NUMERICAL COMPARISON OF IMPROVED METHODS OF TESTING IN CONTINGENCY TABLES WITH SMALL FREQUENCIES

NARIAKI SUGIURA* AND MASANORI ÔTAKE

(Received Nov. 6, 1967)

1. Introduction and summary

The significance levels of various tests for a general $c\times k$ contingency table are usually given by large sample theory. But they are not accurate for the one having small frequencies. In this paper, a numerical evaluation was made to determine how good the approximation of significance level is for various improved tests that have been developed by Nass [6], Yoshimura [9], Gart [2] etc. for $c\times k$ contingency table with small frequencies in some of cells. For this purpose we compared the significance levels of the various approximate methods (i) with those of one-sided tail defined in terms of exact probabilities for given marginals in 2×2 table; (ii) with those of exact probabilities accumulated in the order of magnitude of χ^2 statistic or likelihood ratio (=LR) statistic in 2×3 table mentioned by Yates [8]. In 2×2 table it is well known that Yates' correction gives satisfactory result for small cell frequencies and the other methods that we have not referred here, can be considered if we devote our attention only to 2×2 or $2\times k$ table. But we are mainly interested in comparing the methods that are applicable to a general $c\times k$ table. It appears that such a comparison for the various improved methods in the same example has not been made explicitly, even though these tests are frequently used in biological and medical research.

Our numerical experience shows the following facts. The approximate significance levels due to Gart [2] and the modified Dandekar's method are somewhat better than the others for 2×2 tables, though they are somewhat too small for exact levels in the range one to five percent. The method of Gart [2] is conservative for 2×3 tables, but the others are not always conservative. The corrected LR statistic by Yoshimura [9] is always a little better than the LR statistic. However, the approximate significance level of the LR statistic was always smaller

* This research was supported by the Sakko-kai Foundation.

than the others except for tables having comparatively large marginals. The other methods that we investigated give almost the same results as the standard χ^2 test. The method of Nass [6] proved not to give as good an approximation as he conjectured.

The following notation of observed frequencies for a $c\times k$ contingency table is used throughout this paper.

$$
\begin{array}{c|c}
X_{11}, X_{12}, \cdots, X_{1k} & X_{1\cdot} \\
X_{21}, X_{22}, \cdots, X_{2k} & X_{2\cdot} \\
\cdots\cdots\cdots\cdots & \cdots \\
X_{c1}, X_{c2}, \cdots, X_{ck} & X_{c\cdot} \\
\hline
X_{\cdot 1}, X_{\cdot 2}, \cdots, X_{\cdot k} & N
\end{array}
$$

2. Description of methods with numerical comparisons in Rao's example

Rao ([7], p. 202) considered the following 2×2 table, testing whether the city soldiers are more sociable than village soldiers. In this example,

	Sociable	Nonsociable	Total
City soldiers	13	4	17
Village soldiers	6	14	20
Total	19	18	37

he compared the various improved techniques for small frequencies with the exact method, that is, the one sided tail probability ($X_{12}=0\sim 4$) for fixed marginals is given by equation (1) and the approximate significance levels of the other methods described by him are shown in (2)$\sim$(5).

(1) exact probability$=0.0059$
(2) χ^2 test, $P(\chi^2>7.9435)/2=0.0024$
(3) Yates' correction, $P(\chi^2>6.1922)/2=0.0064$
(4) LR test, $P(-2\log\lambda>8.2811)/2=0.0020$
(5) Dandekar's correction, $P(\chi^2>6.1086)/2=0.0068$.

These results show that χ^2 method (2) and LR method (4) do not give satisfactory approximation and that Yates' correction (3) and Dandekar's correction (5) are fairly good. But Yates' correction as well as Dandekar's cannot be extended to a general $c\times k$ table. We shall investigate some other techniques applicable to $c\times k$ tables and giving better approximations in this example.

(6) Corrected LR statistic. Yoshimura [9] calculated the correc-

tion factor K for the LR statistic, $-2\log\lambda$, such that the first and the second conditional moments for given marginals of the statistic, $-2K\log\lambda$, are equal to those of the $\chi^2_{(c-1)(k-1)}$ distribution up to the order of $1/N$, that is,

$$-2K\log\lambda=-2K\Big[\sum_{i=1}^{c}\sum_{j=1}^{k}X_{ij}\log X_{ij}-\sum_{i=1}^{c}X_{i\cdot}\log X_{i\cdot}\\ -\sum_{j=1}^{k}X_{\cdot j}\log X_{\cdot j}+N\log N\Big]$$

where

$$K=1-[6N(c-1)(k-1)]^{-1}\Big(N\sum_{i=1}^{c}X_{i\cdot}^{-1}-1\Big)\Big(N\sum_{j=1}^{k}X_{\cdot j}^{-1}-1\Big).$$

Then he proposed using the statistic $-2K\log\lambda$, which is approximately distributed according to the $\chi^2_{(c-1)(k-1)}$ distribution. In Rao's example, $K=0.95906$ and $P(-2K\log\lambda>7.9421)/2=0.0024$. The accuracy of this correction gives the same significance level as that of the χ^2 method (2), though it is a little better than the LR method (4).

(7) Correction of χ^2 statistic. We shall determine the correction factors a and b such that the first and the second conditional moments of the statistic $a\chi^2+b$ are equal to those of $\chi^2_{(c-1)(k-1)}$ distribution. The conditional moments of the χ^2 statistic are given by Haldane [4].

$$E_1=E[\chi^2\mid X_{i\cdot},X_{\cdot j}]=N(N-1)^{-1}(c-1)(k-1)\,,$$

$$E_2=E[\chi^4\mid X_{i\cdot},X_{\cdot j}]=N^3[(N-1)(N-2)(N-3)]^{-1}\\ \cdot(A_0+A_1N^{-1}+A_2N^{-1})\,,$$

$$A_0=(c-1)^2(k-1)^2+2(c-1)(k-1)\,,$$

$$A_1=-4(c-1)^2(k-1)^2+c^2k^2+2ck-2-(k^2+2k-2)N\sum_{i=1}^{c}X_{i\cdot}^{-1}\\ -(c^2+2c-2)N\sum_{j=1}^{k}X_{\cdot j}^{-1}+N^2\Big(\sum_{i=1}^{c}X_{i\cdot}^{-1}\Big)\Big(\sum_{j=1}^{k}X_{\cdot j}^{-1}\Big)\,,$$

$$A_2=ck(c-2)(k-2)+k(k-2)N\sum_{i=1}^{c}X_{i\cdot}^{-1}+c(c-2)N\sum_{j=1}^{k}X_{\cdot j}^{-1}\\ +N^2\Big(\sum_{i=1}^{c}X_{i\cdot}^{-1}\Big)\Big(\sum_{j=1}^{k}X_{\cdot j}^{-1}\Big)\,.$$

Then we get $a=[2(c-1)(k-1)/(E_2-E_1^2)]^{1/2}$ and $b=(c-1)(k-1)-aE_1$, regarding the statistic $a\chi^2+b$ as a $\chi^2_{(c-1)(k-1)}$ variate. In Rao's example, $a=0.9864$, $b=0.01382$ and $P(a\chi^2+b>7.8218)/2=0.0026$. This correction is slightly better than the χ^2 method (2), though it is not satisfactory to us.

(8) Method by Nass I. Nass [6] proposes to use the test statistic

$a\chi^2$ as a χ_f^2 variate, where the quantities a and f are determined such that the first and the second conditional moments of the statistic $a\chi^2$ are equal to those of the χ_f^2 distribution, that is, $a=2E_1/(2E_2-E_1^2)$, $f=aE_1$ where E_1, E_2 are given by (7). This method differs from the others in that the test statistic has a χ^2 distribution with non-integer number of degrees of freedom. In order to get the tail probability for the χ_f^2 distribution, we made use of the Chebyshev series expansions for gamma functions mentioned by Clenshaw [1] and the continued fraction expansion

$$x^{-a}e^{x}\int_x^\infty e^{-x}x^{a-1}dx=\frac{1}{x+}\,\frac{1-a}{1+}\,\frac{1}{x+}\,\frac{2-a}{1+}\,\frac{2}{x+}\,\frac{3-a}{1+}\cdots\frac{n-1}{x+}\,\frac{n-a}{1+}\cdots$$

for $0<a<1$, which is discussed by Gupta and Waknis [3]. In Rao's example, $a=1.0000$, $f=1.0278$ and $P(a\chi_f^2>7.9439)/2=0.0025$. This result is almost the same as that for the χ^2 method (2) and method (7).

(9) Method by Nass II. Nass [6] conjectured that his method (8) would be improved, if he could use the test statistic, $a\chi^2+b$, as if it were distributed according to the χ_f^2 distribution, where the parameters a, b and f are determined such that the first three conditional moments of $a\chi^2+b$ are equal to those of the χ_f^2 distribution. In a $2\times k$ table we can check this conjecture by using the third moment given by Haldane [5]. We shall write it after some rearrangement.

$$E_3=E[\chi^6\mid X_{i\cdot},X_{\cdot j}]=N^5[(N-1)(N-2)\cdots(N-5)]^{-1}\cdot\left\{(k^2-1)(k+3)+\sum_{\alpha=1}^{5}c_\alpha N^{-\alpha}\right\},$$

$$c_1=2(30k^2+69k-43)-2(9k+47)N\sum_{j=1}^{k}X_{\cdot j}^{-1}+N^2(X_{1\cdot}X_{2\cdot})^{-1}\left\{-3k^3-21k^2-24k+26+(3k+19)N\sum_{j=1}^{k}X_{\cdot j}^{-1}\right\},$$

$$\begin{aligned}c_2=&120(3k-1)-360N\sum_{j=1}^{k}X_{\cdot j}^{-1}+120N^2\sum_{j=1}^{k}X_{\cdot j}^{-2}\\&+N^2(X_{1\cdot}X_{2\cdot})^{-1}\left\{5k^3-57k^2-266k+120+3(9k+67)N\sum_{j=1}^{k}X_{\cdot j}^{-1}-30N^2\sum_{j=1}^{k}X_{\cdot j}^{-2}\right\}\\&-N^4(X_{1\cdot}X_{2\cdot})^{-2}\left\{-2(k^3+9k^2+14k-12)+(3k+22)N\sum_{j=1}^{k}X_{\cdot j}^{-1}-N^2\sum_{j=1}^{k}X_{\cdot j}^{-2}\right\},\end{aligned}$$

$$c_3=N^2(X_1.X_2.)^{-1}\left\{60k(k-2)-30(k-6)N\sum_{j=1}^{k}X_{.j}^{-1}-90N^2\sum_{j=1}^{k}X_{.j}^{-2}\right\}$$
$$-N^4(X_1.X_2.)^{-2}\left\{6k(k-2)(k+5)+2(3k+23)N\sum_{j=1}^{k}X_{.j}^{-1}\right.$$
$$\left.-16N^2\sum_{j=1}^{k}X_{.j}^{-2}\right\},$$

$$c_4=-N^4(X_1.X_2.)^{-2}\left\{4k(k^2-4)-(21k-20)N\sum_{j=1}^{k}X_{.j}^{-1}-11N^2\sum_{j=1}^{k}X_{.j}^{-2}\right\},$$

$$c_5=-4N^4(X_1.X_2.)^{-2}\left\{(3k-4)N\sum_{j=1}^{k}X_{.j}^{-1}+N^2\sum_{j=1}^{k}X_{.j}^{-2}\right\}.$$

By using the moments E_3 together with E_1, E_2 given by (7), we can determine the three parameters as $a=4m_2/m_3$, $b=4m_2(2m_2^2-m_1m_3)/m_3^2$, $f=8m_2^3/m_3^2$, where $m_1=E_1$, $m_2=E_2-E_1^2$ and $m_3=E_3-3E_2E_1+2E_1^3$. In Rao's example, $a=1.0287$, $b=0.03034$, $f=1.0877$ and $P(a\chi_f^2+b>8.2022)/2=0.0024$. The result of this correction is again the same as that of the χ^2 method (2) and methods (6)~(8). This approximation method conjectured by Nass does not seem to be satisfactory too.

(10) Method by Gart I. Gart [2] derived the modified LR statistic M/d, as a χ^2 variate with $(c-1)(k-1)$ degrees of freedom from an interesting viewpoint of regarding the data value X_{ij} as parameters and cell probabilities as random variables based on the equality connecting the multinomial distribution with the multivariate Beta distribution.

$$M=\sum_{i=1}^{c}\sum_{j=1}^{k}(2X_{ij}+1)\log(2X_{ij}+1)-\sum_{i=1}^{c}(2X_{i.}+k)\log(2X_{i.}+k)$$
$$-\sum_{j=1}^{k}(2X_{.j}+c)\log(2X_{.j}+c)+(2N+ck)\log(2N+ck),$$

$$d=1+\frac{1}{3(c-1)(k-1)}\left\{\sum_{i=1}^{c}\sum_{j=1}^{k}\frac{1}{2X_{ij}+1}-\sum_{i=1}^{c}\frac{1}{2X_{i.}+k}\right.$$
$$\left.-\sum_{j=1}^{k}\frac{1}{2X_{.j}+c}+\frac{1}{2N+ck}\right\}.$$

We shall remark that this correction factor d contains random variables X_{ij}, which distinguishes this method from others. In Rao's example, $M=7.8096$, $d=1.0565$ and $P(M/d>7.3920)/2=0.0033$. This result is somewhat better than the previous ones.

(11) Method by Gart II. Gart [2] also proposed to use the more accurate correction factor d' instead of d in the method (10).

$$d'=\frac{1}{(c-1)(k-1)}\left[\sum_{i=1}^{c}\sum_{j=1}^{k}\left\{1-\frac{1}{3(2X_{ij}+1)}+\frac{1}{8(2X_{ij}+1)^2}\right\}^{-1}\right.$$
$$-\sum_{i=1}^{c}\left\{1-\frac{1}{3(2X_{i.}+k)}+\frac{1}{8(2X_{i.}+k)^2}\right\}^{-1}$$

$$-\sum_{j=1}^{k}\left\{1-\frac{1}{3(2X_{.j}+c)}+\frac{1}{8(2X_{.j}+c)^2}\right\}^{-1}$$
$$+\left\{1-\frac{1}{3(2N+ck)}+\frac{1}{8(2N+ck)^2}\right\}^{-1}\Bigg].$$

In Rao's example, $d'=1.0562$ and $P(M/d'>7.3944)/2=0.0033$. The result of this correction is the same as that of the method (10). The differences between methods (10) and (11) will be seen in another example later.

(12) Modified Dandekar's method. Dandekar's correction mentioned in Rao ([7], p. 203) cannot be extended to a general $c\times k$ table. This is because for given marginals we cannot uniquely increase or decrease the observed frequency by 1 according as the minimum observed frequency is decreased or increased by 1. We shall consider the following modification suggested by Mr. Ueda. Let the minimum number of all the cell frequencies be attained by the $(i_1, j_1), \cdots, (i_l, j_l)$ cells. Then we calculate the χ^2 statistic for the table having the modified frequency X'_{rs}, where $X'_{rs}=X_{rs}\pm l^{-1}$, $X'_{r.}=X_{r.}\pm$(number of $i_1, \cdots, i_l$ equal to $r)\times l^{-1}$, $X'_{.s}=X_{.s}\pm$(number of $j_1, \cdots, j_l$ equel to $s)\times l^{-1}$ for $(r, s)=(i_1, j_1), \cdots, (i_l, j_l)$, $X'_{rs}=X_{rs}$ for other cells and $N'=N\pm 1$. We shall define this value as $\chi^2_{\mp 1}$. The modified Dandekar's method is to use the statistic $\chi^2_c=\chi^2_0-|(\chi^2_0-\chi^2_{-1})(\chi^2_{+1}-\chi^2_0)/(\chi^2_{+1}-\chi^2_{-1})|$ as a χ^2 variate with $(c-1)(k-1)$ degrees of freedom, where χ^2_0 means the value of χ^2 statistic for the observed frequency X_{ij}. Here we must note that the inequalities $\chi^2_{+1}>\chi^2_0>\chi^2_{-1}$ are not necessarily hold. In fact, for $X_{12}=9\sim 12$ in figure 4 we have $\chi^2_{+1}<\chi^2_0<\chi^2_{-1}$. In Rao's example, $\chi^2_0=7.9435$, $\chi^2_{+1}=9.3678$, $\chi^2_{-1}=6.7556$ and $P(\chi^2_c>7.2958)/2=0.0035$. This result is slightly better than Gart's methods (10) and (11).

3. Comparison in some 2×2 and 2×3 tables

In the previous section we have checked the accuracy for the various approximation methods in Rao's example only. We shall first investigate the effects of these corrections for $X_{12}=0\sim 7$ or $X_{11}=0\sim 8$ (accumulated in opposite direction) in the same example, whose results are shown in figures 1 and 2. Then we also examine in case of other tables as follows:

(i) Rao's example (Figures 1 and 2)

X_{11}	X_{12}	17
X_{21}	X_{22}	20
19	18	37

,

(ii) 2×2 symmetric case (Figure 3)

X_{11}	X_{12}	20
X_{21}	X_{22}	20
20	20	40

,

(iii) 2×2 table with large marginals (Figure 4)

X_{11}	X_{12}	18
X_{21}	X_{22}	555
133	440	573

,

(iv) 2×3 table by Yates ([8], p. 233) (Figures 5 and 6)

X_{11}	X_{12}	X_{13}	17
X_{21}	X_{22}	X_{23}	13
13	11	6	30

.

Figures 1 and 2 show the following facts. Gart (10), (11) and modified Dandekar (12) give fairly better results and the LR method (4) always gives the smallest values. Yoshimura (6) is uniformly better than (4). The methods (7)~(9) are almost the same as the χ^2 method (2). We must note in figure 3 that the value obtained by Gart (10) for $X_{12}=0$ is larger than for $X_{12}=1$. This fact seems to be a weak point of method (10). Gart (11) is preferable to Gart (10) in the presence of zero frequencies for some cells; otherwise they are nearly equal. Figure 4 shows that the LR method (4) and Yoshimura (6) are better and they are almost equal to Gart (10) and (11). In this case, the χ^2 method (2) as well as modified Dandekar's method (12) do not give good approximations. In figures 5 and 6, the methods (2) and (7)~(9) give almost the same results as in the previous figures, but they are close to exact method (1). Gart (10) and (11) are conservative, but modified Dandekar (12) is not.

Acknowledgement

The authors wish to express their gratitude to Mr. S. Ueda, Ministry of Foreign Affairs and to Dr. C. E. Land, ABCC, for their valuable suggestions. They also thank Miss Futagami, Computing Laboratory, Hiroshima University for her computational help.

UNIVERSITY OF NORTH CAROLINA AND HIROSHIMA UNIVERSITY
ATOMIC BOMB CASUALTY COMMISSION

REFERENCES

[1] C. W. Clenshaw, *Mathematical Tables 5, Chebyshev Series for Mathematical Function*, Her Majesty's stationary office, National Physical Laboratory, 1962.

[2] John J. Gart, "Alternative analyses of contingency tables," *J. R. Statist. Soc. B*, 28 (1966), 164-179.

[3] Shanti S. Gupta and Mrudulla N. Waknis, "A system of inequalities for the incomplete gamma function and the normal integral," *Ann. Math. Statist.*, 36 (1965), 139-149.

[4] J. B. S. Haldane, "The mean and variance of χ^2, when used as a test of homogeneity, when expectations are small," *Biometrika*, 31 (1940), 346-355.

[5] J. B. S. Haldane, "The use of χ^2 as a test of homogeneity in a $(n\times 2)$ fold table when expectations are small," *Biometrika*, 33 (1945), 234-238.

[6] C. A. G. Nass, "The χ^2 test for small expectations in contingency tables, with special reference to accidents and absenteeism," *Biomerika*, 46 (1959), 365-385.

[7] C. R. Rao, *Advanced Statistical Methods in Biometric Research*, John Wiley and Sons Inc., 1952.

[8] F. Yates, "Contingency tables involving small numbers and the χ^2 test," *Suppl. J. R. Statist. Soc.*, 1 (1934), 217-235.

[9] I. Yoshimura, "Scale correction of likelihood ratio test statistics related to contingency tables," *Rep. Statist. Appl. Res.*, 10 (1963), 1-10.

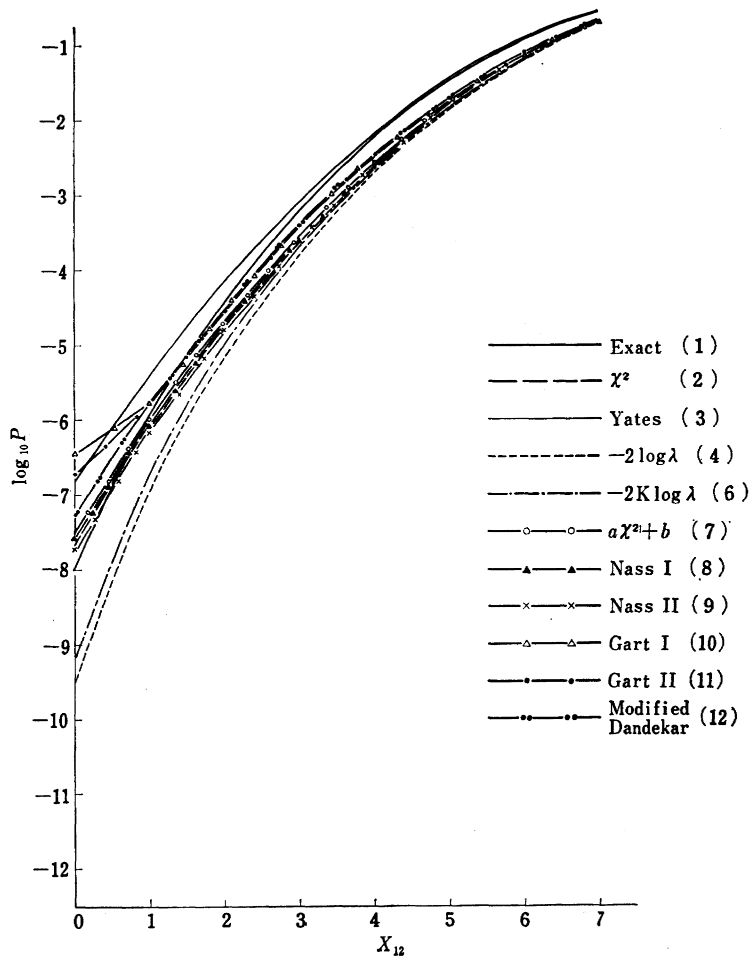

Fig. 1. One sided tail probabilities for $X_{12}=0\sim8$ in Rao's example.

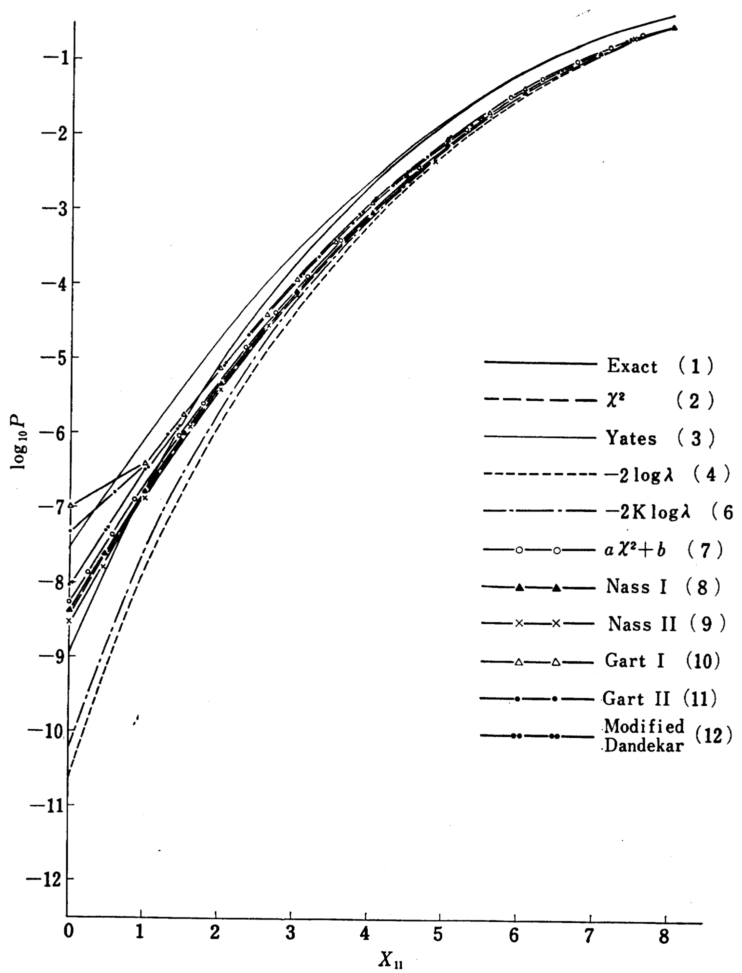

Fig. 2. One sided tail probabilities for $X_{11}=0\sim 8$ in Rao's example.

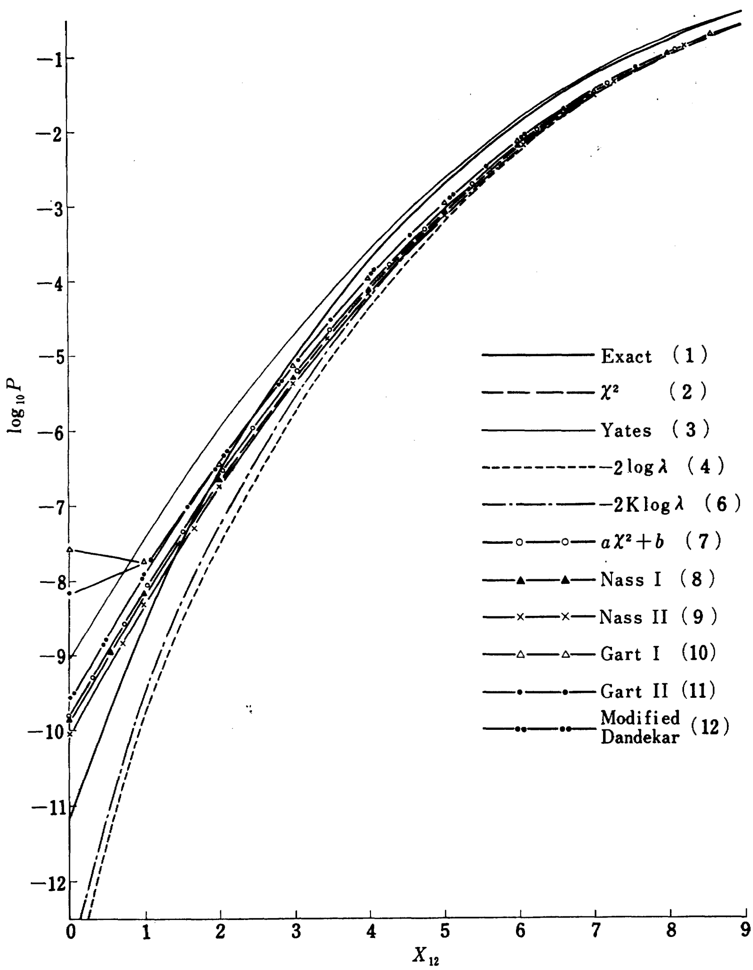

Fig. 3. One sided tail probabilities for $X_{12}=0\sim9$ in symmetric case.

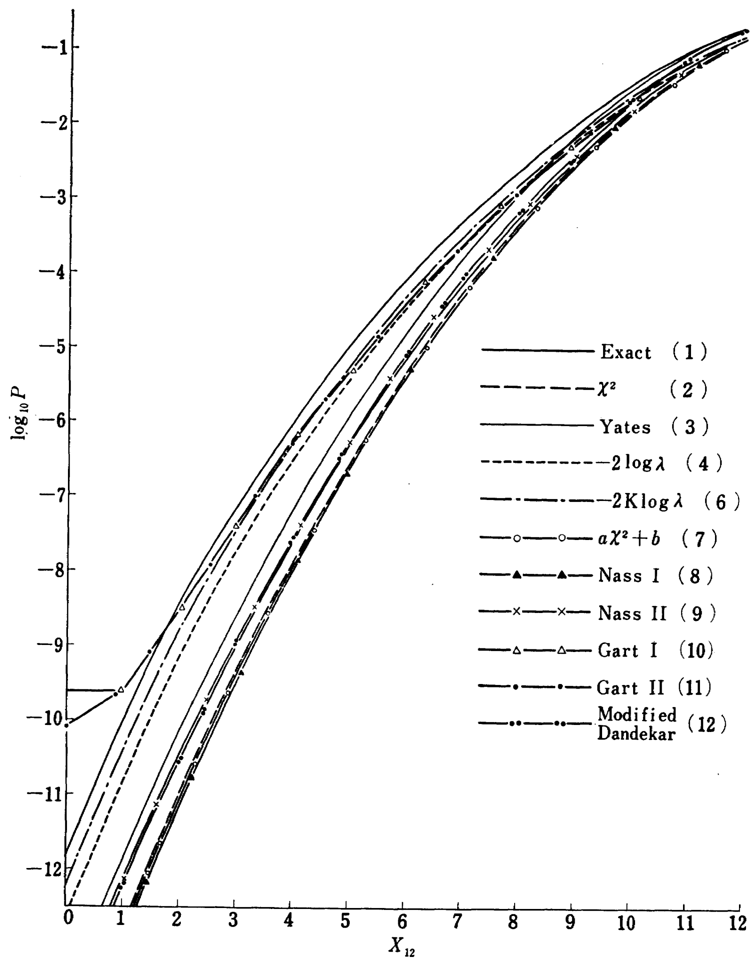

Fig. 4. One sided tail probabilities for $X_{12}=0\sim12$ with large marginals.

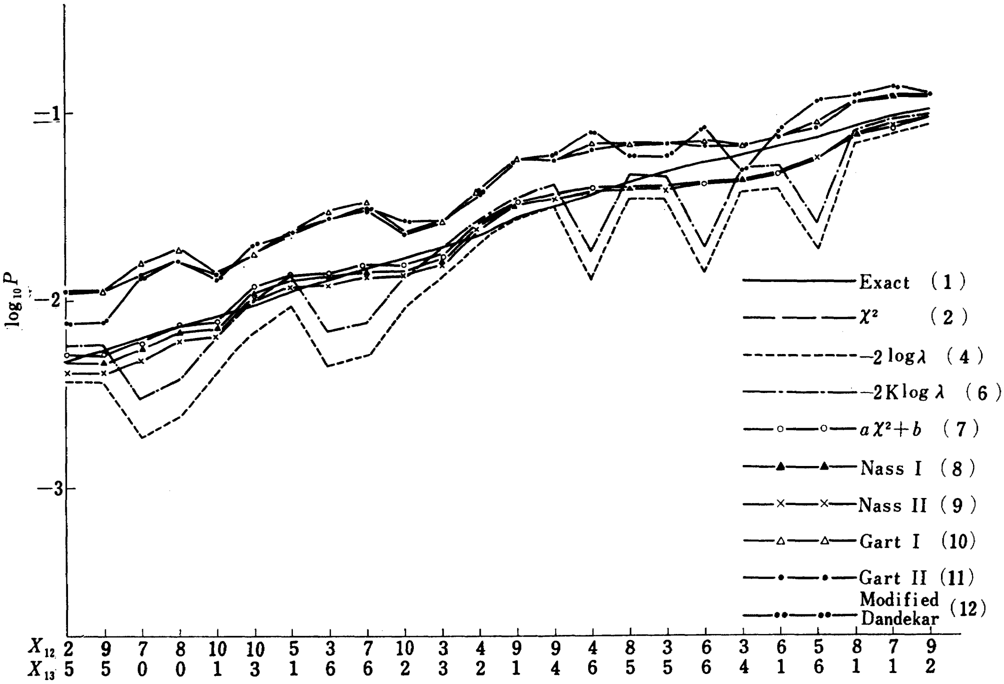

Fig. 5. Tail probabilities in 2×3 table accumulated by the magnitude of χ^2 statistic.

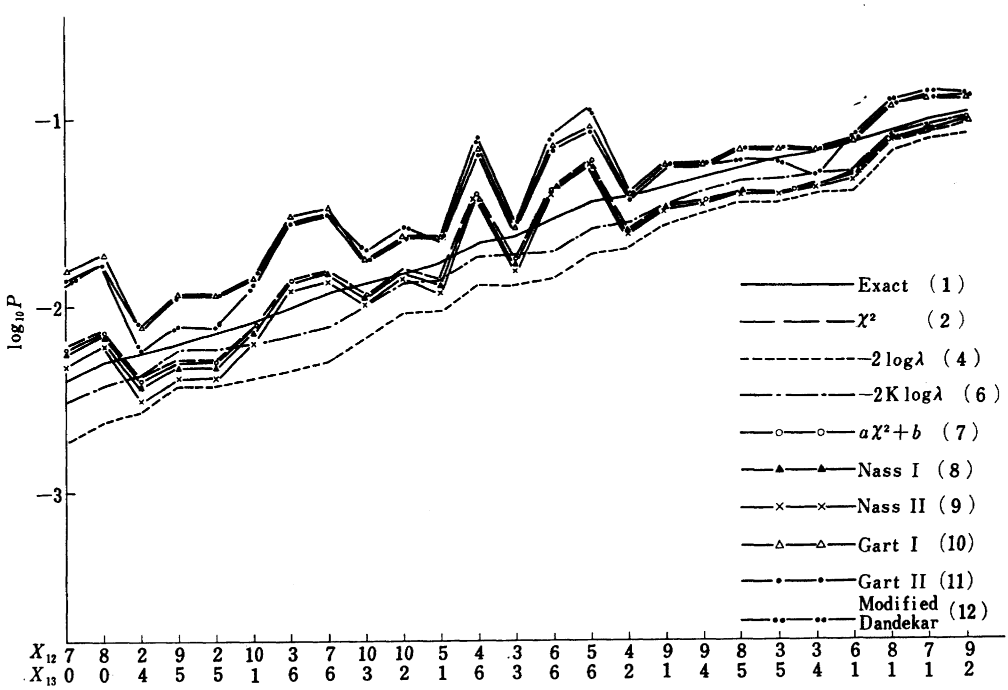

Fig. 6. Tail probabilities in 2×3 tables accumulated by the magnitude of *LR* statistics.