The Annals of Statistics
1976, Vol. 4, No. 3, 542–553

TESTS FOR INDEPENDENCE IN INFINITE CONTINGENCY TABLES

BY SHINGO SHIRAHATA

Kyushu University

This paper deals with distribution-free tests for independence under the constraint that the population has a bivariate discrete distribution. The locally most powerful conditional test, given the marginal empirical distributions, is derived. The unconditional asymptotic distribution ot the conditional test statistic standardized by the conditional mean and variance is also given under the hypothesis of independence and under contiguous alternatives. Furthermore, some discussions on asymptotic relative efficiency are made. Two competitive test statistics having asymptotically chi-square distributions with different degrees of freedom are compared by means of the local asymptotic relative efficiency.

1. Introduction and summary. Let us consider the problem of testing independence of a bivariate distribution on the basis of a random sample $(X_1, Y_1), \cdots, (X_n, Y_n)$. Suppose the functional form of the joint as well as the marginal distribution functions is unknown. Then, distribution-free tests have to be considered. Concerning distribution-free tests for independence, many authors such as Bhuchongkul (1964), Jogdeo (1968), Ruymgaart (1974), Ruymgaart et al. (1972) and Shirahata (1974a, 1975) studied rank tests. However, the assumption of the continuity of the underlying distribution function in these works is not a realistic one since statistical data are usually obtained by counting numbers or by rounding off a continuous quantity.

When ties occur with positive probability, the usual rank tests do not have a distribution-free character. Convenient ways to maintain the distribution-free character are to consider the randomized ranks or, preferably, conditional tests, given the marginal vectors of ties. Behnen (1973) gave the general asymptotic theory of conditional rank tests. As for the univariate problem, Behnen (1976), Conover (1973) and Vorličková (1970, 1972) studied conditional rank tests. This paper deals with a conditional test for discrete distributions conditioned, as in Shirahata (1974b), on the marginal empirical distributions, not on the marginal vectors of ties. Furthermore, the asymptotic distribution is also considered since the determination of the exact critical point is quite tedious for large n. Other works about distribution-free tests for independence when ties are present are given by Burr (1960), Cureton (1958), Lyerly (1952), Robillard (1972) and Sillitto (1947).

Received November 1974; revised November 1975.

AMS 1970 *subject classifications.* Primary 62G10; Secondary 62E20, 62G20.

Key words and phrases. Distribution-free test for independence, discrete distribution, conditional test, asymptotic normality, rank test, contingency table, local asymptotic relative efficiency.

The statistic to be investigated is not based on ranks but on the original data. In Section 2 the locally most powerful conditional test, given the marginal empirical distributions, is derived. A large class of statistics is also proposed and the conditional mean and variance are calculated under the hypothesis of independence H_0. In Sections 3 and 4, the test statistic standardized by the conditional mean and variance is dealt with. The unconditional asymptotic normality is proved under a subhypothesis of H_0 and under contiguous alternatives. In Section 5, asymptotic efficiency of our test is given. A comparison between the test based on the square of our statistic and the chi-square test is also discussed by employing the notion of local asymptotic relative efficiency suggested by Hájek and Šidák (1967). In addition, some attentions are paid to the rank test theory developed by [1].

2. The locally most powerful conditional test for independence. Assume the common distribution of the (X_k, Y_k) is discrete. For simplicity, assume that each (X_k, Y_k) is distributed on $J \times J$ where J denotes the set of integers. The set $J \times J$ may be replaced by a product $J_1 \times J_2$ of arbitrary countable sets J_1 and J_2. Define

$$\tau_{ij} = \# \{k \mid (X_k, Y_k) = (i, j)\} \qquad (i, j) \in J \times J$$

$$\tau_{i\cdot} = \# \{k \mid X_k = i\}, \qquad \tau_{\cdot j} = \# \{k \mid Y_k = j\} \qquad i, j \in J \tag{2.1}$$

$$\tau_1 = \{\tau_{i\cdot}; i \in J\}, \qquad \tau_2 = \{\tau_{\cdot j}; j \in J\}$$

and

$$\tau = \{\tau_{ij}; (i, j) \in J \times J\}$$

where the symbol $\#$ denotes the number of the elements of the specified set. Then, clearly (τ_1, τ_2) is a sufficient statistic under H_0 and τ is sufficient for the class of distributions on $J \times J$. Thus, in order to obtain distribution-free tests of H_0, we can restrict attention to tests based on τ on condition that (τ_1, τ_2) is given. Converting the sample into τ is equivalent to regarding the sample as a contingency table. Throughout this paper, the word 'conditional' means 'on condition that (τ_1, τ_2) is given' unless otherwise stated.

Suppose that the common distribution of (X_k, Y_k)'s is governed by a parameter θ and that X_k and Y_k are independent of each other if and only if $\theta = 0$. Denote by $p_{ij}(\theta)$ the probability of the event $\{(X_k, Y_k) = (i, j)\}$ at θ and by H_θ the associated hypothesis. Put $p_{ij}(0) = p_{i\cdot}p_{\cdot j}$ where $\sum_i p_{i\cdot} = \sum_j p_{\cdot j} = 1$, $p_{i\cdot} \geqq 0$ and $p_{\cdot j} \geqq 0$ with $\sum_i$ taken over J. The first result in this paper is the following theorem which can be proved by the usual manner of deriving locally most powerful tests in [8].

THEOREM 2.1. *Assume that there exists*

$$p'_{ij} = (d/d\theta)p_{ij}(\theta)|_{\theta=0} \qquad \textit{for all} \quad (i, j) \in J \times J \tag{2.2}$$

and that if $p_{ij}(0) = 0$ *then there exists a positive number* θ_{ij} *such that* $p_{ij}(\theta) = 0$ *for* $0 < \theta < \theta_{ij}$. *Then the locally most powerful conditional test of* H_0 *against* H_θ,

$\theta > 0$ is given by the test with critical region

$$S_1 = \textstyle\sum_{ij} \tau_{ij} p'_{ij}/p_{i\bullet} p_{\bullet j} > c(\tau_1, \tau_2) \tag{2.3}$$

where $c(\tau_1, \tau_2)$ is chosen so that the test has the desired level conditionally, $\sum_{ij}$ being taken over $J \times J$.

It seems difficult to apply S_1 in practice since p'_{ij} may be complicated. However, since S_1 can be rewritten as $S_1 = \sum_{k=1}^{n} p'_{X_k Y_k}/p_{X_k \bullet} p_{\bullet Y_k}$, S_1 is easy to treat when p'_{ij} is a product. Suppose that observations are obtained by rounding off continuous random variables to integers and that the continuous model has the Farlie type (1960) distribution function

$$F(x, y) = F(x)G(y)(1 + \theta A(F(x))B(G(y)) + o(\theta)),$$
$$\text{then} \quad p_{i\bullet} = F(i + \tfrac{1}{2}) - F(i - \tfrac{1}{2}), p_{\bullet j} = G(j + \tfrac{1}{2}) - G(j - \tfrac{1}{2})$$

and p'_{ij} is a product, or specifically

$$\begin{aligned} p'_{ij} = {} & [A(F(i + \tfrac{1}{2}))F(i + \tfrac{1}{2}) - A(F(i - \tfrac{1}{2}))F(i - \tfrac{1}{2})] \\ & \times [B(G(j + \tfrac{1}{2}))G(j + \tfrac{1}{2}) - B(G(j - \tfrac{1}{2}))G(j - \tfrac{1}{2})] \,. \end{aligned}$$

In the following theorem the conditional mean and variance of a linear statistic in τ,

$$S_2(\mathbf{a}) = \textstyle\sum_{ij} a_{ij} \tau_{ij} \,, \tag{2.4}$$

which generalizes the statistic S_1, is calculated under H_0 where a_{ij} is a given constant. Let us denote by E_C, Var_C and P_C the conditional mean, variance and probability, respectively.

THEOREM 2.2. *Under H_0, the conditional mean and variance of $S_2(\mathbf{a})$ are given by*

$$E_C S_2(\mathbf{a}) = \textstyle\sum_{ij} a_{ij} \tau_{i\bullet} \tau_{\bullet j}/n \tag{2.5}$$

and

$$\begin{aligned} \mathrm{Var}_C\, S_2(\mathbf{a}) = {} & (n - 1)^{-1} \textstyle\sum_{ij} a_{ij}^2 \tau_{i\bullet} \tau_{\bullet j} \\ & - [n(n - 1)]^{-1} [\textstyle\sum_j \tau_{\bullet j} (\sum_i a_{ij} \tau_{i\bullet})^2 + \sum_i \tau_{i\bullet} (\sum_j a_{ij} \tau_{\bullet j})^2] \\ & + [n^2(n - 1)]^{-1} (\textstyle\sum_{ij} a_{ij} \tau_{i\bullet} \tau_{\bullet j})^2 \,. \end{aligned} \tag{2.6}$$

The theorem is an easy consequence of the following lemma, which is a generalization of Fisher's expression for the 2 × 2 table.

LEMMA 2.1. *Under H_0, the conditional distribution of τ is given by*

$$P_C(\tau) = (\textstyle\prod_i \tau_{i\bullet}!)(\prod_j \tau_{\bullet j}!)/n! \prod_{ij} \tau_{ij}! \tag{2.7}$$

and. hence

$$\begin{aligned} E_C \tau_{ij} &= \tau_{i\bullet} \tau_{\bullet j}/n \,, \\ \mathrm{Var}_C\, \tau_{ij} &= \tau_{i\bullet} \tau_{\bullet j} (n - \tau_{i\bullet})(n - \tau_{\bullet j})/n^2(n - 1) \,, \\ E_C \tau_{ij} \tau_{ij'} &= \tau_{i\bullet} \tau_{\bullet j} \tau_{\bullet j'} (\tau_{i\bullet} - 1)/n(n - 1) \qquad j \neq j' \\ E_C \tau_{ij} \tau_{i'j} &= \tau_{i\bullet} \tau_{i'\bullet} \tau_{\bullet j} (\tau_{\bullet j} - 1)/n(n - 1) \qquad i \neq i' \end{aligned}$$

and

$$E_C \tau_{ij}\tau_{i'j'} = \tau_{i\bullet}\tau_{i'\bullet}\tau_{\bullet j}\tau_{\bullet j'}/n(n-1) \qquad i \neq i', j \neq j' .$$

When both X_k and Y_k can take only two values, the test based on $S_2(\mathbf{a})$ is, if it is not trivial, equivalent to Fisher's exact test. Thus, the test using $S_2(\mathbf{a})$ can be regarded as an extension of Fisher's exact test to the case of an infinite contingency table.

3. The asymptotic distribution of the test statistic under the hypothesis of independence. In this section the unconditional asymptotic normality of

$$(3.1) \qquad S_3(\mathbf{a}) = (\mathrm{Var}_C\, S_2(\mathbf{a}))^{-\frac{1}{2}}(S_2(\mathbf{a}) - E_C S_2(\mathbf{a}))$$

is investigated under the hypothesis of independence where $S_2(\mathbf{a})$, $E_C S_2(\mathbf{a})$ and $\mathrm{Var}_C\, S_2(\mathbf{a})$ are given in the previous section.

It would be desirable to do without any limitation on the underlying distribution $\{p_{i\bullet}p_{\bullet j}\}$ but indeed we introduce a class of distributions

$$(3.2) \qquad C_{\mathbf{a}} = \{\{p_{i\bullet}p_{\bullet j}\}\,; \qquad \textstyle\sum_{ij} a_{ij}^2 p_{i\bullet}p_{\bullet j} < \infty\}$$

and consider a subhypothesis of H_0,

$H_{0\mathbf{a}}$: the common distribution of (X_k, Y_k)'s belongs to $C_{\mathbf{a}}$.

When $\{a_{ij}\}$ is bounded, $H_{0\mathbf{a}}$ coincides with H_0. Furthermore every distribution with a finite domain belongs to $C_{\mathbf{a}}$ for any $\{a_{ij}\}$. Hence, the limitation may be not so strong. The main result in this section is

THEOREM 3.1. *Under $H_{0\mathbf{a}}$, the convergence*

$$(3.3) \qquad S_3(\mathbf{a}) \rightarrow_d N(0, 1) \qquad as \quad n \rightarrow \infty$$

holds provided $\sigma_1^2(\mathbf{a})$ is positive where

$$(3.4) \qquad \sigma_1^2(\mathbf{a}) = \textstyle\sum_{ij} a_{ij}^2 p_{i\bullet}p_{\bullet j} - \sum_i p_{i\bullet}(\sum_j a_{ij}p_{\bullet j})^2 \\ - \textstyle\sum_j p_{\bullet j}(\sum_i a_{ij}p_{i\bullet})^2 + (\sum_{ij} a_{ij}p_{i\bullet}p_{\bullet j})^2 .$$

Note that the asymptotic normality of $n^{-\frac{1}{2}}(S_2(\mathbf{a}) - E(S_2(\mathbf{a}))$ holds under $H_{0\mathbf{a}}$ since $S_2(\mathbf{a}) = \sum_{k=1}^n a_{X_k Y_k}$ is a sum of independent identically distributed random variables with finite variance. Thus Theorem 3.1 implies that $S_3(\mathbf{a})$ is asymptotically normal under the same condition under which $S_2(\mathbf{a})$ is asymptotically normal.

In order to prove Theorem 3.1, let us put

$$(3.5) \qquad Z_{ij} = n^{-\frac{1}{2}}(\tau_{ij} - np_{i\bullet}p_{\bullet j}) ,$$

$$(3.6) \qquad Z_{i\bullet} = n^{-\frac{1}{2}}(\tau_{i\bullet} - np_{i\bullet})$$

and

$$(3.7) \qquad Z_{\bullet j} = n^{-\frac{1}{2}}(\tau_{\bullet j} - np_{\bullet j}) .$$

Then

$$(3.8) \qquad n^{-\frac{1}{2}}(S_2(\mathbf{a}) - E_C S_2(\mathbf{a})) = \textstyle\sum_{ij} a_{ij} V_{ij} + n^{-\frac{1}{2}} \sum_{ij} a_{ij} Z_{i\bullet} Z_{\bullet j}$$

where

$$V_{ij} = Z_{ij} - p_{i\bullet} Z_{\bullet j} - p_{\bullet j} Z_{i\bullet} \,. \tag{3.9}$$

Thus, Theorem 3.1 can be easily derived by the following lemma.

LEMMA 3.1. *Under H_{0a}, it holds that*

(i) $A_1 = n^{-\frac{1}{2}} \sum_{ij} a_{ij} Z_{i\bullet} Z_{\bullet j} = o_P(1)$,

(ii) $\sum_{ij} a_{ij} V_{ij} \rightarrow_d N(0, \sigma_1{}^2(\mathbf{a}))$ *as* $n \rightarrow \infty$

and

(iii) $n^{-1} \operatorname{Var}_C S_2(\mathbf{a}) \rightarrow_P \sigma_1{}^2(\mathbf{a})$ *as* $n \rightarrow \infty$.

PROOF. (i) Since $\tau_{i\bullet}$ and $\tau_{\bullet j}$ are independent of each other and since $(\tau_{i\bullet}, \tau_{i'\bullet})$ and $(\tau_{\bullet j}, \tau_{\bullet j'})$ each follows trinomial distribution,

$$\begin{aligned} P(|A_1| > \varepsilon) &\leqq E(A_1{}^2)/\varepsilon^2 \\ &= (n\varepsilon^2)^{-1}\Big(\sum_{ij} a_{ij}^2 p_{i\bullet} p_{\bullet j}(1 - p_{i\bullet})(1 - p_{\bullet j}) \\ &\qquad - \sum_{i,j\neq j'} a_{ij} a_{ij'} p_{i\bullet} p_{\bullet j} p_{\bullet j'}(1 - p_{i\bullet}) \\ &\qquad - \sum_{i\neq i',j} a_{ij} a_{i'j} p_{i\bullet} p_{i'\bullet} p_{\bullet j}(1 - p_{\bullet j}) \\ &\qquad + \sum_{i\neq i',j\neq j'} a_{ij} a_{i'j'} p_{i\bullet} p_{i'\bullet} p_{\bullet j} p_{\bullet j'}\Big) \,. \end{aligned} \tag{3.10}$$

By the Schwarz inequality, (3.10) is bounded by $4(n\varepsilon^2)^{-1} \sum_{ij} a_{ij}^2 p_{i\bullet} p_{\bullet j}$ which tends to zero as $n \rightarrow \infty$.

(ii) The proposition required is a well-known property of the multinomial distribution if the domain of the distribution $\{p_{i\bullet} p_{\bullet j}\}$ is finite. Thus we have only to take care of the complication due to the infiniteness of J. For each finite set $K \subset J \times J$, $\sum_{(i,j)\in K} a_{ij} V_{ij}$ is asymptotically normal with mean zero and limiting variance $\sigma_K{}^2(\mathbf{a})$, say. It is easily seen that

$$\sigma_K{}^2(\mathbf{a}) \rightarrow \sigma_1{}^2(\mathbf{a}) \qquad \text{as} \quad K \rightarrow J \times J \,. \tag{3.11}$$

Denote by $\Phi(x)$ and $\Phi_K(x)$ the distribution functions of $N(0, \sigma_1{}^2(\mathbf{a}))$ and $N(0, \sigma_K{}^2(\mathbf{a}))$, respectively. Then

$$\begin{aligned} &|P(\textstyle\sum_{ij} a_{ij} V_{ij} \leqq x) - \Phi(x)| \\ &\qquad \leqq |\Phi_K(x) - \Phi(x)| + |P(\textstyle\sum_{(i,j)\in K} a_{ij} V_{ij} \leqq x) - \Phi_K(x)| \\ &\qquad\qquad + |P(\textstyle\sum_{ij} a_{ij} V_{ij} \leqq x) - P(\textstyle\sum_{(i,j)\in K} a_{ij} V_{ij} \leqq x)| \,. \end{aligned} \tag{3.12}$$

The last term of the right-hand side of (3.12) is bounded by

$$P(|\textstyle\sum_{(i,j)\notin K} a_{ij} V_{ij}| > h) + P(|\textstyle\sum_{(i,j)\in K} a_{ij} V_{ij} - x| \leqq h) \tag{3.13}$$

for every $h > 0$. The second term of (3.13) converges to $\Phi_K(x + h) - \Phi_K(x - h)$ as $n \rightarrow \infty$. Let ε be an arbitrary but small positive number. Then there exists a positive number δ such that

$$\Phi_K(x + h) - \Phi_K(x - h) < \varepsilon \qquad \text{whenever} \quad 2h < \delta\sigma_K \,. \tag{3.14}$$

If we take $h = (\sigma_1(\mathbf{a}) - \varepsilon)\delta/2$, then there exists a finite set K_0 such that for every

$K \supset K_0$ (3.14) holds. Furthermore, for sufficiently large but fixed K and the stated h, the first term of (3.13) is smaller than ε for large n.

On the other hand, the first and the second terms of the right-hand side of (3.12) are small for large K and large n. Therefore, for an appropriate h and sufficiently large but fixed finite set K, there exists an $n_0 = n_0(h, K)$ such that for every $n \geqq n_0$ (3.12) is bounded by 5ε. This completes the proof of (ii).

(iii) From (2.6), (3.4) and Schwarz inequality, it sufficies to show that each of

$$A_2 = \sum_{ij} a_{ij}^2 |p_{i\bullet} - \tau_{i\bullet}/n| \times |p_{\bullet j} - \tau_{\bullet j}/n| \,,$$
$$A_3 = \sum_{ij} a_{ij}^2 p_{i\bullet} |p_{\bullet j} - \tau_{\bullet j}/n| \qquad \text{and} \qquad A_4 = \sum_{ij} a_{ij}^2 p_{\bullet j} |p_{i\bullet} - \tau_{i\bullet}/n|$$

converges to zero in probability as $n \to \infty$.

Put $B_{ij} = a_{ij}^2 |p_{i\bullet} - \tau_{i\bullet}/n| \times |p_{\bullet j} - \tau_{\bullet j}/n|$ and take a sufficiently large but finite index set K so that $\sum_{(i,j)\notin K} a_{ij}^2 p_{i\bullet} p_{\bullet j} < \varepsilon^2$ for a small $\varepsilon > 0$. Then

$$P(\textstyle\sum_{ij} B_{ij} > \varepsilon) \leqq P(\sum_{(i,j)\in K} B_{ij} > \varepsilon/2) + P(\sum_{(i,j)\notin K} B_{ij} > \varepsilon/2)\,. \tag{3.15}$$

Since K is a finite set and each B_{ij} converges to zero in probability, the first term of the right-hand side of (3.15) is smaller than ε provided n is sufficiently large. On the other hand the second term is bounded by

$$P(\textstyle\sum_{(i,j)\notin K} a_{ij}^2 p_{i\bullet} \tau_{\bullet j}/n > \varepsilon/8) + P(\sum_{(i,j)\notin K} a_{ij}^2 p_{\bullet j} \tau_{i\bullet}/n > \varepsilon/8)$$
$$+ P(\textstyle\sum_{(i,j)\notin K} a_{ij}^2 \tau_{i\bullet} \tau_{\bullet j}/n^2 > \varepsilon/8)$$

which is smaller than 24ε for $\varepsilon < \frac{1}{8}$. Thus, it is found that $A_2 \to_P 0$ as $n \to \infty$. The convergences of A_3 and A_4 can be shown similarly. This completes the proof.

4. The asymptotic distribution of the test statistic under contiguous alternatives.

Returning to the probability distribution $\{p_{ij}(\theta)\}$, let us postulate the following assumptions:

ASSUMPTION 4.1. The distribution $\{p_{i\bullet} p_{\bullet j}\} = \{p_{ij}(0)\}$ belongs to C_a defined by (3.2).

ASSUMPTION 4.2. There exists a positive number θ' such that the distribution $\{p_{ij}(\theta)\}$ is absolutely continuous with respect to $\{p_{i\bullet} p_{\bullet j}\}$ for $0 < \theta < \theta'$.

ASSUMPTION 4.3. The derivative $p'_{ij}(\theta) = (d/d\theta) p_{ij}(\theta)$ exists in an interval including $\theta = 0$ which is common for $(i, j) \in J \times J$.

ASSUMPTION 4.4. The convergence

$$\textstyle\sum_{ij} |p'_{ij}(\theta)| \to \sum_{ij} |p'_{ij}| < \infty \qquad \text{as } \theta \to 0$$

holds where $p'_{ij} = p'_{ij}(0)$.

ASSUMPTION 4.5. The Fisher information satisfies

$$I_\theta \equiv \textstyle\sum_{ij} [p'_{ij}(\theta)]^2 / p_{ij}(\theta) \to I_0 \equiv \sum_{ij} (p'_{ij})^2 / p_{i\bullet} p_{\bullet j} < \infty \qquad \text{as } n \to \infty\,.$$

ASSUMPTION 4.6. It holds that

$$p'_i \equiv \sum_j p'_{ij} = 0 \quad i \in J\,, \qquad p'_{\cdot j} \equiv \sum_i p'_{ij} = 0 \quad j \in J\,.$$

The asymptotic normality of $S_3(\mathbf{a})$ given by (3.1) is proved under $H_{\theta_n{}^{\mathbf{a}}}$: the common distribution of (X_k, Y_k)'s follows $\{p_{ij}(\theta_n)\}$, $\theta_n = n^{-\frac{1}{2}}\theta_0$ for some $\theta_0 > 0$.

THEOREM 4.1. *If Assumptions* 4.1 *through* 4.6 *are satisfied, then the convergence*

$$S_3(\mathbf{a}) \rightarrow N(\rho_1(\mathbf{a}), 1) \qquad \text{as } n \rightarrow \infty \tag{4.1}$$

holds under $H_{\theta_n{}^{\mathbf{a}}}$ *where*

$$\rho_1(\mathbf{a}) = \sigma_1^{-1}(\mathbf{a})\theta_0 \sum_{ij} a_{ij}p'_{ij}\,. \tag{4.2}$$

The proof of this theorem is based on the notion of contiguity introduced by Le Cam (1960) and elucidated in [8].

Let

$$L_n = \prod_{k=1}^{n} [p_{X_k Y_k}(\theta_n)/p_{X_k \cdot}p_{\cdot Y_k}]\,, \tag{4.3}$$

$$W_n = 2\sum_{k=1}^{n} \{[p_{X_k Y_k}(\theta_n)/p_{X_k \cdot}p_{\cdot Y_k}]^{\frac{1}{2}} - 1\} \tag{4.4}$$

and

$$T_n = \theta_n \sum_{k=1}^{n} p'_{X_k Y_k}/p_{X_k \cdot}p_{\cdot Y_k})\,. \tag{4.5}$$

To show Theorem 4.1, we need the following lemma which can be proved along the similar line of arguments as Lemma VI 2.1a and Theorem VI 2.1 of [8].

LEMMA 4.1. *Suppose Assumptions* 4.1 *through* 4.5 *are satisfied, then the convergences*

$$T_n \rightarrow_d N(0, \theta_0^2 I_0) \qquad \text{as } n \rightarrow \infty\,, \tag{4.6}$$

$$\log L_n - T_n + \theta_0^2 I_0/2 \rightarrow_P 0 \qquad \text{as } n \rightarrow \infty \tag{4.7}$$

and

$$\log L_n \rightarrow_d N(-\theta_0^2 I_0/2, \theta_0^2 I_0) \qquad \text{as } n \rightarrow \infty \tag{4.8}$$

hold under $H_{0\mathbf{a}}$.

In view of Le Cam's second lemma [8], the convergence (4.8) implies contiguity of $H_{\theta_n{}^{\mathbf{a}}}$ to $H_{0\mathbf{a}}$. Therefore, from Le Cam's third lemma [8], in order to prove Theorem 4.1 it suffices to show that $(S_3(\mathbf{a}), \log L_n)$ is asymptotically bivariate normal with covariance $\rho_1(\mathbf{a})$, under $H_{0\mathbf{a}}$.

PROOF OF THEOREM 4.1. From the arguments in Section 3 and (4.7), $(S_3(\mathbf{a}), \log L_n)$ is asymptotically equivalent in probability to $(\sigma_1^{-1}(\mathbf{a})n^{-\frac{1}{2}} \sum_{ij} a_{ij} V_{ij}, T_n - \theta_0^2 I_0/2)$. Therefore, it is sufficient to show that

$$an^{-\frac{1}{2}} \sum_{ij} a_{ij} V_{ij} + bT_n \rightarrow_d N(0, a^2\sigma_1^2(\mathbf{a}) + b^2\theta_0^2 I_0 + 2ab\sigma_1(\mathbf{a})\rho_1(\mathbf{a})) \tag{4.9}$$

as $n \rightarrow \infty$ for any but fixed real numbers a and b. Now, in view of Assumption 4.6, T_n is identical with $\theta_0 \sum_{ij} (p'_{ij}/p_{i\cdot}p_{\cdot j})V_{ij}$. Hence the asymptotic normality (4.9) can be shown by the same method in Lemma 3.1.

5. Some arguments on asymptotic relative efficiency. Let us first consider the situations in Sections 3 and 4. The asymptotic efficiency (AE) of the test using $S_3(\mathbf{a})$ is defined in [8] as

$$\text{AE}(S_3(\mathbf{a})) = (\textstyle\sum_{ij} a_{ij} p'_{ij})^2 / I_0 \sigma_1{}^2(\mathbf{a}) , \tag{5.1}$$

which is the square of the asymptotic correlation coefficient between $S_3(\mathbf{a})$ and the log-likelihood ratio (4.3). Assumption 4.6 entails that (5.1) is maximized and equal to one if and only if $a_{ij} = p'_{ij}/p_{i\cdot}p_{\cdot j}$. Therefore, the locally most powerful conditional test is an asymptotically most powerful test.

The ratio of the AE's of two competitive tests is their asymptotic relative efficiency (ARE). Consider the testing $H_{0\mathrm{a}} \cap H_{0\mathrm{b}}$, then the ARE of the test using $S_3(\mathbf{a})$ with respect to that using $S_3(\mathbf{b})$ is

$$\text{ARE}(S_3(\mathbf{a}), S_3(\mathbf{b})) = (\textstyle\sum_{ij} a_{ij} p'_{ij} / \sum_{ij} b_{ij} p'_{ij})^2 \times (\sigma_1{}^2(\mathbf{b})/\sigma_1{}^2(\mathbf{a})) .$$

Second, let us pay a little attention to rank tests. The ranks R_i and Q_i are defined by

$$R_i = \# \{k \mid X_k \leqq X_i\} , \qquad Q_i = \# \{k \mid Y_k \leqq Y_i\} .$$

and denote by $B_1 = (B_{11}, \cdots, B_{1\beta})$ and $B_2 = (B_{21}, \cdots, B_{2\gamma})$ the marginal vectors of ties of X's and Y's respectively.

Let $F(x, y; \theta)$ be a bivariate distribution function such that it is absolutely continuous with respect to $F(x, y; 0) = F_1(x)F_2(y)$ where F_1 and F_2 are distribution functions. Here $F(x, y; \theta)$ is not assumed to be continuous or discrete. Furthermore, assume that the Radon-Nikodym derivative $f(x, y; \theta) = dF(x, y; \theta)/dF(x, y; 0)$ has derivative $(\partial/\partial\theta)f(x, y; \theta) = h(x, y; \theta)$ and that $h(x, y; \theta)$ satisfies

$$\textstyle\int\int |h(x, y; \theta)| \, dF_1(x) \, dF_2(y) \to \int\int |h(x, y; 0)| \, dF_1(x) \, dF_2(y) < \infty \qquad \text{as } \theta \to 0 .$$

Define the ties-conditional (this word means 'on condition that B_1 and B_2 are given') scores by

$$a_n(i, j \mid B_1, B_2) = E(h(X^{(i)}, Y^{(j)}; 0) \mid B_1, B_2, F_1, F_2)$$

where $X^{(i)}$ $(Y^{(j)})$ is the ith (jth) order statistic among $(X_1, \cdots, X_n)$ $((Y_1, \cdots, Y_n))$ and where the ties-conditional expectation is calculated under $F_1(x)F_2(y)$. Then the ties-conditional rank test using

$$S_4 = \textstyle\sum_{i=1}^n a_n(R_i, Q_i \mid B_1, B_2)$$

is the locally most powerful ties-conditional rank test to test H_0 against the alternative associated with $F(x, y; \theta)$, $\theta > 0$.

Furthermore, let us again consider the sequence of distributions $\{p_{ij}(\theta)\}$ such that $p_{ij}(0) = p_{i\cdot}p_{\cdot j}$ and consider the scores constants $b_n(i, j)$ satisfying

$$\textstyle\int_0^1 \int_0^1 [b_n(1 + [un], 1 + [vn]) - \varphi(u, v)]^2 \, du \, dv \to 0 \qquad \text{as } n \to \infty$$

for a nonconstant square integrable function $\varphi(u, v)$ defined on (0.1) x (0.1). Let

$$S_5(\varphi) = n^{-\frac{1}{2}} \textstyle\sum_{k=1}^n b_n(R_k, Q_k \mid B_1, B_2)$$

be an averaged scores rank statistic where

$$b_n(i, j \mid B_1, B_2) = [(B_{1,k} - B_{1,k-1})(B_{2,k'} - B_{2,k'-1})]^{-1} \sum b_n(m, m') ,$$

$\sum$ being taken over the set of pairs (m, m') such that $B_{1,k-1} < m \leqq B_{1,k}$, $B_{2,k'-1} < m' \leqq B_{2,k'}$ for $B_{1,k-1} < i \leqq B_{1,k}$ and $B_{2,k'-1} < j \leqq B_{2,k'}$ and put

$$S_6(\varphi) = [\mathrm{Var}\,(S_5(\varphi) \mid B_1, B_2)]^{-\frac{1}{2}}[S_5(\varphi) - E(S_5(\varphi) \mid B_1, B_2)]$$

where ties-conditional mean and variance are calculated under H_0. Then Behnen [1] shows that under certain weak conditions

$$S_6(\varphi) \to N(0, 1) \quad \text{as} \quad n \to \infty \quad \text{under} \quad H_0$$

and

$$S_6(\varphi) \to N(\rho_2(\varphi), 1) \quad \text{as} \quad n \to \infty \quad \text{under} \quad H_{\theta_n} , \qquad \theta_n = n^{-\frac{1}{2}}\theta_0$$

where

$$\rho_2(\varphi) = \theta_0(\textstyle\sum_{ij} p_{i\bullet} p_{\bullet j} \varphi^2_{p_{i.} p_{.j}})^{-\frac{1}{2}} \sum_{ij} p'_{ij} \varphi_{p_{i.} p_{.j}}$$

for

$$\varphi_{p_{i.} p_{.j}} = (p_{i\bullet} p_{\bullet j})^{-1} \textstyle\iint_{I_{ij}} \varphi(u, v)\, du\, dv$$

with $I_{ij} = (\sum_{k=-\infty}^{i-1} p_{k\bullet}, \sum_{k=-\infty}^{i} p_{k\bullet}] \times (\sum_{k=-\infty}^{j-1} p_{\bullet k}, \sum_{k=-\infty}^{j} p_{\bullet k}]$. If the scores function φ is chosen so that $\varphi_{p_{i.} p_{.j}} = p'_{ij}/p_{i\bullet} p_{\bullet j}$, then the test using $S_6(\varphi)$ is asymptotically most powerful and its AE is equal to one. The locally most powerful ties-conditional rank test has the scores function stated above.

Finally, let us consider the case in which (X_k, Y_k)'s take a point in a finite set $\{(i, j); i = 1, \cdots, m_1\ j = 1, \cdots, m_2\}$. Suppose that $p_{ij}(\theta)$ is continuously differentiable with respect to θ and that if $p_{ij}(0) = p_{i\bullet} p_{\bullet j} = 0$ then $p_{ij}(\theta) = 0$ for $|\theta|$ smaller than a positive number θ'. The sequence of the alternatives to be considered is H_{θ_n} under which the common distribution of (X_k, Y_k)'s is given by $\{p_{ij}(\theta_n); \theta_n = n^{-\frac{1}{2}}\theta_0, \theta_0 \neq 0\}$. A conventional test for the two-sided situation is the chi-square test based on

$$\chi^2 = \textstyle\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} n(\tau_{ij} - \tau_{i\bullet} \tau_{\bullet j}/n)^2/\tau_{i\bullet} \tau_{\bullet j} .$$

Denote by $\chi_m{}^2(\Delta^2)$ a chi-square distribution with m degrees of freedom and noncentrality parameter Δ^2. Then the convergences

$$\chi^2 \to \chi^2_{(m_1-1)(m_2-1)}(0) \qquad \text{under} \quad H_0$$

and

$$\chi^2 \to \chi^2_{(m_1-1)(m_2-1)}(\theta_0'^2 I_0) \qquad \text{under} \quad H_{\theta_n}$$

hold where I_0 is given in Assumption 4.5. Furthermore, $S_3{}^2(\mathbf{a})$ and $S_6{}^2(\varphi)$ are asymptotically $\chi_1{}^2(0)$ under H_0 and are asymptotically $\chi_1{}^2(\rho_1{}^2(\mathbf{a}))$ and $\chi_1{}^2(\rho_2{}^2(\varphi))$ respectively under H_{θ_n}.

Now, we want to compare the test using $S_3{}^2(\mathbf{a})$ or $S_6{}^2(\varphi)$ with the chi-square test. In this case, simple determination of ARE is not possible because of the difference between their degrees of freedom. However, the notion of the local asymptotic efficiency defined in [8], page 271 can be utilized.

Let d be the asymptotic distance between the alternatives and the null hypothesis and let us consider two competitive tests using T_1 or T_2. Denote by $\beta_i(\alpha; d)$ $i = 1, 2$ the asymptotic power of the test T_i at level α.

DEFINITION 5.1. The local asymptotic efficiency at level α of the test T_i is defined by

$$e(T_i; \alpha) = (\partial/\partial d)[\beta_i(\alpha; d) - \alpha]|_{d=0}\,, \qquad i = 1, 2 \tag{5.1}$$

if it exists and the local asymptotic relative efficiency at level α is

$$e(T_1, T_2; \alpha) = e(T_1; \alpha)/e(T_2; \alpha)\,. \tag{5.2}$$

The efficiency (5.2) suggested in [8] is convenient when the functional form of the asymptotic distribution of T_1 is different from that of T_2. But it still depends on α. Thus it is more convenient to introduce the following.

DEFINITION 5.2. The local asymptotic relative efficiency of the test T_1 with respect to the test T_2 is defined by

$$e(T_1, T_2) = \lim_{\alpha\to 0} e(T_1, T_2; \alpha)\,. \tag{5.3}$$

In this paper we are interested in the situation when T_1 and T_2 are asymptotically chi-square.

THEOREM 5.1. *Suppose the asymptotic distance between the null hypothesis and the sequence of alternatives is represented by* θ_0^2 *and that* T_i *is asymptotically* $\chi^2_{k_i}(0)$ *under the null hypothesis and* $\chi^2_{k_i}(\theta_0^2\delta_i^2)$ *under the alternatives,* $i = 1, 2$. *Then*

$$e(T_1, T_2; \alpha) = \frac{\delta_1^2}{\delta_2^2} \times \frac{1 - \alpha - \int_0^{\alpha_1} f_{k_1+2}(x; 0)\,dx}{1 - \alpha - \int_0^{\alpha_2} f_{k_2+2}(x; 0)\,dx} \tag{5.4}$$

and

$$e(T_1, T_2) = k_2\delta_1^2/k_1\delta_2^2\,, \tag{5.5}$$

where $f_m(x; \Delta^2)$ *denotes the density function of* $\chi_m^2(\Delta^2)$ *and* $\alpha_i = \alpha_i(\alpha)$ *is determined by*

$$\alpha = \int_{\alpha_i}^{\infty} f_{k_i}(x; 0)\,dx\,, \qquad i = 1, 2\,. \tag{5.6}$$

PROOF. The asymptotic power of the test T_i at level α is

$$\beta_i(\alpha; \theta_0^2) = 1 - \int_0^{\alpha_i} f_{k_i}(x; \theta_0^2\delta_i^2)\,dx\,, \qquad i = 1, 2\,. \tag{5.7}$$

Hence, (5.4) follows by a short calculation.

Since the derivative of $\alpha_i(\alpha)$ is given by

$$\alpha_i'(\alpha) = -1/f_{k_i}(\alpha_i; 0) \tag{5.8}$$

and since $\alpha_i \to \infty$ as $\alpha \to 0$, it follows from (5.4) that

$$\begin{aligned} e(T_1, T_2) &= \lim_{\alpha\to 0} (\delta_1^2/\delta_2^2)[-1 - f_{k_1+2}(\alpha_1; 0)\alpha_1'(\alpha)]/[-1 - f_{k_2+2}(\alpha_2; 0)\alpha_2'(\alpha)] \\ &= (k_2\delta_1^2/k_1\delta_2^2) \lim_{\alpha\to 0} \alpha_1/\alpha_2\,. \end{aligned}$$

Furthermore we have $\lim_{\alpha\to 0} \alpha_1/\alpha_2 = \lim_{\alpha\to 0} \alpha_1'(\alpha)/\alpha_2'(\alpha)$. Hence, from (5.8), it remains to show that

$$\lim_{\alpha\to 0} f_{k_2}(\alpha_2; 0)/f_{k_1}(\alpha_1; 0) = 1 . \tag{5.9}$$

From (5.6), it follows that

$$\alpha = 2f_{k_i}(\alpha_i; 0) + (k_i - 2) \int_{\alpha_i}^{\infty} x^{-1} f_{k_i}(x; 0)\, dx \tag{5.10}$$

and the second terms of the right-hand side of (5.10) is $0(\alpha/\alpha_i)$. Hence

$$\alpha^{-1} f_{k_i}(\alpha_i; 0) = \tfrac{1}{2} + O(\alpha_i^{-1})$$

and consequently (5.9) follows. This completes the proof.

Theorem 5.1 implies

$$e(S_3^2(\mathbf{a}), \chi^2) = (m_1 - 1)(m_2 - 1)(\textstyle\sum_{ij} a_{ij} p'_{ij})^2 / I_0 \sigma_1^2(\mathbf{a}) \tag{5.11}$$

and

$$e(S_6^2(\varphi), \chi^2) = (m_1 - 1)(m_2 - 1)(\textstyle\sum_{ij} a_{ij} \varphi_{p_{i.} p_{.j}})^2 / I_0 \textstyle\sum_{ij} p_{i\bullet} p_{\bullet j} \varphi^2_{p_i p_j} . \tag{5.12}$$

If we employ the locally most powerful conditional test or the locally most powerful ties-conditional test, then each of (5.11) and (5.12) is maximized and equal to $(m_1 - 1)(m_2 - 1)$. Thus, when m_1 or m_2 is not small, our test with appropriate constants a_{ij} or the rank test with appropriate scores function φ will assure us higher asymptotic power for the alternatives not too far from H_0 and for large n than the chi-square test does.

Acknowledgment. The author wishes to thank Professor M. Okamoto at Osaka University for his encouragement and valuable comments. He is also grateful to Professor K. Behnen for his sending his papers [1] and [2] and to the referees for their detailed comments.

REFERENCES

[1] Behnen, K. (1973). Asymptotic properties of averaged scores rank tests of independence. *Proc. Prague Symposium on Asymptotic Statist.* 5–30.

[2] Behnen, K. (1976). Asymptotic comparison on rank tests for the regression problem when ties are present. *Ann. Statist.* **4** 157–174.

[3] Bhuchongkul, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Ann. Math. Statist.* **35** 138–149.

[4] Burr, E. J. (1960). The distribution of Kendall's score *S* for a pair of tied rankings. *Biometrika* **47** 151–171.

[5] Conover, W. J. (1973). Rank tests for one sample, two samples, and *k* samples without the assumption of a continuous distribution function. *Ann. Statist.* **1** 1105–1125.

[6] Cureton, E. E. (1958). The average Spearman rank criterion correlation when ties are present. *Psychometrika* **23** 271–272. Correction (1965) **30** 377.

[7] ·Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika* **47** 307–323.

[8] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

[9] Jogdeo, K. (1968). Asymptotic normality in nonparametric methods. *Ann. Math. Statist.* **39** 905–922.

[10] Le Cam, L. (1960). Locally asymptotically normal families of distributions. *Univ. Calif. Publ. Statist.* **3** 37–98.

[11] LYERLY, S. B. (1952). The average Spearman rank correlation coefficient. *Psychometrika* **17** 421-428.

[12] ROBILLARD, P. (1972). Kendall's S distribution with ties in one ranking. *J. Amer. Statist. Assoc.* **67** 453-455.

[13] RUYMGAART, F. H. (1974). Asymptotic normality of nonparametric tests for independence. *Ann. Statist.* **2** 892-910.

[14] RUYMGAART, F. H., SHORACK, G. R. and VAN ZWET, W. R. (1972). Asymptotic normality of nonparametric tests for independence. *Ann. Math. Statist.* **43** 1122-1135.

[15] SHIRAHATA, S. (1974a). Locally most powerful rank tests for independence. *Bull. Math. Statist.* **16** 11-21.

[16] SHIRAHATA, S. (1974b). On tests of symmetry for discrete populations. *Austral J. Statist.* **16** 83-90.

[17] SHIRAHATA, S. (1975). Locally most powerful rank tests for independence with censored data. *Ann. Statist.* **3** 241-245.

[18] SILLITTO, G. P. (1947). The distribution of Kendall's τ coefficient of rank correlation in rankings containing ties. *Biometrika* **34** 36-40.

[19] VORLICKOVÁ, D. (1970). Asymptotic properties of rank tests under discrete distributions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **14** 275-289.

[20] VORLICKOVÁ, D. (1972). Asymptotic properties of rank tests of symmetry under discrete distributions. *Ann. Math. Statist.* **43** 2013-2018.

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
KYUSHU UNIVERSITY
HAKOZAKI, HIGASHI-KU
FUKUOKA 812
JAPAN