Estimation and Testing in
Large Binary Contingency Tables

W. C. M. KALLENBERG

University of Twente, Enschede, The Netherlands

Communicated by the Editors

Very sparse contingency tables with a multiplicative structure are studied. The number of unspecified parameters and the number of cells are growing with the number of observations. Consistency and asymptotic normality of natural estimators are established. Also uniform convergence of the estimators to the parameters is investigated, and an application to the construction of confidence intervals is presented. Further, a family of goodness-of-fit tests is proposed for testing multiplicativity. It is shown that the test statistics are asymptotically normal. The results can be applied in such different fields as production testing or psychometrics.


1. INTRODUCTION

Consider a sequence \( \{X_{ij}\}, i = 1, \ldots, I, j = 1, \ldots, J \) of independent r.v.’s with distribution

\[
P(X_{ij} = 1) = p_{ij} = 1 - P(X_{ij} = 0).
\]  

(1.1)

The observations may be displayed in a very sparse contingency table, with \( X_{ij} \) the entry of the \((i, j)\)th cell. The model will further be restricted to the multiplicative structure describing independence between rows and columns,

\[
p_{ij} = \alpha_i \beta_j \quad i = 1, \ldots, I, j = 1, \ldots, J,
\]  

(1.2)

where \( 0 < \alpha_i < 1 \) and \( 0 < \beta_j < 1 \) are unknown probabilities. It is supposed that \( I \) and \( J \) are large, implying that we have also a large number of unspecified parameters \( \alpha_i, \beta_j \) involved. Note the difference between the
present multiplicative model with only one observation per cell and the more familiar log-linear models, where generally speaking the expected values of the cell counts are required to be not too small or only a few of them are tolerated to be small, cf. Cox [2], Bishop, Fienberg, and Holland [11, and references there.

It is immediately seen from (1.2) that there is a problem of identifiability. If $c$ is a constant close to one, $\alpha_i^* = c\alpha_i$ and $\beta_j^* = c^{-1}\beta_j$ lead to the same product, i.e., $\alpha_i^*\beta_j^* = \alpha_i\beta_j$, and both $\alpha_i^*$ and $\beta_j^*$ are again probabilities if $c \neq 1$ is suitably chosen. However, we are only interested in the product $\alpha_i\beta_j$, thus the problem of identifiability does not bother us. (We may for instance put $\alpha_1 = \frac{1}{2}$ without loss of generality and the problem of identifiability is eliminated.)

An example of the preceding situation is met in the context of production testing of an "IC chip" in the integrated circuit industry. Many IC chips are produced simultaneously, say $J$, as they are formed on the silicon wafer and as they go through the process steps. All $J$ chips are subjected to the same treatments, but they have different probabilities $\beta_j$ of being functional, depending upon their positions on the silicon wafer. Some of the treatments are performed separately for each silicon wafer, but sometimes wafers are handled in groups or batches. A batch is constituted by $I$ wafers and each wafer has a probability $\alpha_i$ of being properly treated. The last production step consists in testing whether the IC chip is correctly functioning or not. Noting 1 for functionality we have the model given by (1.1) and (1.2). A detailed discussion of this model has been given by Pesotchinsky [9].

There are many other potential applications, for example in psychometrics. At first sight the Rasch model

$$p_{ij} = \frac{\text{e}^{\theta_j}}{1 + \text{e}^{\theta_j}}$$

may look similar to (1.2). However, in the Rasch model

$$X_{i+} = \sum_{i=1}^{J} X_{i\cdot}, \quad i = 1, \ldots, I,$$  \hspace{1cm} (1.3)

is the sufficient statistic for $\epsilon_i$, and

$$X_{+j} = \sum_{k=1}^{I} X_{kj}, \quad j = 1, \ldots, J,$$  \hspace{1cm} (1.4)

is the sufficient statistic for $\theta_j$. Here we do not have such simple sufficient statistics. The lack of sufficiency shows also the difference between the present model and classical log-linear models.
Although it is possible to derive maximum likelihood estimators, minimum distance estimators or Pitman-type estimators, we investigate the following far more simple and easy to calculate natural estimator of \( \alpha_i \beta_j \). Define

\[
X_{++} = \sum_{k=1}^{I} \sum_{l=1}^{J} X_{kl}.
\] (1.5)

For fixed \( i \in \{1, \ldots, I\} \) and \( j \in \{1, \ldots, J\} \) the estimator \( T_{ij} \) of \( \alpha_i \beta_j \) is now defined by

\[
T_{ij} = \frac{X_{i+} X_{+j}}{X_{++}}
\] (1.6)

if \( X_{++} \neq 0 \), define \( T_{ij} = 0 \) (cf. also (5") on p. 1265 of Pesotchinsky [9], but note that \( T_{ij} \) is not an unbiased estimator of \( \alpha_i \beta_j \)). Denoting

\[
\alpha_+ = \sum_{k=1}^{I} \alpha_k \quad \text{and} \quad \beta_+ = \sum_{l=1}^{J} \beta_l
\] (1.7)

the idea is that \( X_{i+} \) estimates \( \alpha_i \beta_+ \), that \( X_{+j} \) estimates \( \alpha_+ \beta_j \) and that \( X_{++} \) estimates \( \alpha_+ \beta_+ \). Hence the ratio in (1.6) serves as an estimator of \( \alpha_i \beta_j \). In Section 2 consistency and asymptotic normality of this estimator is established. Further the uniform convergence of \( T_{ij} \) to \( \alpha_i \beta_j \) is studied and an application to the construction of confidence intervals is presented. It is always open to question whether a multiplicative probability model properly describes the data at hand. In Section 3 a family of goodness-of-fit tests is investigated based on directed divergence measures, introduced by Cressie and Read [3]. The main problem here is the large number of nuisance parameters, since \( I \) and \( J \) are supposed to be large. At first sight one may hope that the influence of the estimators on the asymptotic null distribution of the Cressie–Read family of test statistics is negligible as in testing the fit of i.i.d. observations with distribution function \( F_{\theta} \), where \( \theta \) is a location-scale nuisance parameter, when the number of classes increases with the number of observations, cf. Drost [4]. However, here the number of nuisance parameters is so large, that the estimators do have a serious influence on the limiting distribution. Therefore we have to modify the Cressie–Read family to correct for a bias term and to adjust the asymptotic variance. Nevertheless it can be shown that the modified test statistics are asymptotically standard normal under the null hypothesis of multiplicativity. Related results in a different context are given in Haberman [6] and Koehler [8].
2. Estimation

As mentioned in Section 1, the unknown parameters $\alpha_i$ and $\beta_j$ themselves are not identifiable. We restrict attention to the estimation of the product $\alpha_i\beta_j$ or $\alpha_i\beta_+$, etc., which are the natural parameters in the model. For estimation of $\alpha_i\beta_j$ we apply the estimator $T_{ij}$ given by (1.6).

This estimator may be modified a little bit to ensure that its values are in $[0, 1]$, taking into account that $\alpha_i$ and $\beta_j$ are probabilities. This may be done by using

$$\tilde{T}_{ij} = \min(T_{ij}, 1)$$

as an estimator of $\alpha_i\beta_j$. Under very weak regularity conditions both estimators $\tilde{T}_{ij}$ and $T_{ij}$ are asymptotically equivalent.

Another modification of (1.6) is worth mentioning. Replace in (1.6) $X_{i+}$ by

$$X_{i+}^{(i)} = X_{i+} - X_{ij},$$

$X_{+j}$ by

$$X_{+j}^{(i)} = X_{+j} - X_{ij},$$

and $X_{++}$ by

$$X_{++}^{(i)} = X_{++} - X_{+j} - X_{i+} + X_{ij}.$$

The attraction of this modification is of mathematical nature and lies in the independence of the statistics $X_{i+}^{(i)}$, $X_{+j}^{(i)}$, $X_{++}^{(i)}$, and $X_{ij}$. The results of this paper continue to hold for this modified estimator, cf. however, Remark 3.3 and Remark A.1.

The first result of this section describes the consistency of the estimators for each fixed $i, j$ as $I, J \to \infty$.

**Theorem 2.1.** Let $i, j$ be fixed.

(i) If $\beta_+ \to \infty$ as $J \to \infty$, then

$$\frac{X_{i+}}{\beta_+} \to P \alpha_i \quad \text{as} \quad J \to \infty. \tag{2.1}$$

(ii) If $\alpha_+ \to \infty$ as $I \to \infty$, then

$$\frac{X_{+j}}{\alpha_+} \to P \beta_j \quad \text{as} \quad I \to \infty. \tag{2.2}$$
(iii) If $\alpha_+ \beta_+ \to \infty$ as $IJ \to \infty$, then
$$\frac{X_{++}}{\alpha_+ \beta_+} \xrightarrow{P} 1 \quad \text{as} \quad IJ \to \infty. \quad (2.3)$$

(iv) If $\alpha_+ \to \infty$ as $I \to \infty$ and $\beta_+ \to \infty$ as $J \to \infty$, then
$$T_{ij} \xrightarrow{P} \alpha_i \beta_j \quad \text{as} \quad I \to \infty \text{ and } J \to \infty. \quad (2.4)$$

**Proof.** (i) For each $\varepsilon > 0$ we have by Chebyshev's inequality
$$P \left( \left| \frac{X_{++}}{\beta_+} - \alpha_+ \right| > \varepsilon \right) \leq \varepsilon^{-2} \beta_+^{-2} \sum_{i=1}^I \alpha_i \beta_i (1 - \alpha_i \beta_i) \leq \varepsilon^{-2} \beta_+^{-1} \to 0 \quad \text{as} \quad J \to \infty.$$

(ii) and (iii) are proved in a similar way. Combination of (i)-(iii) yields (iv). □

By Theorem 2.1(iv) we see that for each fixed $i$, $j$ we can estimate $\alpha_i \beta_j$ consistently if both $I$ and $J \to \infty$ and both $\alpha_+$ and $\beta_+ \to \infty$. Since usually the $\alpha$'s and $\beta$'s stay away from zero, the latter will mostly be the case. Although in general $\alpha_i$ is not identifiable, it is seen from (2.1) that if $\beta_+$ is known and hence $\alpha_i$ is identifiable, $\alpha_i$ can be estimated consistently. The same argument applies to $\beta_j$. Note that for (2.3) to hold it is not necessary that both $I$ and $J$ tend to infinity.

Next *asymptotic normality* is discussed. Define for fixed $i$, $j$,
$$v_i^2(\alpha_i) = \sum_{l=1}^J \alpha_i \beta_i (1 - \alpha_i \beta_i)$$
$$v_j^2(\beta_j) = \sum_{k=1}^I \alpha_k \beta_j (1 - \alpha_k \beta_j)$$
$$v_2 = \sum_{k=1}^I \sum_{l=1}^J \alpha_k \beta_l (1 - \alpha_k \beta_l)$$
$$v_2(\alpha_i, \beta_j) = \alpha_i^2 \beta_+^2 v_2^2(\beta_j) + \beta_j^2 \alpha_+^2 v_2^2(\alpha_i). \quad (2.5)$$

The following inequalities are immediate
$$\alpha_i (1 - \alpha_i) \beta_+ \leq v_i^2(\alpha_i) \leq \alpha_i \beta_+$$
$$\beta_j (1 - \beta_j) \alpha_+ \leq v_j^2(\beta_j) \leq \beta_j \alpha_+$$
$$\sum_{k=1}^I \alpha_k (1 - \alpha_k) \beta_+ \leq v_2 \leq \alpha_+ \beta_+$$
$$\sum_{l=1}^J \beta_l (1 - \beta_l) \alpha_+ \leq v_2 \leq \alpha_+ \beta_+ \quad (2.6)$$
and hence
\[
\begin{align*}
    v_1^2(\alpha_i) &\to \infty \Leftrightarrow \beta_+ \to \infty \\
    v_2^2(\beta_j) &\to \infty \Leftrightarrow \beta_+ \to \infty \\
    v_3^2 &\to \infty \Leftrightarrow \alpha_+ \beta_+ \to \infty \\
    v_4^2(\alpha_i, \beta_j) &\to 0 \Leftrightarrow \alpha_+ \to \infty, \beta_+ \to \infty.
\end{align*}
\] (2.7)

These results are used in the proof of the following theorem.

**Theorem 2.2.** Let i, j be fixed.

(i) If \( \beta_+ \to \infty \) as \( J \to \infty \), then
\[
\frac{X_{i+} - \alpha_i \beta_+}{v_1(\alpha_i)} \xrightarrow{D} N(0, 1) \quad \text{as} \quad J \to \infty.
\]

(ii) If \( \alpha_+ \to \infty \) as \( I \to \infty \), then
\[
\frac{X_{i+j} - \alpha_+ \beta_j}{v_2(\beta_j)} \xrightarrow{D} N(0, 1) \quad \text{as} \quad I \to \infty.
\]

(iii) If \( \alpha_+ \beta_+ \to \infty \) as \( IJ \to \infty \), then
\[
\frac{X_{i+j} - \alpha_+ \beta_+}{v_3} \xrightarrow{D} N(0, 1) \quad \text{as} \quad IJ \to \infty.
\]

(iv) If \( \alpha_+ \to \infty, \beta_+ \to \infty \) as \( I \) and \( J \to \infty \), then
\[
\frac{T_{ij} - \alpha_i \beta_j}{v_4(\alpha_i, \beta_j)} \xrightarrow{D} N(0, 1) \quad \text{as} \quad I \text{ and } J \to \infty.
\]

**Proof.** (i)–(iii). Direct application of the Lindeberg–Feller Central Limit Theorem, cf. Feller [5, p. 519], using the relations (2.7).

(iv) Write
\[
\frac{T_{ij} - \alpha_i \beta_j}{v_4(\alpha_i, \beta_j)} = T_{ij} \left( \frac{\alpha_+ \beta_+ - X_{i+j}}{v_3} \right) \frac{v_3}{v_4(\alpha_i, \beta_j) \alpha_+ \beta_+}
\]
\[
+ v_4^{-1}(\alpha_i, \beta_j) \left\{ \frac{X_{i+j} - \alpha_+ \beta_+}{\alpha_+ \beta_+ - \alpha_i \beta_j} \right\}.
\] (2.8)

Since
\[
T_{ij} \xrightarrow{p} \alpha_i \beta_j, \quad \frac{\alpha_+ \beta_+ - X_{i+j}}{v_3} = O_p(1)
\]
and

\[
\frac{v_3}{v_4(a_i, b_j) \alpha_+ \beta_+} \leq \frac{(\alpha_+ \beta_+)^{1/2}}{\alpha_i \alpha_+^{-1} \beta_+ (1 - \beta_j) \alpha_+^{1/2} \alpha_+ \beta_+} = \frac{1}{\alpha_i \beta_+^{1/2} (1 - \beta_j)^{1/2} \beta_+^{1/2}} \to 0.
\]

It follows that the first term on the right-hand side of (2.8) tends to zero in probability as \( I \) and \( J \to \infty \). The second term equals

\[
v_4^{-1}(a_i, b_j) \left\{ \left( \frac{X_i^{(j)} - \alpha_i \beta_+}{\beta_+} \right) \beta_j + \frac{(X_i^{(j)} - \alpha_i \beta_+)}{\alpha_+} \alpha_i \right\} + v_4^{-1}(a_i, b_j) \left\{ X_{ij} \left( \frac{\beta_j}{\beta_+} + \frac{\alpha_i}{\alpha_+} \right) + \frac{(X_{i+j} - \alpha_i \beta_+) (X_{i+j} - \alpha_+ \beta_j)}{\alpha_+ \beta_+} \right\}.
\]

(2.9)

In view of (i) and (ii) the second term of (2.9) equals

\[
v_4^{-1}(a_i, b_j) O_p(\beta_+^{-1} + \alpha_+^{-1} + v_1(a_i) v_2(b_j) \alpha_+^{-1} \beta_+^{-1}) = O_p(1).
\]

It is easily seen that in (i) and (ii) \( X_{i+} \) and \( X_{+j} \) may be replaced by \( X_{i+}^{(j)} \) and \( X_{+j}^{(i)} \), respectively. Therefore, using the independence of \( X_{i+}^{(j)} \) and \( X_{+j}^{(i)} \), it follows that the first term of (2.9) converges in distribution to an \( N(0, 1) \)-distribution. This completes the proof of Theorem 2.2.

So far we investigated the behaviour of \( T_{ij} \) for fixed \( i, j \). For several applications uniform convergence of \( T_{ij} \) to \( \alpha_i \beta_j \) is needed. To obtain uniformity we use Hoeffding's inequality.

PROPOSITION 2.3 [7]. Let \( Y_1, \ldots, Y_n \) be independent r.v.'s with \( 0 \leq Y_i \leq 1 \) and \( \mathbb{E}Y_i = p_i \), for \( i = 1, \ldots, n \). Then for every \( x > 0 \),

\[
P \left( \sum_{i=1}^n Y_i - \sum_{i=1}^n p_i \geq nx \right) \leq \exp(-2nx^2)
\]

(2.10)

and hence

\[
P \left( \left| \sum_{i=1}^n Y_i - \sum_{i=1}^n p_i \right| \geq nx \right) \leq 2 \exp(-2nx^2).
\]

(2.11)

As a consequence we have the following lemma.
Lemma 2.4. For each $\delta > 0$ we have

$$P \left( \max_{1 \leq i \leq J} \left| \frac{X_{+i}}{\beta_+} - \alpha_i \right| > \delta \right) \leq 1 - \left\{ 1 - 2 \exp(-2\delta^2 \beta_+^2 J^{-1}) \right\}^J \tag{2.12}$$

and

$$P \left( \max_{1 \leq i \leq J} \left| \frac{X_{+i}}{\alpha_+} - \beta_j \right| > \delta \right) \leq 1 - \left\{ 1 - 2 \exp(-2\delta^2 \alpha_+^2 I^{-1}) \right\}^J \tag{2.13}$$

and

$$P \left( \left| \frac{X_{+i}}{\alpha_+} + 1 \right| > \delta \right) \leq 2 \exp(-2\delta^2 \alpha_+^2 \beta_+^2 I^{-1} J^{-1}) \tag{2.14}$$

Proof. If $U_1, \ldots, U_n$ are independent, then

$$P \left( \max_{1 \leq i \leq n} |U_i| > \delta \right) = 1 - P \left( \max_{1 \leq i \leq n} |U_i| \leq \delta \right)$$

$$= 1 - \prod_{i=1}^n P(|U_i| \leq \delta) = 1 - \prod_{i=1}^n \left\{ 1 - P(|U_i| > \delta) \right\}. \tag{2.15}$$

By (2.15) and Proposition 2.3 the results easily follow.

The uniform convergence of $T_{ij}$ to $\alpha_i \beta_j$ is given in the following theorem. We write $N = IJ$ for the total number of observations. The numbers $I$ and $J$ are considered as functions of $N$.

Theorem 2.5. Suppose that for all $\delta > 0,$

$$\lim_{N \to \infty} I \exp(-\delta \beta_+^2 J^{-1}) = 0 \tag{2.16}$$

and

$$\lim_{N \to \infty} J \exp(-\delta \alpha_+^2 I^{-1}) = 0, \tag{2.17}$$

then

$$\max_{1 \leq i \leq I, 1 \leq j \leq J} |T_{ij} - \alpha_i \beta_j| \overset{p}{\longrightarrow} 0 \quad \text{as} \quad N \to \infty. \tag{2.18}$$

Proof. It suffices to show that

$$\max_{1 \leq i \leq I, 1 \leq j \leq J} \left| \frac{X_{+i}}{\beta_+} - \alpha_i \right| = \max_{1 \leq i \leq I} \left| \frac{X_{+i}}{\beta_+} - \alpha_i \right| \overset{p}{\longrightarrow} 0 \tag{2.19}$$

and

$$\max_{1 \leq i \leq I, 1 \leq j \leq J} \left| \frac{X_{+j}}{\alpha_+} - \beta_j \right| = \max_{1 \leq j \leq J} \left| \frac{X_{+j}}{\alpha_+} - \beta_j \right| \overset{p}{\longrightarrow} 0 \tag{2.20}$$
and

$$\max_{1 \leq i, j \leq f} \left| \frac{X_{++}}{\alpha_+ \beta_+} - 1 \right| = \left| \frac{X_{++}}{\alpha_+ \beta_+} - 1 \right| \rightarrow 0.$$  \hspace{1cm} (2.21)$$

By (2.12) and (2.16) we obtain (2.19), while (2.20) follows from (2.13) and (2.17). Since (2.16) implies $\beta^2_+ J^{-1} \rightarrow \infty$ and (2.17) implies $\alpha^2_+ I^{-1} \rightarrow \infty$, we have $\alpha^2_+ \beta^2_+ I^{-1} J^{-1} \rightarrow \infty$ as $N \rightarrow \infty$. Application of (2.14) now yields (2.21), thus completing the proof.

Conditions (2.16) and (2.17) imply that almost empty rows or columns are not permitted. This agrees with what we want. A lot of empty cells may occur, but the expected number of non-empty cells in a row or column must be larger than the root of the number of columns or rows, respectively.

As an application of the preceding theorems we consider the construction of a confidence interval for $a_i \beta_j$, where $i$ and $j$ are fixed. We assume for simplicity

$$\alpha_k \beta_l \in [\varepsilon_0, 1 - \varepsilon_0] \quad \text{for all } k, l \text{ and some } \varepsilon_0 > 0.$$  

Further assume for all $\delta > 0$,

$$I \exp(-\delta J) \rightarrow 0, \quad J \exp(-\delta I) \rightarrow 0.$$  

Under these weak conditions an approximate level $(1 - \eta)$ confidence interval for $\alpha \beta_j$ can be constructed as follows. Since

$$\left| \sum_{i=1}^f T_{ii}(1 - T_{ii})/v^2_1(\alpha_i) \right| \leq \frac{J \max_{1 \leq i \leq j} |T_{ii} - \alpha_i \beta_i|}{J \varepsilon_0^2},$$

it follows from Theorem 2.5 that

$$\sum_{l=1}^J T_{ii}(1 - T_{ii})/v^2_1(\alpha_i) \rightarrow^P 1.$$  

Similarly we obtain

$$\sum_{k=1}^J T_{kk}(1 - T_{kk})/v^2_2(\beta_j) \rightarrow^P 1.$$  

In view of (2.1) and (2.3) we have

$$\left( \frac{X_{++}}{X_{++}} \right)^2 \left( \frac{\alpha_+}{\alpha_+} \right)^2 \rightarrow^P 1.$$
and by (2.2) and (2.3)
\[
\left( \frac{X_{++}}{X_{++}} \right)^2 \left( \frac{\beta_j}{\hat{\beta}_j} \right)^2 \xrightarrow{D} 1.
\]

It is now easily seen that
\[
(T_y - \Phi^{-1}(1 - \frac{1}{2}\eta)S, T_y + \Phi^{-1}(1 - \frac{1}{2}\eta)S)
\]
with
\[
s^2 = \left( \frac{X_{++}}{X_{++}} \right)^2 \sum_{k=1}^{I} T_{kj}(1 - T_{kj}) + \left( \frac{X_{++}}{X_{++}} \right)^2 \sum_{l=1}^{J} T_{li}(1 - T_{li})
\]
is an approximate level \((1 - \eta)\) confidence interval for \(\alpha_i\beta_j\). Here \(\Phi\) denotes the standard normal distribution function.

3. Goodness-of-Fit Tests

The estimation theory developed in Section 2 is valid under the basic condition of independence between rows and columns. In this section we will investigate this basic assumption by testing the null hypothesis
\[
H_0: p_{ij} = \alpha_i\beta_j, \quad i = 1, \ldots, I, j = 1, \ldots, J,
\]
where \(p_{ij} = P(X_{ij} = 1)\) and \(0 < \alpha_i < 1\) and \(0 < \beta_j < 1\) are unknown probabilities. As usual in goodness-of-fit testing problems we are interested in tests with good power properties against a broad class of alternatives. A well-known class of goodness-of-fit tests with, in general, nice overall power properties is based on so-called directed divergence measures. Here it is supposed that \(I\) and \(J\) are large, and hence under \(H_0\) a large number of nuisance parameters is involved. Therefore standard techniques in contingency tables cannot be applied.

To construct the class of test statistics for the preceding testing problem we first define directed divergence measures between \(0 \leq q \leq 1\) and \(0 < p < 1\) by
\[
I'(q; p) = \frac{2}{\lambda(\lambda + 1)} \left[ q \left\{ \frac{q}{p} \right\}^{\lambda} - 1 \right] + (1 - q) \left\{ \frac{1 - q}{1 - p} \right\}^{\lambda} - 1 \right].
\]
\text{cf. Cressie and Read [3]. We restrict attention to } \lambda > -1, \text{ since } I'(0; p) =
$I^\lambda(1 : p) = \infty$ for $\lambda \leq -1$. For $\lambda = 0$ the measure $I^\lambda$ is defined by continuity, yielding

$$I^0(q : p) = 2 \left\{ q \log\frac{q}{p} + (1 - q) \log\frac{1 - q}{1 - p} \right\}. \quad (3.3)$$

These measures can be used to embed classical multinomial goodness-of-fit statistics in a family indexed by $\lambda$. Note that $\lambda = 1$ gives

$$I^1(q : p) = \frac{(q - p)^2}{p} + \frac{((1 - q) - (1 - p))^2}{1 - p},$$

corresponding to Pearson's chi-square statistic, and that $I^0$ equals twice the Kullback–Leibler information number, corresponding to a likelihood ratio statistic.

Next consider the statistics

$$Y^\lambda = \sum_{i=1}^I \sum_{j=1}^J I^\lambda(X_{ij}; \alpha_i \beta_j), \quad (3.4)$$

measuring the distance between the observed $X_{ij}$'s and the null probabilities $\alpha_i \beta_j$. First, conditions on $\alpha_i \beta_j$ are presented ensuring asymptotic normality of

$$\tilde{Y}^\lambda = \frac{Y^\lambda - EY^\lambda}{\sqrt{\text{var} Y^\lambda}} \quad (3.5)$$

as $IJ \to \infty$. The conditions are rather mild. However, if $\alpha_i \beta_j$ is close to $\frac{1}{2}$, asymptotic normality may fail. If we look at $I^\lambda(X_{ij}; \frac{1}{2})$, it becomes clear why this occurs. As one might expect of a distance measure, the distance from 0 to $\frac{1}{2}$ equals the distance from 1 to $\frac{1}{2}$. For $I^\lambda$ this is indeed the case, and hence

$$I^\lambda(X_{ij}; \frac{1}{2}) = 2\lambda^{-1}(\lambda + 1)^{-1}(2^\lambda - 1),$$

irrespective whether $X_{ij} = 0$ or 1. Therefore, if $\alpha_i \beta_j = \frac{1}{2}$ for all $i, j$, the statistic $Y^\lambda$ is degenerate, cf. also Remark 3.3.

Of course we cannot use $\tilde{Y}^\lambda$ as a test statistic since it contains the unknown parameters $\alpha_i \beta_j$. We therefore insert in $\tilde{Y}^\lambda$ everywhere the estimator $T_\theta$ of $\alpha_i \beta_j$. However, the number of parameters to be estimated is too large relative to the number of observations to obtain the same asymptotic distribution for this statistic as for $\tilde{Y}^\lambda$ itself. It turns out that inserting the estimators yields a serious bias term. Moreover, the asymptotic variance has to be adjusted. After modification we obtain a
more complicated family of test statistics, which however are again asymptotically normal and can therefore be used in testing multiplicativity.

3.1. Asymptotic Normality of $\bar{Y}^\lambda$ under $H_0$

To investigate the asymptotic normality of $\bar{Y}^\lambda$, firstly we calculate the expectation and variance of $Y^\lambda$, cf. (3.4),

$$EY^\lambda = \sum_{i=1}^{I} \sum_{j=1}^{J} \mu_\lambda(\alpha_i \beta_j)$$

with

$$\mu_\lambda(p) = 2\lambda^{-1}(\lambda + 1)^{-1}\{p^{1-\lambda} + (1 - p)^{1-\lambda} - 1\}, \quad \lambda \neq 0$$

(3.6)

and

$$\mu_0(p) = -2p \log p - 2(1 - p) \log(1 - p)$$

and

$$\text{var } Y^\lambda = \sum_{i=1}^{I} \sum_{j=1}^{J} \sigma_\lambda^2(\alpha_i \beta_j)$$

with

$$\sigma_\lambda^2(p) = 4\lambda^{-2}(\lambda + 1)^{-2}p(1 - p)\{p^{-\lambda} - (1 - p)^{-\lambda}\}^2, \quad \lambda \neq 0$$

(3.7)

$$\sigma_0^2(p) = 4p(1 - p)\left\{\log\left(\frac{p}{1 - p}\right)\right\}^2.$$

Note that $\mu_0(p) = \lim_{\lambda \to 0} \mu_\lambda(p)$ and $\sigma_0^2(p) = \lim_{\lambda \to 0} \sigma_\lambda^2(p)$.

The proof of the following theorem is based on the Lindeberg–Feller Central Limit Theorem.

**Theorem 3.1.** Assume that $\alpha_i \beta_j \neq \frac{1}{2}$ for some $i, j$, and that

$$\frac{\max\{[(\alpha_i \beta_j)^{-\lambda} - (1 - \alpha_i \beta_j)^{-\lambda}]^2; i = 1, \ldots, I, j = 1, \ldots, J\}}{\sum_{i=1}^{I} \sum_{j=1}^{J} \alpha_i \beta_j(1 - \alpha_i \beta_j)[(\alpha_i \beta_j)^{-\lambda} - (1 - \alpha_i \beta_j)^{-\lambda}]^2} \to 0$$

(3.8)

as $IJ \to \infty$; then

$$\bar{Y}^\lambda \overset{D}{\to} N(0, 1)$$

(3.9)

as $IJ \to \infty$.

**Proof.** Writing

$$Y^\lambda = I^\lambda(X^\alpha; \alpha_i \beta_j) - \mu_\lambda(\alpha_i \beta_j),$$
we obtain a sequence \( \{ Y_{ij}^\lambda \} \) of independent r.v.'s with

\[
EY_{ij}^\lambda = 0 \quad \text{and} \quad E(Y_{ij}^\lambda)^2 = \sigma_j^2(\alpha_i, \beta_j).
\]

By direct calculation we have

\[
I^\lambda(0: p) - \mu_\lambda(\alpha_i, \beta_j) = 2\lambda^{-1}(\lambda + 1)^{-1}p\{(1 - p)^{-\lambda} - p^{-\lambda}\}
\]

and hence

\[
| Y_{ij}^\lambda | \leq 2\lambda^{-1}(\lambda + 1)^{-1}|(\alpha_i, \beta_j)^{-\lambda} - (1 - \alpha_i, \beta_j)^{-\lambda} | \quad (3.11)
\]

for all \( i = 1, \ldots, I, j = 1, \ldots, J \).

The condition that \( \alpha_i, \beta_j \neq \frac{1}{2} \) for some \( i, j \) ensures that \( Y^\lambda \) is not degenerate. Moreover, this condition together with (3.8) implies that

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} \sigma_j^2(\alpha_i, \beta_j) \to \infty,
\]

because the numerator of (3.8) does not tend to zero. By (3.8) and (3.11) it now easily follows that the Lindeberg condition is satisfied, implying (3.9) by application of Theorem 1 on p. 519 of Feller (1971).

**Remark 3.1.** The conditions of Theorem 3.1 are satisfied if \( \alpha_i, \beta_j \) stays away from 0, \( \frac{1}{2} \), and 1. Condition (3.8) states how close \( \alpha_i, \beta_j \) may be to these possible limiting points of (sub)sequences of \( \alpha_i, \beta_j \).

**Remark 3.2.** For \( \lambda = 1 \) corresponding to Pearson's chi-square test condition (3.8) reads

\[
\max \left\{ \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} (1 - 2\alpha_i \beta_j)^2 (\alpha_i - 1)^2 (1 - \alpha_i, \beta_j)^{-2}}{\sum_{i=1}^{I} \sum_{j=1}^{J} (1 - 2\alpha_i \beta_j)^2 (\alpha_i, \beta_j)^{-1} (1 - \alpha_i, \beta_j)^{-1}} \right\} \to 0 \quad (3.12)
\]

as \( IJ \to \infty \). Again this tells us how close to 0, \( \frac{1}{2} \), and 1 the \( \alpha_i, \beta_j \)'s may be.

**Remark 3.3.** If \( \alpha_i, \beta_j = \frac{1}{2} \) for all \( i, j \) the statistic \( Y^\lambda \) is degenerate and \( \text{var} Y^\lambda = 0 \). To see what happens in the neighbourhood of this special case with \( \hat{Y}^\lambda \), suppose that \( \alpha_i, \beta_j \to \frac{1}{2} \) in such a way that \( \sum_{i=1}^{I} \sum_{j=1}^{J} (1 - 2\alpha_i \beta_j)^2 \to \infty \) as \( IJ \to \infty \). In that case the numerator of (3.8) is bounded, while the denominator tends to \( \infty \), since

\[
\lambda^{-2}\left[p^{-\lambda} - (1 - p)^{-\lambda}\right]^2 = 2^{2\lambda}4(1 - 2p)^2 + O((1 - 2p)^4) \quad \text{as} \quad p \to \frac{1}{2}.
\]

Therefore we can again apply Theorem 3.1 to obtain asymptotic normality of \( \hat{Y}^\lambda \). If \( \alpha_i, \beta_j \to \frac{1}{2} \) in such a way that \( \sum_{i=1}^{I} \sum_{j=1}^{J} (1 - 2\alpha_i \beta_j)^2 \) remains...
bounded, then asymptotic normality usually fails. For instance, if \( \alpha_i \beta_j = \frac{1}{2} \)
for all \( i \geq I_0 \) and \( j \geq J_0 \), we have in fact only finitely many terms in \( \hat{Y}^2 \), since in this case \( Y_0^2 = 0 \) for all \( i \geq I_0 \) and \( j \geq J_0 \).

**Remark 3.4.** Note that in Theorem 3.1 we only assume \( JJ \to \infty \), including the case of fixed \( I \) and \( J \to \infty \) or fixed \( J \) and \( I \to \infty \). So it is not necessary that both \( I \) and \( J \) tend to \( \infty \).

3.2. The Test Statistics and Their Asymptotic Distribution

The statistic \( \hat{Y}^2 \) cannot be used for testing \( H_0 \) since it contains the nuisance parameters \( \alpha_i \beta_j \). Therefore one would like to insert in \( \hat{Y}^2 \) everywhere the estimator \( T_{ij} \) to obtain a test statistic for testing \( H_0 \). However, there are so many parameters in \( \hat{Y}^2 \) to be estimated, relative to the number of observations in \( \hat{Y}^2 \), that the contribution of the estimating process is not negligible. Therefore we have to make a careful analysis of the influence of the estimating procedure, leading to a modification of the test statistic in such a way that again asymptotic normality is obtained. Then this statistic can be used for testing \( H_0 \).

We assume that \( \alpha_i \beta_j \) stays away from 0 and 1. More precisely, we assume that there exists \( \varepsilon_0 > 0 \) such that

\[
\alpha_i \beta_j \in [\varepsilon_0, 1 - \varepsilon_0] \quad (3.13)
\]

for all \( i \in \{1, \ldots, I\} \), \( j \in \{1, \ldots, J\} \).

Taking \( \delta = (J \log I)^{1/2} \beta_{-1}^{-1} \) in (2.12) and \( \delta = (I \log J)^{1/2} \alpha_{-1}^{-1} \) in (2.13) it follows by (3.13) that

\[
\max_{1 \leq j \leq J} |T_{ij} - \alpha_i \beta_j| = O_p((I^{-1} \log J)^{1/2} + (J^{-1} \log J)^{1/2}). \quad (3.14)
\]

Further, we assume

\[
I^{-2}J \log J + J^{-2}I \log I = o(1) \quad (3.15)
\]

as \( N \to \infty \). It is seen from (3.13), (3.14), and (3.15) that \( T_{ij} \) stays away from 0 and 1 in probability uniformly in \( i, j \). Writing

\[
f(p) = \begin{cases} 
2\lambda^{-1}(\lambda + 1)^{-1} \{ (1 - p)^{-\lambda} - p^{-\lambda} \} & \lambda \neq 0 \\
\log \{ p/(1 - p) \} & \lambda = 0
\end{cases}
\]

we have, in view of (3.10),

\[
I^2(X_0; T_{ij}) \sim \mu_0(T_{ij}) = (T_{ij} - X_0) f(T_{ij}). \quad (3.16)
\]
We write \( T_{ij} - X_{ij} = (T_{ij} - \alpha_i\beta_j) + (\alpha_i\beta_j - X_{ij}) \) and use a Taylor expansion of \( f(T_{ij}) \) about \( \alpha_i\beta_j \). Combining terms with the same power of \( T_{ij} - \alpha_i\beta_j \) we obtain

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} \left\{ I_i^*(X_{ij}; T_{ij}) - \mu_i(T_{ij}) \right\}(IJ)^{-1/2} = 
\sum_{i=1}^{I} \sum_{j=1}^{J} \left[ ((\alpha_i\beta_j - X_{ij}) f(\alpha_i\beta_j) 
+ \{ T_{ij} - \alpha_i\beta_j \} \{ f(\alpha_i\beta_j) + f'(\alpha_i\beta_j)(\alpha_i\beta_j - X_{ij}) \} 
+ \{ T_{ij} - \alpha_i\beta_j \}^2 \{ \frac{1}{2} f''(\alpha_i\beta_j)(\alpha_i\beta_j - X_{ij}) \} 
+ \{ T_{ij} - \alpha_i\beta_j \}^3 \{ \frac{1}{6} f'''(\xi_{ij})(\alpha_i\beta_j - X_{ij}) \} 
+ \{ T_{ij} - \alpha_i\beta_j \}^4 \frac{1}{24} f^{(4)}(\xi_{ij}) \} \right](IJ)^{-1/2},
\]

where \( \xi_{ij} \) are (random) points between \( T_{ij} \) and \( \alpha_i\beta_j \). Note that the first term on the right-hand side of (3.17) equals (cf. (3.10) or (3.16))

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} \left\{ I_i^*(X_{ij}; \alpha_i\beta_j) - \mu_i(\alpha_i\beta_j) \right\}(IJ)^{-1/2}.
\]

Define

\[
\tau_{ij} = \sum_{k=1}^{I} \alpha_k \alpha_k^{-1} f(\alpha_k \beta_j) + \sum_{l=1}^{J} \beta_l \beta_l^{-1} f(\alpha_i \beta_l) 
- \sum_{k=1}^{I} \sum_{l=1}^{J} \alpha_k \alpha_k^{-1} \beta_l \beta_l^{-1} f(\alpha_k \beta_l) - f(\alpha_i \beta_j),
\]

(3.18)

and

\[
\zeta = \sum_{i=1}^{I} \sum_{j=1}^{J} (\alpha_i \alpha_i^{-1} + \beta_j \beta_j^{-1}) f'(\alpha_i \beta_j) \alpha_i \beta_j (1 - \alpha_i \beta_j)(IJ)^{-1/2},
\]

\[
\gamma = \sum_{i=1}^{I} \sum_{j=1}^{J} f'(\alpha_i \beta_j) \{ \alpha_i^2 \alpha_i^{-2} v^2_1(\beta_j) + \beta_j^2 \beta_j^{-2} v^2_2(\alpha_i) \} (IJ)^{-1/2}
\]

(3.20)

with \( v^2_1(\alpha_i) \) and \( v^2_2(\beta_j) \) given by (2.5).

Application of Lemma A.2 in the Appendix yields

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} f(\alpha_i \beta_j)(T_{ij} - \alpha_i\beta_j)(IJ)^{-1/2} = 
\sum_{i=1}^{I} \sum_{j=1}^{J} \{ \tau_{ij} + f(\alpha_i \beta_j) \}(X_{ij} - \alpha_i\beta_j)(IJ)^{-1/2} + o_p(1)
\]

(3.21)
and
\[
\sum_{i=1}^{I} \sum_{j=1}^{J} f'(x_i, y_j)(x_i y_i - x_i, y_j)y_j(\beta_j - y_j)(IJ)^{-1/2} = -\zeta + o_p(1). \tag{3.22}
\]

It follows from Lemma A.3, Remark A.1 in the Appendix, and (3.15) that
\[
\sum_{i=1}^{I} \sum_{j=1}^{J} f'(x_i, y_j)(T_{ij} - x_i, y_j)^2(IJ)^{-1/2} = \gamma + o_p(1) \tag{3.23}
\]
and
\[
\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{1}{2} f''(x_i, y_j)(x_i y_i - x_i, y_j)^2(\beta_j - y_j)^2(IJ)^{-1/2} = o_p(1). \tag{3.24}
\]

The remainder terms of the right-hand side of (3.17) are of order
\[
O_p \left( \sum_{i=1}^{I} \sum_{j=1}^{J} |T_{ij} - x_i, y_j|^3(IJ)^{-1/2} \right),
\]
which in view of (3.14), (3.15), and Lemma A.3 is of order
\[
O_p \left( \max_{1 \leq i, j \leq J} |T_{ij} - x_i, y_j| (IJ)^{1/2}(I^{-1} + J^{-1}) \right)
= O_p(I^{-2}J \log J)^{1/2} + (J^{-2}I \log I)^{1/2}) = o_p(1).
\]
Together with (3.21), (3.22), (3.23), and (3.24) this implies (cf. (3.17))
\[
\sum_{i=1}^{I} \sum_{j=1}^{J} \left\{ I^2(X_i, y_i; T_{ij}) - \mu_x(T_{ij}) \right\}(IJ)^{-1/2}
= \sum_{i=1}^{I} \sum_{j=1}^{J} g_y(X_i, y_i)(IJ)^{-1/2} + \gamma - \zeta + o_p(1). \tag{3.25}
\]
So we see that the estimating procedure gives a bias term $\gamma - \zeta_1$ which does not tend to zero. Moreover, although the first term on the right-hand side of (3.25) has expectation 0, its asymptotic variance differs from the asymptotic variance of $Y^2(IJ)^{-1/2}$.

By the Lindeberg–Feller Central Limit Theorem it follows similarly as in the proof of Theorem 3.1 that
\[
\frac{\sum_{i=1}^{I} \sum_{j=1}^{J} g_y(X_i, y_i)(IJ)^{-1/2}}{(I^{-1}J^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \tau_y^2 x_i, y_i(1 - x_i, y_j))^{1/2}} \overset{D}{\rightarrow} N(0, 1). \tag{3.26}
\]
This condition seems not very restrictive, but if $\alpha_i\beta_j$ is the same for all $i$ and $j$, $\tau_{ij} = 0$ for all $i, j$. Next we replace in $\tau_{ij}^2$ everywhere $\alpha_i\beta_j$ by the estimator $T_{ij}$, $\alpha_i\alpha_i^{-1}$ by the estimator $X_+X_+^{-1}$ and $\beta_j\beta_j^{-1}$ by the estimator $X_+X_+^{-1}$ and denote this statistic by $\hat{\tau}_{ij}^2$. In view of (3.13), (3.27) and the uniform convergence of the estimators we have

$$
\sum_{i=1}^I \sum_{j=1}^J \hat{\tau}_{ij}^2 T_{ij}(1 - T_{ij}) \right \} \sum_{i=1}^I \sum_{j=1}^J \tau_{ij}^2 \alpha_i \beta_j (1 - \alpha_i \beta_j) \right \} \rightarrow 1
$$

as $N \rightarrow \infty$. Similarly, we insert in $\gamma$ and $\zeta$ everywhere the estimators, yielding $\hat{\gamma}$ and $\hat{\zeta}$ with the property

$$\hat{\gamma} - \gamma = O_p((I^{-2}J \log J)^{1/2} + (J^{-2}I \log I)^{1/2}) = o_p(1)$$

$$\hat{\zeta} - \zeta = o_p(1)$$

as $N \rightarrow \infty$.

The family of test statistics is now defined as

$$W^* = \frac{\sum_{i=1}^I \sum_{j=1}^J \{ \bar{I}_i(X_{ij}; T_{ij}) - \mu_i(T_{ij}) \} - \hat{\gamma}(IJ)^{1/2} + \hat{\zeta}(IJ)^{1/2}}{\sum_{i=1}^I \sum_{j=1}^J \hat{\tau}_{ij}^2 T_{ij}(1 - T_{ij})^{1/2}}$$

and we have proved the following theorem.

**Theorem 3.2.** Suppose that (3.13), (3.15), and (3.27) hold. Then

$$W^* \xrightarrow{D} N(0, 1)$$

as $N \rightarrow \infty$.

**Remark 3.1.** With the test statistic $W^*$ we can in fact test the null hypothesis

$$H_0^*: p_{ij} = \alpha_i \beta_j \quad i = 1, \ldots, I, j = 1, \ldots, J$$

with

$$\alpha_i \beta_j \in [\varepsilon_0, 1 - \varepsilon_0] \quad \text{for some } \varepsilon_0 > 0$$

by rejecting $H_0^*$ when

$$|W^*| > \Phi^{-1}(1 - \eta/2).$$
Provided that the conditions of Theorem 3.2 hold, the asymptotic level of this test will be \( \eta \). Whether this test is also a test of \( H_0 \) with asymptotically level \( \eta \), depends on the asymptotic behaviour of \( W^2 \) under null probabilities \( \alpha_i \beta_j \) with \( \alpha_i \beta_j < \epsilon_0 \), \( \alpha_i \beta_j > 1 - \epsilon_0 \), or \( I^{-1} J^{-1} \sum_{i=1}^I \sum_{j=1}^J \tau_{ij}^2 \) tending to zero as \( N \to \infty \).

**Remark 3.2.** The results of this paper can be generalized to the case where \( X_{ij} \) follows a binomial distribution given by

\[
P(X_{ij} = x_{ij}) = \binom{M}{x_{ij}} p_{ij}^{x_{ij}}(1 - p_{ij})^{M - x_{ij}}
\]

with \( M \) some fixed number. The estimator \( T_{ij} \) of \( \alpha_i \beta_j \) is now defined by

\[
T_{ij} = \frac{X_{ij}^+ X_{+j}^-}{MX_{++}}.
\]

The family of test statistics are found by expanding \( I(X_{ij}^M^{-1}; T_{ij}) - EI^2(X_{ij}^M^{-1}; T_{ij}) \) with respect to \( T_{ij} \) about \( \alpha_i \beta_j \) and arguing in the same way as in proving Theorem 3.2. The formulae for the correction terms \( \delta \) and \( \tau_{ij}^2 \) are rather complicated. So we do not present them here.

**Remark 3.3.** If the modified estimator

\[
T_{ij} = \frac{X_{ij}^{(i)} X_{ij}^{(i)}}{X_{ij}^{(i)} X_{ij}^{(i)}} \quad (3.29)
\]

is used, (3.22) holds with \( \zeta = 0 \) and hence Theorem 3.2 holds in that case with \( \zeta = 0 \).

**APPENDIX**

Let \( N = IJ \) be the total number of observations. The numbers \( I \) and \( J \) are considered as functions of \( N \).

**Lemma A.1.** Let \( h_{ij} \) be uniformly bounded functions. Then we have

\[
\sum_{i=1}^I \sum_{j=1}^J h_{ij}(X_{ij}) I^{-1}(X_{ij} - \alpha_i \beta_j)(IJ)^{-1/2}
\]

\[
= \sum_{i=1}^I \sum_{j=1}^J \left[ \left( \sum_{k=1}^K E h_{ij}(X_{ki}) I^{-1} \right)(X_{ij} - \alpha_i \beta_j) + E\{h_{ij}(X_{ij})(X_{ij} - \alpha_i \beta_j) I^{-1}\} \right] (IJ)^{-1/2} + o_p(1) \quad (A.1)
\]

as \( N \to \infty \).
Proof. Without loss of generality assume $Eh_i(X_{ij}) = 0$ for all $i, j$. Since the functions $h_{ij}$ are uniformly bounded,

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \left[ h_{ij}(X_{ij})(X_{ij} - \alpha_i \beta_j) - E\{h_{ij}(X_{ij})(X_{ij} - \alpha_i \beta_j)^2\}\right] I^{-1}(IJ)^{-1/2} = O_p(I^{-1}) = o_p(1)$$

as $N \to \infty$. Hence it suffices to show

$$U_N = \sum_{i=1}^{I} \sum_{j=1}^{J} h_{ij}(X_{ij}) I^{-1}(X_{ij}^{(i)} - \alpha_*^{(i)} \beta_j)(IJ)^{-1/2} = o_p(1) \quad (A.2)$$

as $N \to \infty$, where

$$X_{ij}^{(i)} = X_{i+j} - X_{ij} \quad \text{and} \quad \alpha_*^{(i)} = \alpha_* - \alpha_i. \quad (A.3)$$

We have $EU_N = 0$ and

$$\text{var } U_N = \sum_{j=1}^{J} \text{var } U_{jN} \quad \text{with } U_{jN} = \sum_{i=1}^{I} U_{ijN}$$

and

$$U_{ijN} = h_{ij}(X_{ij}) I^{-1}(X_{ij}^{(i)} - \alpha_*^{(i)} \beta_j)(IJ)^{-1/2}.$$

Since

$$\text{var } U_{ijN} = O(I^{-2}I(JI)^{-1}) = O(I^{-2}J^{-1})$$

and

$$\text{cov}(U_{ijN}, U_{kjN}) = O(I^{-3}J^{-1}) \quad \text{for } k \neq i,$$

we obtain

$$\text{var } U_{jN} = O(I^{-1}J^{-1}),$$

implying

$$\text{var } U_N = O(I^{-1}),$$

and hence $U_N \to^p 0$.  

Lemma A.2. Let $h_{ij}$ be uniformly bounded functions. Then we have
\[
\sum_{i=1}^{I} \sum_{j=1}^{J} h_{ij}(X_{ij})(T_{ij} - \alpha_I \beta_j)(IJ)^{-1/2} \\
= \sum_{i=1}^{I} \sum_{j=1}^{J} \left\{ \sum_{k=1}^{I} E h_{ki}(X_{ki}) \alpha_k \beta_i^{-1} + \sum_{i=1}^{I} E h_{ij}(X_{ij}) \beta_i \beta_j^{-1} \right\} (X_{ij} - \alpha_i \beta_j)(IJ)^{-1/2} \\
+ \sum_{i=1}^{I} \sum_{j=1}^{J} (\alpha_i \beta_i^{-1} + \beta_j \beta_j^{-1}) E(h_{ij}(X_{ij})(X_{ij} - \alpha_i \beta_j))(IJ)^{-1/2} + o_p(1) \\
(A.4)
\]

as \(N \to \infty\).

**Proof.** By (1.6) we have

\[
T_{ij} - \alpha_i \beta_j = \left( \frac{X_{ij} + X_{ij} - \alpha_i \beta_j}{X_{ij} + \beta_j} \right)(X_{ij} + \beta_j)^{-1} + \frac{X_{ij} + X_{ij} - \alpha_i \beta_j}{X_{ij} + \beta_j}. \\
(A.5)
\]

Since the right-hand side of (A.4) equals \(O_p(1)\) and \(\alpha_i \beta_j(X_{ij} + \beta_j)^{-1} - 1 \to^p 0\) by (2.3), it suffices to prove (A.4) with \(T_{ij} - \alpha_i \beta_j\) on the left-hand side replaced by \((X_{ij} + X_{ij} - \alpha_i \beta_j)(X_{ij} + \beta_j)^{-1}\).

Next we write

\[
X_{i+} X_{i+} - \alpha_i \beta_j X_{i+} = (X_{i+} - \alpha_i \beta_j)(X_{i+} - \alpha_i \beta_j) + \alpha_i \beta_j(X_{i+} - \alpha_i \beta_j) \\
+ \alpha_i \beta_j(X_{i+} - \alpha_i \beta_j) + \alpha_i \beta_j(X_{i+} - \alpha_i \beta_j). \\
(A.6)
\]

We first show that

\[
\sum_{i=1}^{I} \sum_{j=1}^{J} h_{ij}(X_{ij})(X_{i+} - \alpha_i \beta_j)(X_{i+} - \alpha_i \beta_j) \alpha_i^{-1} \beta_j^{-1} (IJ)^{-1/2} = o_p(1). \\
(A.7)
\]

It is easily seen by Lemma A.1 that it suffices to show (A.7) with \(X_{i+} - \alpha_i \beta_j\) replaced by \(X_{i+}^{(i)} - \alpha_i \beta_j^{(i)}\), where \(X_{i+}^{(i)} = X_{i+} - X_{ij}\) and \(\beta_j^{(i)} = \beta_j - \beta_j\), and \(X_{i+} - \alpha_i \beta_j\) replaced by \(X_{i+}^{(i)} - \alpha_i \beta_j^{(i)}\), cf. (A.3) for notation.

Writing

\[
U_N = \sum_{i=1}^{I} \sum_{j=1}^{J} U_{ij/N} \\
\]

with

\[
U_{ij/N} = h_{ij}(X_{ij})(X_{i+}^{(i)} - \alpha_i \beta_j^{(i)})(X_{i+}^{(i)} - \alpha_i \beta_j^{(i)}) \alpha_i^{-1} \beta_j^{-1} (IJ)^{-1/2},
\]

we have \(E U_N = 0\), since \(E U_{ij/N} = 0\) for all \(i, j\).
Further, we obtain

\[ \text{var } U_{ij} = E(U_{ij}^2) = O(JI^{-2}J^{-2}I^{-1}) = O(I^{-2}J^{-2}) \]

\[ \text{cov}(U_{ij}, U_{kl}) = O(I^{-3}J^{-3}) \quad i \neq k, j \neq l \]

\[ \text{cov}(U_{ij}, U_{il}) = 0 \quad j \neq l \]

\[ \text{cov}(U_{ij}, U_{kj}) = 0 \quad i \neq k \]

and hence

\[ \text{var } U = O(I^{-1}J^{-1}), \]

implying \( U \to 0 \), thus completing the proof of (A.7).

Application of Lemma A.1 twice (once interchanging the roles of \( i \) and \( j \)) and noting that

\[ \sum_i \sum_j \{ h_q(X_{iy}) - E h_q(X_{iy}) \} \alpha_i \beta_j \alpha_+^{-1} \beta_+^{-1} \overset{P}{\to} 0 \]

yields (A.4). 1

Remark A.1. If \( T_{ij} \) is replaced by \( T_{ij}^* \) (cf. (3.29)) in (A.4), the term

\[ \sum_i \sum_j (\alpha_i \alpha_+^{-1} + \beta_j \beta_+^{-1}) E \{ h_q(X_{ij} - \alpha_i \beta_j) \} (IJ)^{-1/2} \]

disappears on the right-hand side.

**Lemma A.3.** Let \( h_q \) be uniformly bounded functions. Then we have

\[ \sum_{i=1}^{I} \sum_{j=1}^{J} h_q(X_{ij}^*) (T_{ij} - \alpha_i \beta_j)^2 (IJ)^{-1/2} \]

\[ = \sum_{i=1}^{I} \sum_{j=1}^{J} E \{ h_q(X_{ij}) (X_{ij} - \alpha_+ \beta_j)^2 \} \alpha_+^2 (IJ)^{-1/2} \]

\[ + \sum_{i=1}^{I} \sum_{j=1}^{J} E \{ h_q(X_{ij}) (X_{ij} - \alpha_i \beta_+)^2 \} \beta_+^2 (IJ)^{-1/2} + o_p(1) \quad (A.8) \]

as \( N \to \infty \).

**Proof.** Using (A.5) and (A.6) it is seen that in fact we have to prove

\[ U_N = \sum_{i=1}^{I} \sum_{j=-1}^{J} \{ h_q(X_{ij}) (X_{ij}^{(i)} - \alpha_+^{(i)} \beta_j)^2 \}

\[ - E h_q(X_{ij}) E (X_{ij}^{(i)} - \alpha_+^{(i)} \beta_j)^2 \} \alpha_+^2 (IJ)^{-1/2} = o_p(1), \]

since the essentially different terms are of lower order. We have \( E U_N = 0 \) and \( \text{var } U_N = O(I^{-1}) \), which follows from \( E (X_{ij}^{(i)} - \alpha_+^{(i)} \beta_j)^4 = O(I^2) \). Hence (A.8) is established. 1
Remark A.1. If $E h_{ij}(X_{ij}) = 0$ and $I^{-J} J^{-J} I 	o 0$, then the right-hand side of (A.8) equals $o_P(1)$, since in that case

$$
\sum_{i=1}^{I} \sum_{j=1}^{J} E\{h_{ij}(X_{ij})(X_{ij} - \alpha_i \beta_i)^2\} \alpha_i^2 \beta_i^2 (IJ)^{-1/2} = O(I^{-3/2} J^{1/2}) = o(1)
$$

and

$$
\sum_{i=1}^{I} \sum_{j=1}^{J} E\{h_{ij}(X_{ij})(X_{ij} - \alpha_i \beta_i)^2\} \beta_j^2 \beta_i^2 (IJ)^{-1/2} = O(J^{-3/2} I^{1/2}) = o(1).
$$

ACKNOWLEDGMENTS

The author thanks William J. J. Rey for suggesting the problem and J. Engel for fruitful discussions.

REFERENCES