{q_1} \right\} & \sqrt{d_2} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ \sqrt{d_0} \left\{ 1 + \frac{z_1}{q_0} \right\} & \sqrt{d_1} \left\{ 1 + \frac{z_2}{q_1} \right\} & \dots & \dots & \dots & \sqrt{d_{(2m-4)}} \end{pmatrix}, \end{aligned}$$

where  $z_j = \sum_{i=j}^{2m-4} p_i$  for  $j = 1, 2, \dots, 2m - 4$ . Since the column vectors of  $\mathcal{L}_V$  are all non-negative with non-zero first elements,

$$\mathcal{R}_{2m-3}^+ \subseteq \mathcal{C}(\mathcal{L}_V) \subseteq \mathcal{C}(e^T),$$

where  $\mathcal{C}(R) = \{\Theta : R^T \beta \geq 0\}$  and  $e = (1, 0_{1 \times (2m-4)})^T$ .

We finally claim that  $\mathcal{C}(\mathcal{L}_V)$  is not a proper subset of  $\mathcal{C}(e^T)$  by choosing a sequence in  $\mathcal{H}_0$  whose  $\mathcal{C}(\mathcal{L}_V)$  converges to  $\mathcal{C}(e^T)$ . If we let  $\mathbf{p}_1(\epsilon) = 2^{-1}(1 - (2m - 4)\epsilon, \epsilon \cdot \mathbf{1}_{1 \times (2m-4)})$ , then  $\mathbf{p}_1(\epsilon)$  converges to  $2^{-1}(1, 0, \dots, 0)$ . Hence,  $\mathcal{L}_V$  converges to

$$\frac{1}{2}(\mathbf{1}_{(2m-3) \times 1}, 0_{(2m-3) \times (2m-4)}),$$

and, therefore,  $\mathcal{C}(\mathcal{L}_V)$  converges to  $\mathcal{C}(e^T)$ .

## References

Arrow, K. J. (1973). *Information and Economic Behavior*. Stockholm: Federation of Swedish Industries.

- Barlow, W. (1996). Measurement of interrater agreement with adjustment for covariates. *Biometrics*, 52:659–702.
- Bartolucci, F. and Scaccia, L. (2004). Testing for positive association in contingency tables with fixed margins. *Computational Statistics and Data Analysis*, 47:195–210.
- Becker, G. (1975). *Human Capital. 2nd ed.* New York: National Bureau of Economic Research.
- Blau, P. M. and Duncan, O. D. (1967). *The American Occupational Structure.* New York: Wiley.
- Blaug, M. (1985). Where are we now in the economics of education? *Economics of Education Review*, 4:17–28.
- Cohen, J. (1960). A coefficient of agreement for normal scales. *Educational and Psychological Measurement*, 20:37–46.
- Feng, Y. and Wang, J. (2007). Likelihood ratio test against simple stochastic ordering among several multinomial populations. *Journal of Statistical Planning and Inference*, 137:1362–1374.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76:378–382.
- Fleiss, J. L. and Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33:613–619.
- Fleiss, J. L. and Cuzick, J. (1979). The reliability of dichotomous judgments: unequal numbers of judges per subject. *Applied Psychological Measurement*, 3:537–542.
- Freeman, R. (1976). *The Overeducated American.* New York: Academic Press.

- Goodman, L. A. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, 74:537–552.
- Goodman, L. A. (1985). The analysis of cross-classified data having ordered and/or un-ordered categorized: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries. *Annals of Statistics*, 13:10–69.
- Hand, D., Daly, F., Lunn, A., McConway, K., and Ostrowski, E. (1994). *A Handbook of Small Data Sets*. London: Chapman and Hall.
- Landis, J. R. and Koch, G. G. (1977a). The measurement of observer agreement for categorical data. *Biometrics*, 33:159–174.
- Landis, J. R. and Koch, G. G. (1977b). A one-way components of variance model for categorical data. *Biometrics*, 33:671–679.
- Lau, T. (1993). High-order kappa-type statistics for a dichotomous attribute in multiple ratings. *Biometrics*, 49:535–542.
- Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalization and alternatives. *Psychological Bulletin*, 76:365–377.
- Liu, I. and Agresti, A. (2005). The analysis of ordered categorical data: An overview and a survey of recent developments. *Test*, 14:1–73.
- Loughin, T. M. and Scherer, P. N. (1998). Testing for general association in contingency tables with multiple column responses. *Biometrics*, 54:630–637.
- Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: Columbia University Press.
- Moreh, J. (1971). Human capital and economic growth: United kingdom, 1951-1961. *Economic and Social Review*, October.

- Morris, V. and Ziderman, A. (1971). The economic return on investment in higher education in england and wales. *Economic Trends*, 211:x–xxmi.
- Psacharopoulos, G. (1973). *Returns to Education: An International Comparison*. Amsterdam: Elsevier.
- Silvapulle, M. and Sen, P. (2005). *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions*. New York: Wiley.
- Spence, A. M. (1974). *Market Signaling: Informational Transfer in Hiring and Related Screening Process*. Cambridge:Harvard University Press.
- Stuart, A. (1953). The estimation and comparison of strengths of association in contingency tables. *Biometrika*, 40:105–110.
- Tomizawa, S. Miyamoto, N. and Hatanaka, Y. (2001). Measure of symmetry for square contingency tables having ordered categories. *The Australian and New Zealand Journal of Statistics*, 43:335–349.
- Tomizawa, S. and Tokunaga, S. (2006). Incomplete symmetry and conditional symmetry models for square contingency tables with ordered categories. *Far East Journal of Statistics*, 19:33–42.
- Umesh, U. N. (1995). Predicting nominal variable relationships with multiple response. *Journal of Forecasting*, 14:585–596.