TESTING NONLINEAR TREND IN PROPORTIONS UNDER BINOMIAL AND EXTRA-BINOMIAL VARIABILITIES

Astuti, Tri Erni
Central Bureau of Statistics | Graduate School of Mathematics, Kyushu University

Yanagawa, Takashi
Graduate School of Mathematics, Kyushu University

https://doi.org/10.5109/13499

出版情報：Bulletin of informatics and cybernetics. 32 (2), pp.157-168, 2000-12. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics, Vol. 32, No. 2, December, 2000

TESTING NONLINEAR TREND IN PROPORTIONS UNDER BINOMIAL AND EXTRA-BINOMIAL VARIABILITIES

By

Erni TRI ASTUTI * and **Takashi YANAGAWA**†

Abstract

Trend tests for nonlinear response in proportion are proposed under binomial and extra-binomial variabilities for dose-response developmental toxicology data. The proposed test under binomial variability is constructed by applying orthonormal dose vector to score test under logistic model, and the test for extra binomial variability is constructed by applying the same idea to generalized score test. It is shown that the test for data with binomial variability is asymptotically equivalent to the Q_r test proposed by Jayasekara et al. (1999); and that the proposed test for data with overdispersion has higher powers than the generalized Cochran-Armitage test for detecting nonlinear response. It is also shown by simulation that the trend tests for binomial variability should not be used when the data exhibit extra binomial variability since it inflates the type I error substantially.

Key Words and Phrases: beta-binomial distribution, Cochran-Armitage test, developmental toxicology, generalized score test, orthonormal dose vector, score test.

1. Introduction

An interest in a dose-response experiment is in testing the trend of mean responses over increasing dose levels. For proportion which we deal with in this paper the Cochran-Armitage (C-A) test (Cochran, 1954: Armitage, 1955) has been applied for detecting the dose-response. However, sometimes environmental data exhibit a nonmonotone or nonlinear response. For example, the data may respond in an increasing fashion over low doses, but then has a downturn in higher doses. If this is the case the C-A test could lose powers for detecting the dose-response.

Using the idea of Jayasekara et al. (1999) that developed the Q_r test for testing nonlinear response in 2xk contingency tables, we propose in this paper trend tests for proportion which has high power in detecting nonlinear response under binomial and extra-binomial variabilities. The proposed test are constructed by applying orthonormal dose vector to score test (Rao, 1947), or generalized score test (Boos, 1992), under rth order logistic model. It is shown that the proposed test under binomial variability is asymptotically equivalent to the Q_r test. When $r = 1$, that is, for simple logistic model, the test is equivalent to the conventional C-A test, and to the generalized C-A test (Carr and Gorelick, 1995) in the case of overdispersion. Carr and Gorelick (1995) show that the generalized C-A test can control the Type I error, but C-A test can not in the presence of

* Graduate School of Mathematics, Kyushu University, and Central Bureau of Statistics, Indonesia
† Graduate School of Mathematics, Kyushu University , yanagawa@math.kyushu-u.ac.jp

overdispersion. We focus in this paper on those tests constructed when $r = 2$, and study their behaviour for detecting nonlinear response. It is shown by simulation that the test for binomial variability has higher power than the C-A test for detecting nonlinear trend; that the test for overdispersion has higher power than the generalized C-A test for detecting nonlinear response; and that the trend tests for binomial variability should not be used when the data exhibit extra-binomial variability since it inflates the type I error substantially.

2. Trend Tests

Consider a dose response experiment, where m_i is number of dam (experimental units) that are administered to dose d_i and n_{ij} is number of litters from jth dam ($i = 1, 2, \cdots, k; j = 1, 2, \cdots, m_i$). Denote by Y_{ij}/n_{ij} the proportion of fetuses that has specific response. Assume that dams are independent and $d_1 < d_2 < \cdots < d_k$. Let π_i represents the response probability at d_i. Our goal is to test

$$H_0 : \pi_1 = \pi_2 = \cdots = \pi_k$$

against H_1 that postulates a linear or nonlinear trend in π's.

2.1. Orthonormal Dose vector

Let a dot in subscript denote the summation over that subscript, hence $n_{i\cdot} = \sum_j n_{ij}, n_{\cdot\cdot} = \sum_i \sum_j n_{ij}, Y_{i\cdot} = \sum_j Y_{ij}$ and so on. Consider $\mathbf{d} = (d_1, d_2, \cdots, d_k)'$ as dose level vector. Then define:

$$\mathbf{c_1} = (c_1, c_2, \cdots, c_k)'$$

where $c_i = d_i - \bar{d}$ and $\bar{d} = \sum_i d_i n_{i\cdot}/n_{\cdot\cdot}$, so that $\sum_i c_i n_{i\cdot} = 0$. Also define:

$$\mathbf{c_s} = (c_{s1}, c_{s2}, \cdots, c_{sk})'$$

where $c_{si} = c_i^s$ (sth power of c_i) for $s = 1, 2, \cdots, r$. In particular $\mathbf{c_0} = (1, 1, \cdots, 1)'$. Next define inner product of two vector as $(\mathbf{a}, \mathbf{b}) = \sum_i a_i b_i n_{i\cdot}$ and $\|\mathbf{a}\|^2 = (\mathbf{a}, \mathbf{a})$. It is obvious that $\mathbf{c_0}, \mathbf{c_1}, \cdots, \mathbf{c_r}$ are linearly independent. Let $\mathbf{a_0}, \mathbf{a_1}, \cdots, \mathbf{a_r}$ be orthonormal vectors obtained by applying Gram-Schimdt orthonormalization to these vector, that is

$$\mathbf{a_0} = \frac{\mathbf{c_0}}{\|\mathbf{c_0}\|}$$

$$\mathbf{d_s^*} = \mathbf{c_s} - \sum_{h=0}^{s-1} (\mathbf{c_s}, \mathbf{a_h})\mathbf{a_h}$$

$$\mathbf{a_s} = \frac{\mathbf{d_s^*}}{\|\mathbf{d_s^*}\|}$$

so we have

$$(\mathbf{a_s}, \mathbf{a_l}) = \begin{cases} 0 & \text{if } s \neq l \\ 1 & \text{if } s = l \end{cases}$$

and $\|\mathbf{a_s}\| = 1$ for all $s = 0, 1, \cdots, r$.

2.2. Trend Test under Binomial Variability

Suppose that $\{Y_{ij}\}_{j=1,\cdots,m_i}$ are independent and Y_{ij} follows binomial distribution $B(n_{ij},\pi_i), i = 1,\cdots,k, j = 1,\cdots,m_i$. For testing H_0 vs H_1 we propose a trend tests based on statistics

$$T_{S_r} = \sum_{s=1}^{r} \frac{(\sum_i a_{si} Y_{i\cdot})^2}{\bar{Y}(1-\bar{Y})}, \qquad r = 1, 2, \cdots$$

where $\bar{Y} = Y_{\cdot\cdot}/n_{\cdot\cdot}$ and a_{si} is the ith element of the orthonormal vector $\mathbf{a}_s$. It may be shown similarly as Jayasekara et al. (1999) that under H_0 the summands of T_{S_r} are mutually uncorrelated and that T_{S_r} follows a chi-square distribution with r degree of freedom, asymptotically. When $r = 1$, T_{S_r} is written as

$$S1 = \frac{(\sum_i a_{1i} Y_{i\cdot})^2}{\bar{Y}(1-\bar{Y})}$$

which is easily shown equivalent to the statistics of the C-A test. When $r = 2$ T_{S_r} is written as

$$S2 = \sum_{s=1}^{2} \frac{(\sum_i a_{si} Y_{i\cdot})^2}{\bar{Y}(1-\bar{Y})}.$$

We call the test based on this statistics the S2 test. As $n_{\cdot\cdot} \to \infty$, S2 test is shown equivalent to the Q_2 test by Jayasekara et al. (1994), asymptotically.

T_{S_r} is the test statistics of a score test, which will be shown as follows; Considering the rth order logistic model ($r < k$) defined by:

$$\pi_i = \frac{1}{1 + exp(-\sum_{s=0}^{r} \beta_s a_{si})}$$

with the orthonormal dose vector $\mathbf{a}_s = (a_{s1}, a_{s2}, \cdots, a_{sk})'$, and putting $\boldsymbol{\beta}_{(2)} = (\beta_1, \cdots, \beta_r)'$, the problem of trend test is formulated as testing problem of $H_0^* : \boldsymbol{\beta}_{(2)} = \mathbf{0}$ against $H_1^* : \boldsymbol{\beta}_{(2)} \neq \mathbf{0}$. Then we have

PROPOSITION 2.1. *T_{S_r} is the statistic of the score test for testing H_0^* against H_1^**

PROOF. The log likelihood function is represented by,

$$\ell(\boldsymbol{\beta}) = c + \sum_i \sum_j y_{ij} (\sum_{s=0}^{r} \beta_s a_{si}) - \sum_i \sum_j n_{ij} log(1 - (1 + exp\{-\sum_{s=0}^{r} \beta_s a_{si}\})^{-1})$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_r)$ and c is a constant term. Thus the proof of the proposition is immediate.

2.3. Trend Test under Extra Binomial Variability

Boos(1992) introduced the generalization of score test that able to account for certain model inadequacies or lack of knowledge by use of empirical variance estimates.

We apply his idea to obtain trend test when data are overdispersed. For orthonormal dose vector $\mathbf{a}_s = (a_{s1}, a_{s2}, \cdots, a_{sk})'$, $s = 1, 2, \cdots, r$, let

$$\tilde{S}_{(2)} = \left(\sum_i a_{1i} Y_{i\cdot}, \cdots, \sum_i a_{ri} Y_{i\cdot} \right)'$$

and

$$\tilde{D}_{Y(22)} = \left(\sum_i \sum_j (Y_{ij} - n_{ij}\bar{Y})^2 a_{ti} a_{ui} \right)_{r \times r}$$

For testing $H_0 : \pi_1 = \cdots = \pi_k$, we propose a test based on statistic

$$T_{GSr} = \tilde{S}'_{(2)} \tilde{D}^{-1}_{Y(22)} \tilde{S}_{(2)}$$

It may be shown that T_{GSr} follows chi-square distribution with r degree of freedom under H_0, asymptotically. It is shown in Appendix, that T_{GSr} is the statistic of the generalized score test for testing $H_0^* : \boldsymbol{\beta}_{(2)} = \mathbf{0}$ vs $H_1^* : \boldsymbol{\beta}_{(2)} \neq \mathbf{0}$.

When $r = 1$, T_{GSr} is written as

$$GS1 = \frac{(\sum_i a_{1i} Y_{i\cdot})^2}{\sum_i a_{1i}^2 \sum_j (Y_{ij} - n_{ij}\bar{Y})^2}$$

which is easily shown equivalent to the generalized C-A test (Carr and Gorelick , 1995), thus has high power in detecting linear trend.

When $r = 2$ T_{GSr} is written as

$$GS2 = \left(\sum_i a_{1i} Y_{i\cdot}, \quad \sum_i a_{2i} Y_{i\cdot} \right) \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}^{-1} \begin{pmatrix} \sum_i a_{1i} Y_{i\cdot} \\ \sum_i a_{2i} Y_{i\cdot} \end{pmatrix}$$

where

$$v_{sl} = \sum_i a_{si} a_{li} \sum_j (Y_{ij} - n_{ij}\bar{Y})^2$$

We call the test based on this statistics the GS2 test. We propose it for testing $H_0 : \pi_1 = \cdots = \pi_k$ against nonlinear trend. Note that no specific distribution is assumed for the generalized score test. However, in below we examine the behaviour of the GS2 test under beta-binomial distribution, assuming that $\{Y_{ij}\}_{j=1,\cdots,m_i}$ are independent and Y_{ij} follows,

$$P(Y_{ij} = y) = \begin{pmatrix} n_{ij} \\ y \end{pmatrix} \frac{\Gamma(y + (\pi_i/\varphi))\Gamma(n_{ij} - y + (1 - \pi_i)/\varphi)\Gamma(1/\varphi)}{\Gamma(\pi_i/\varphi)\Gamma((1 - \pi_i)/\varphi)\Gamma((1/\varphi) + n_{ij})},$$

where φ is the dispersion parameter.

3. Numerical Evaluation

We consider linear or quadratic responses shown in Table 1. We call the response patterns of No 1,2,3 and 4 in Table 1, the uniform, convex, concave and increasing monotone response, respectively.

No.	π_1	π_2	π_3	π_4	π_5	Response Patterns
1	0.2	0.2	0.2	0.2	0.2	uniform
2	0.2	0.45	0.8	0.6	0.5	convex
3	0.8	0.6	0.5	0.5	0.6	concave
4	0.2	0.2	0.27	0.4	0.6	increasing monotone

Table 1. Trend Patterns of Population's Succes Probability

Performance of the S1 test (C-A test), S2 test, GS1 test (generalized C-A test) and GS2 test are examined in terms of the Type I error and also in terms of the powers for detecting the true patterns by simulation. We consider binomial distributions with response probabilities given in Table 1 and three beta-binomial distributions with the value of dispersion parameter φ 0.1, 0.5 and 1, respectively. The significance level is taken as 0.05. 10,000 data are generated from each distribution, and empirical Type I errors and the powes are computed. The number of dams (m) per group and the number of litters (n) per dam is given equal.

3.1. Type I Error

Type I errors are shown in Figure 1 and 2 when $n = 5$ and 10, respectively. When the underlying pdf is binomial, Figure 1(a) and Figure 2 (a) show that the Type I errors of all tests are quite close to the nominal level $\alpha = 0.05$, although the GS1 and GS2 tests can not control the Type I error when m is very small. When the underlying pdf is beta-binomial, Figure 1 (b),(c),(d) and Figure 2 (b),(c),(d) show that the Type I errors of the S1 and S2 tests are deviated substantially from the nominal level; that the deviation increases as the increase of the value of φ; and it also increases as the increase of n. For instance from Figure 2 (d), the Type I error of S2 test is as remarkable as 0.6. It means that the S1 and S2 tests too often reject the true H_0, and that these tests should not be used when the data exhibit extra-binomial variability. In contrast the GS1 and GS2 tests perform reasonably well by keeping the Type I errors close to 0.05.

3.2. Power of the Tests

Empirical powers of the tests are shown in Figure 3,4,5,6,7 and 8. The S1, S2, GS1 and GS2 tests are considered in binomial distribution, but the S1 and S2 tests are not considered in the case of beta-binomial distribution since these tests violate the nominal type I error substantially.

Figure 3, 5 and 7 show the powers of these tests when $n = 5$, and Figure 4, 6 and 8 show the powers of these tests when $n = 10$.

Figure 3 and 4 show the powers of the tests for convex type response. Figure 3(a) and Figure 4 (a) show the powers of the tests when the underlying pdf is binomial. The figures show that the power of the S2 test is superior to the others; that the powers of the GS1 and GS2 tests are poor when m is small, in particular, when $m = 1$; and that the powers of the GS1 and GS2 tests get closer to those of the S2 and S1 tests when m becomes large. Figure 3 (b),(c),(d) and Figure 4 (b),(c),(d) show the powers of the tests when the underlying pdf is beta-binomial. The dispersion parameters in (b), (c) and (d), respectively, are 0.1, 0.5 and 1. Figure 3 (b),(c),(d) and Figure 4 (b),(c),(d)

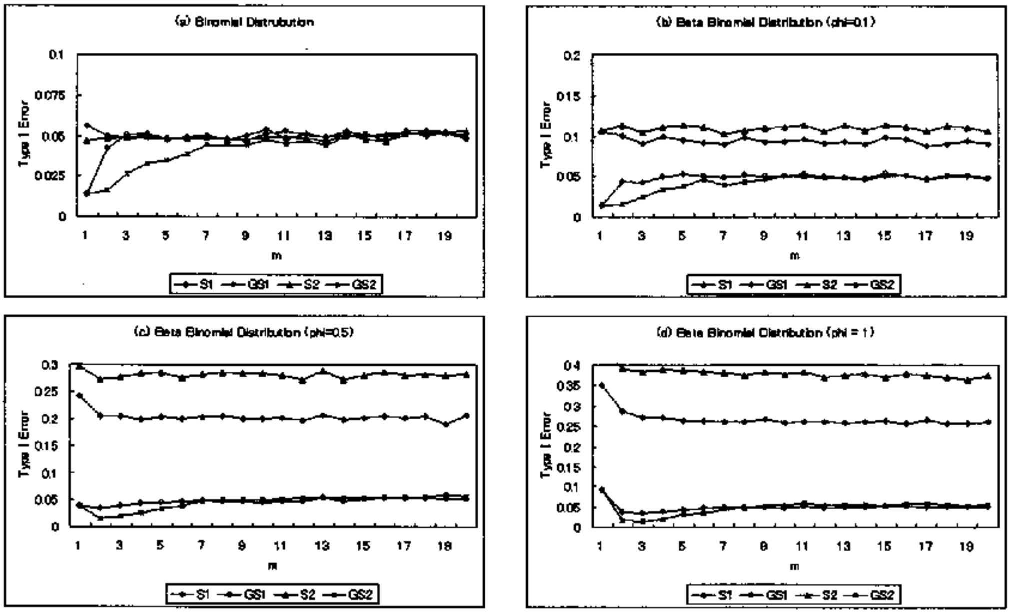

Figure 1: Empirical Type I Error ($n = 5$)

show that the powers of the GS1 test are a little better than the GS2 test when m is very small but otherwise the powers of the GS2 test are higher than those of the GS1 test; and the powers of the GS2 test decrease slightly as the increase of the value of the dispersion parameter.

For concave type response given in Table 1, we get similar results as the case of convex type, which is shown in Figure 5 and 6.

Figure 7 and 8 show powers of the tests for increasing monotone type response. Figure 7 (a) and 8 (a) show that, when underlying pdf is binomial, as is expected, the power of the S1 test (which is C-A test) is superior to the others. And when the underlying pdf is beta-binomial, Figure 7 (b),(c),(d) and Figure 8 (b),(c),(d) show that the power of the GS2 test is superior to the GS1 test.

4. Summary

The S1 and S2 tests for data from binomial pdf, and the GS1 and GS2 tets for data with overdispersion are systematically derived in the framework of score test and generalized score test. The S1 test is shown to be equivalent to the C-A test (Cochran, 1954: Armitage, 1955) and the GS2 test shown to be equivalent to the generalized C-A test (Carr and Gorelick, 1995). Furthermore, it is shown by simulation that

1. The S2 test is superior to the other tests when underlying pdf is binomial and response is nonlinear (convex or concave)

2. The S1 test and S2 test inflate the Type I error substantially and should not be used in the presence of overdispersion

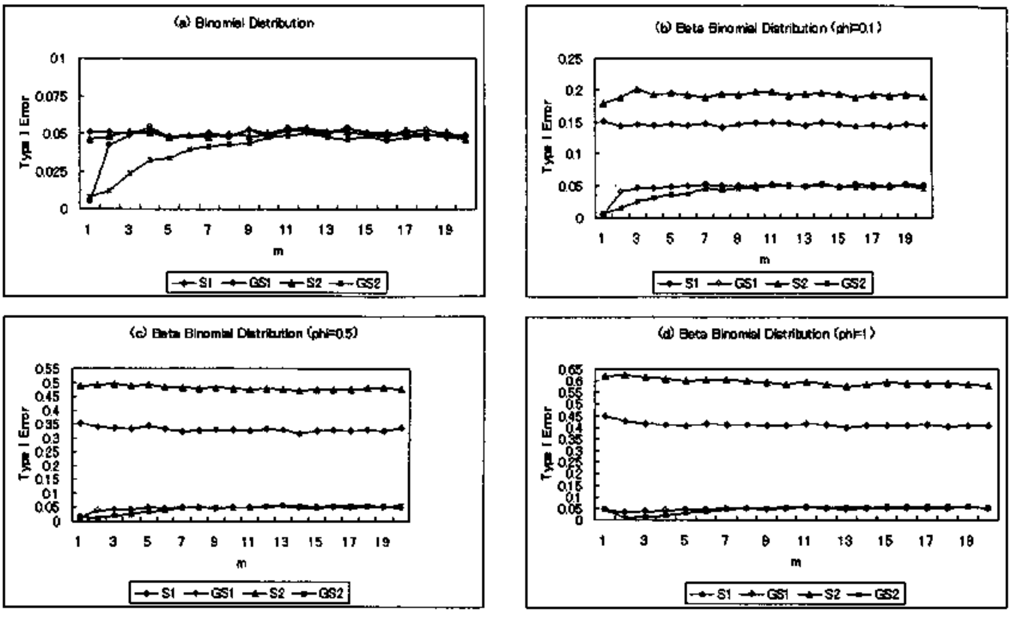

Figure 2: Empirical Type I Error ($n = 10$)

3. The GS2 test is superior to the GS1 test for detecting nonlinear response (convex or concave) in the presence of overdispersion

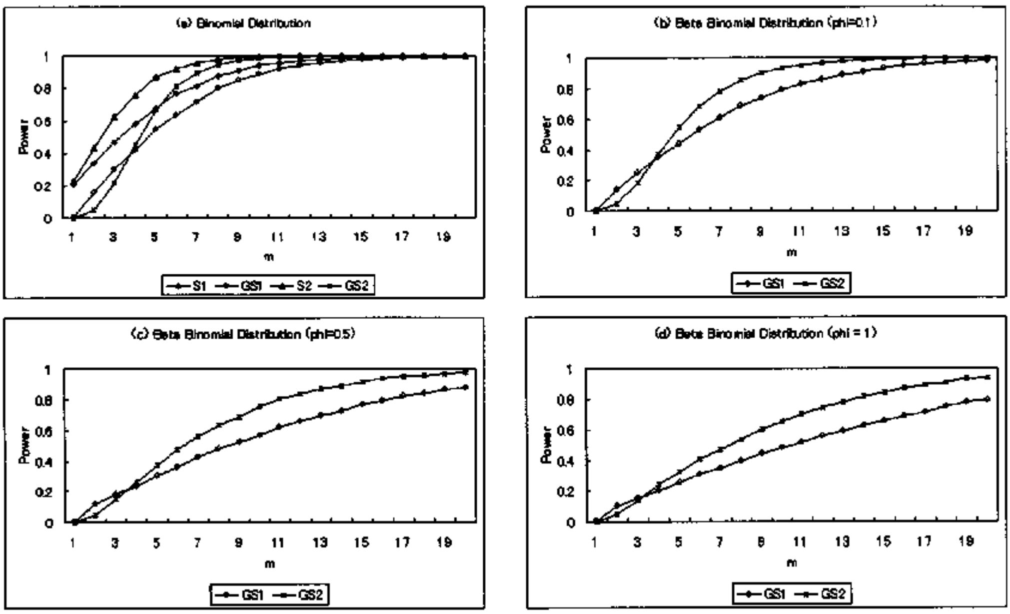

Figure 3: Empirical Power for convex type ($n = 5$)

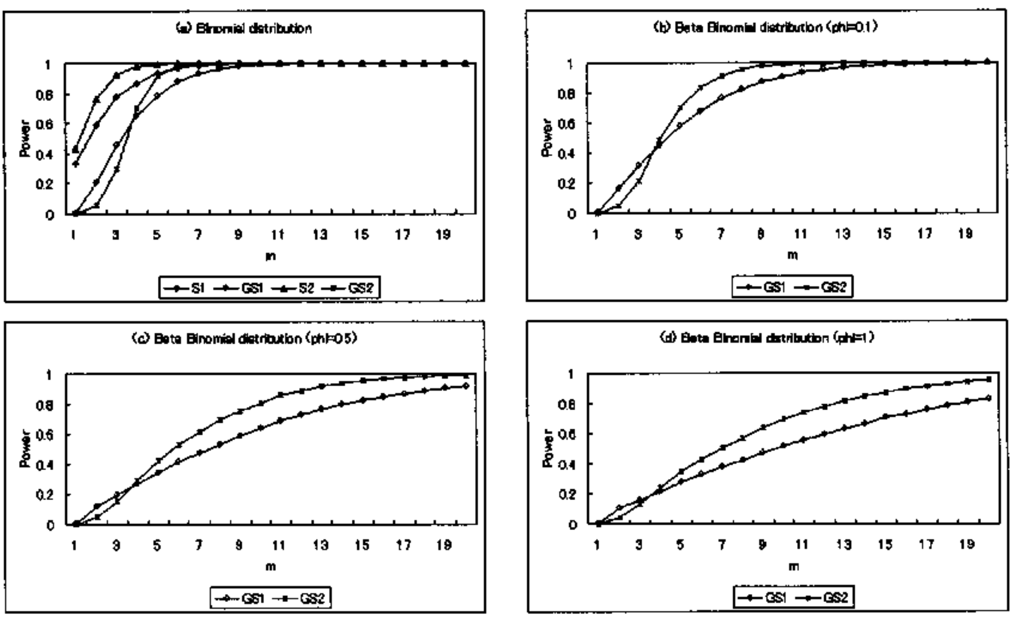

Figure 4: Empirical Power for convex type ($n = 10$)

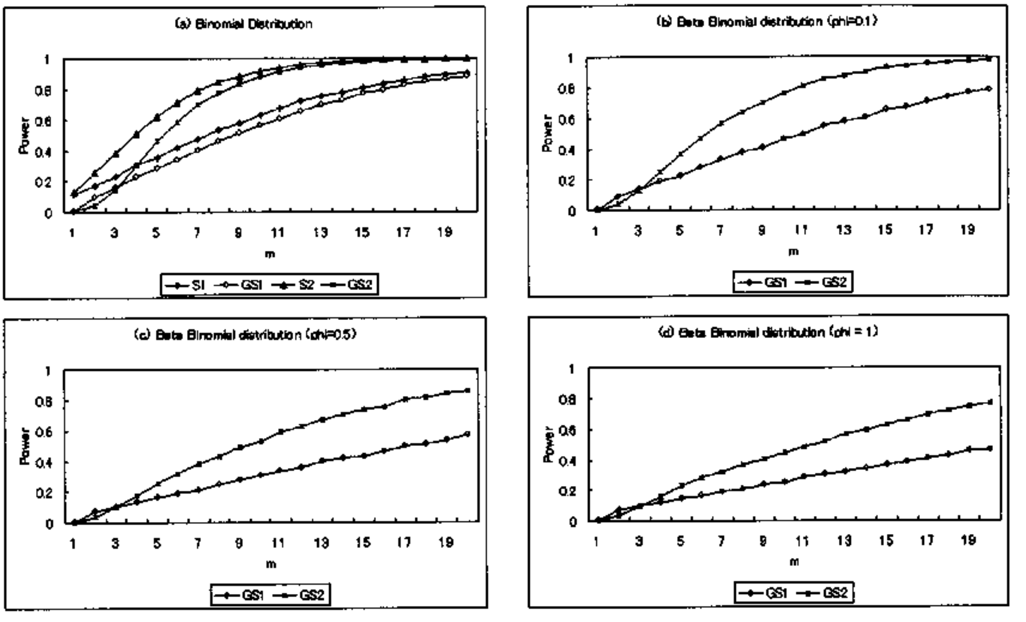

Figure 5: Empirical Power for concave type ($n = 5$)

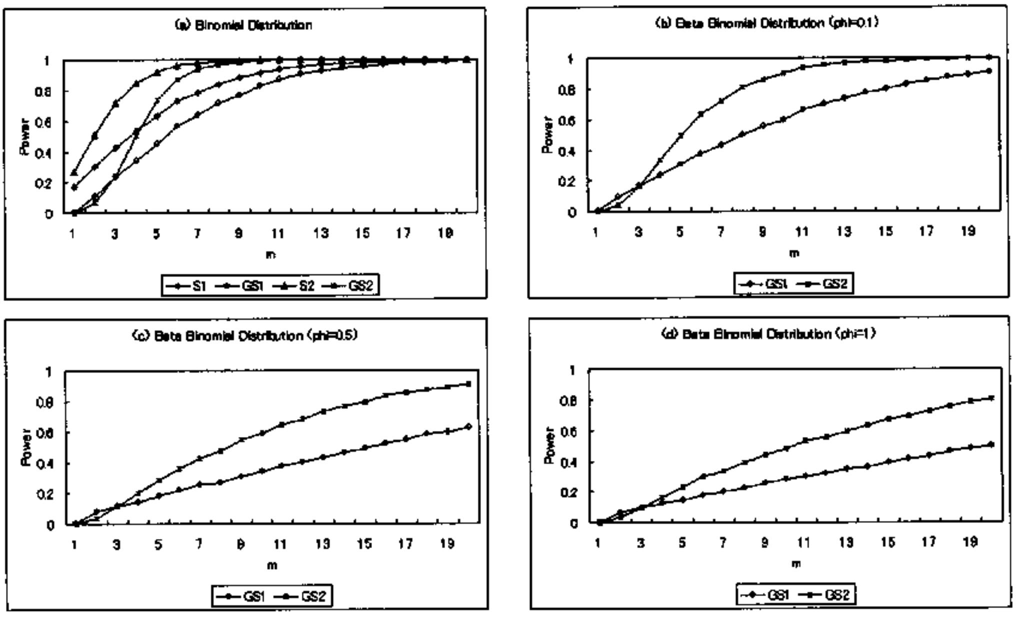

Figure 6: Empirical Power for concave type ($n = 10$)

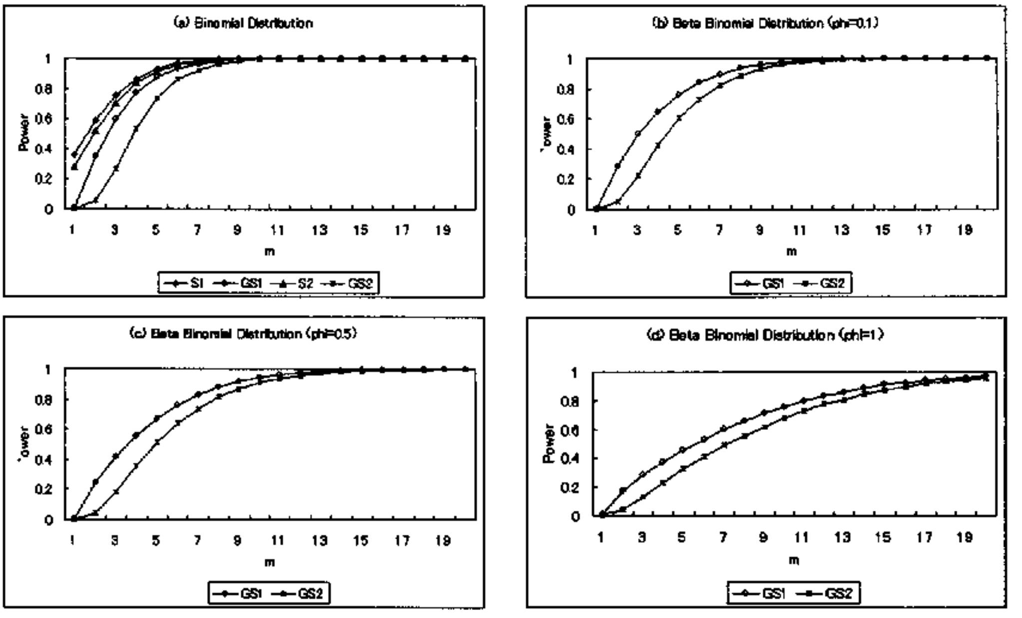

Figure 7: Empirical Power for increasing monotone type ($n = 5$)

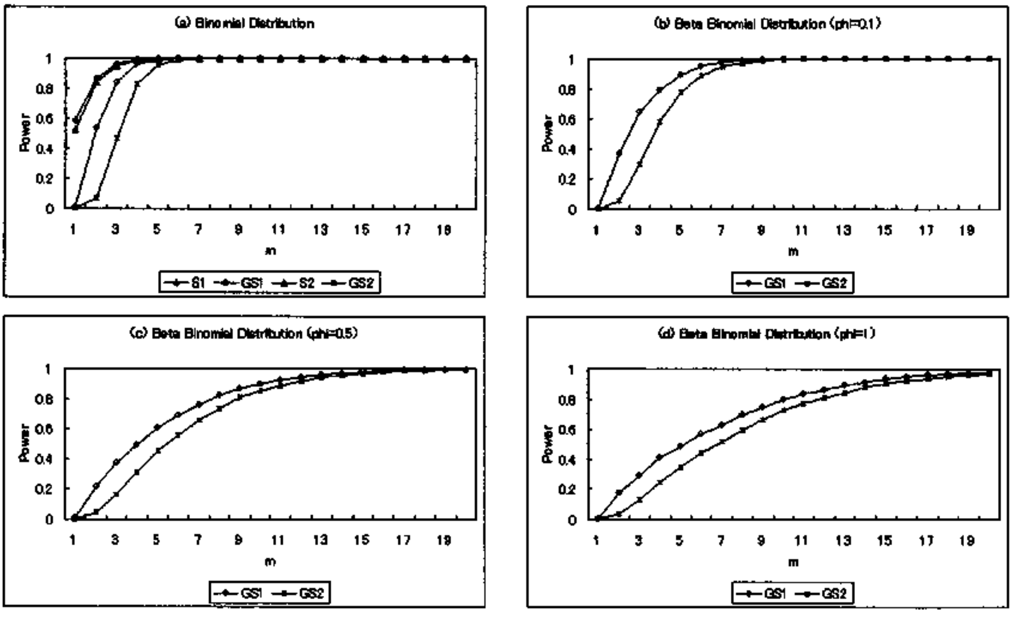

Figure 8: Empirical Power for increasing monotone type ($n = 10$)

References

Armitage, P. (1955). Test for linear trend in proportions and frequencies *Biometrics*, **11**, 375-386.

Boos, D.D (1992). On generalized score test *The American Statistician*, **46**, 327-333.

Carr, G.J and Gorelick, N.J (1995). Statistical design and analysis of mutation studies in transgenic mice *Environmental and Molecular Mutagenesis*, **25**, 245-255.

Cochran, W.G. (1992). Some methods for strengthning the common χ^2 tests *Biometrics*, **10**, 417-451.

Jayasekara, L.,Yanagawa, T. and Tsujitani, M. (1994). A Location-dispersion test for $2 \times k$ tables *Bulletin of Informatics and Cybernetics*, **26**,125-139.

Jayasekara, L., Nishiyama, H. and Yanagawa, T. (1999). Versatile tests for non-linear data in $2 \times k$ tables *Bulletin of Informatics and Cybernetics*, **31**, No.2, 176-190.

Piegorsch, W.W, and Bailer, A.J (1997). *Statistics for Enviromental Biology and Toxicology*, London:Chapman & Hall.

Rao, C.R. (1947). Large sample tests for statistical hypotheses concerning several parameters with applications to problems of estimation *Proceedings of the Cambridge Philosophical Society*, **44**, 50-57.

Appendix

PROPOSITION 4.1. *T_{GS_r} is the statistic of the generalized score test for testing H_0^* ag ainst H_1^**

PROOF. Let

$$\begin{aligned} S &= S(\boldsymbol{\beta}) = \left(\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_t}\right)_{(r+1)\times 1} \\ I_Y &= I_Y(\boldsymbol{\beta}) = \left(-\frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \beta_t \partial \beta_u}\right)_{(r+1)\times(r+1)} \\ D_Y &= D_Y(\boldsymbol{\beta}) = \left(\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_t}\frac{\partial \ell(\boldsymbol{\beta})}{\partial \beta_u}\right)_{(r+1)\times(r+1)} \end{aligned}$$

Reperent $S' = (S'_{(1)}, S'_{(2)})$, where $S_{(1)}$ is 1×1 and $S_{(2)}$ is $r \times r$. The matrices above are partitioned accordingly, e.g., $I_{Y(11)}$ is 1×1, $I_{Y(12)}$ is $1 \times r$ and so on. For testing $H_0^* : \boldsymbol{\beta}_{(2)} = \mathbf{0}$ vs $H_1^* : \boldsymbol{\beta}_{(2)} \neq \mathbf{0}$, Boos (1992) proposed generalized score test as:

$$T_{GS} = \tilde{S}'_{(2)} \tilde{\mathbf{V}}(\tilde{S}_{(2)}) \tilde{S}_{(2)} \qquad (A1)$$

where $\tilde{}$ denote those matrices evaluated at $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}$, and $\tilde{\boldsymbol{\beta}}$ is restricted mle of $\boldsymbol{\beta}$ under H_0^*, and

$$\begin{aligned} \tilde{\mathbf{V}}(\tilde{S}_{(2)}) &= \tilde{D}_{Y(22)} - \tilde{I}_{Y(21)}\tilde{I}_{Y(11)}^{-1}\tilde{D}'_{Y(21)} - \tilde{D}_{Y(21)}\tilde{I}_{Y(11)}^{-1}\tilde{I}'_{Y(21)} \\ &\quad + \tilde{I}_{Y(21)}\tilde{I}_{Y(11)}^{-1}\tilde{D}_{Y(11)}\tilde{I}_{Y(11)}^{-1}\tilde{I}'_{Y(21)} \end{aligned}$$

Since $\mathbf{a}_s = (a_{s1}, a_{s2}, \cdots, a_{sk})'$, $s = 1, 2, \cdots, r$ are orthonormal, it follows that $\tilde{I}_{Y(21)} = \tilde{I}'_{Y(12)} = \mathbf{0}$. Thus we have $\tilde{\mathbf{V}}(\tilde{S}_{(2)}) = \tilde{D}_{Y(22)}$. Furthermore, a straightforward computation shows that

$$\tilde{S}_{(2)} = \left(\sum_i a_{1i} Y_{i\cdot}, \cdots, \sum_i a_{ri} Y_{i\cdot} \right)'$$

Thus submitting those equation to (A1) we have the desired result.

Reveived December 18, 2000
Revised March 8, 2001