浜松医科大学紀要　一般教育　第25号（2011）

シンプソンのパラドックスに関連した2種類の χ^2 統計量の比較（1）

野田　明男
（総合人間科学講座・数学）

A Comparison of Two Kinds of the Chi-square Statistics Associated with Simpson's Paradox (1)

Akio NODA
(Integrated Human Sciences · Mathematics)

Abstract: Let us consider two strata of 2×2 contingency tables consisting of four entries a_h, b_h, c_h, d_h ($h = 1,2$), for which we have the marginal frequencies $M_{1,h} = a_h + b_h$, $M_{0,h} = c_h + d_h$, $N_{1,h} = a_h + c_h$, $N_{0,h} = b_h + d_h$, and the sample size $T_h = M_{1,h} + M_{0,h} = N_{1,h} + N_{0,h}$. The homogeneity of the strata would lead us to combine them into a single 2×2 contingency table that consists of the entries $a_1 + a_2$, $b_1 + b_2$, $c_1 + c_2$ and $d_1 + d_2$. We are then in a position to compute two kinds of chi-square statistics. Taking up the Mantel-Haenszel χ^2 test (one degree of freedom) on the data in the two strata, we first obtain χ_1^2 expressed in the form (1-1) in §1. Another χ^2 test comes from the single combined table for which the usual procedure yields χ_2^2 of the form (1-2) derived in §1. If a big difference between both statistics χ_1^2 and χ_2^2, is observed, such a situation is known as Simpson's paradox. In what follows, we use the notation χ_α^2 $(0 < \alpha < 1)$ indicating the percentage points of the χ^2 distribution with one degree of freedom. We now put

$$u = \left|M_{1,1} - M_{1,2}\right| / (M_{1,1} + M_{1,2})^{1/2} (M_{0,1} + M_{0,2})^{1/2}, \quad v = \left|N_{1,1} - N_{1,2}\right| / (N_{1,1} + N_{1,2})^{1/2} (N_{0,1} + N_{0,2})^{1/2}$$

which correspond to the χ^2 statistics for the data with the entries $M_{1,h}$, $M_{0,h}$ ($h = 1,2$), and with $N_{1,h}$, $N_{0,h}$ ($h = 1,2$), respectively.

The purpose of the present paper is to establish a comparison theorem for χ_1^2 and χ_2^2 tests under the following assumptions:

(i) The sample sizes are the same, $T_1 = T_2 = T$.

(ii) At least one of the equalities $M_{1,1} + M_{1,2} = T$, $N_{1,1} + N_{1,2} = T$ holds.

Indeed, we prove three theorems in §2. Theorem 1 tells us that a sufficient condition of the form Min $(u,v) \leq D_1$ guarantees that the inequality $\chi_1 \geq \chi_{\alpha_1}$ follows from $\chi_2 \geq \chi_{\alpha_2}$ for each pair (α_1,α_2) such that $0 < \alpha_2 < \alpha_1 \leq 5\%$; here, we find a good expression of D_1 as a function of (T,α_2,α_1). In contrast with Theorem 1, we have to suppose in Theorem 2, another sufficient condition of the form Max $(u,v) \leq D_2$, in order to claim that the inequality $\chi_1 \geq \chi_{\alpha_1}$ yields $\chi_2 \geq \chi_{\alpha_2}$ for each pair (α_1,α_2) such that $0 < \alpha_1 < \alpha_2 \leq 5\%$, where D_2 depends upon (T,α_1,α_2). Finally, Theorem 3 describes two different situations in which Simpson's paradox arises.

While the author was making a study of the books [2], [3] and others, he got familiar with the statistical analysis used in medicine and related fields, and succeeded in proving the theorems mentioned above, for which he is grateful to Professor Hirotsu.

Key words: 2×2 contingency table, Chi-square test, Simpson's paradox.

§1.　層の併合に付随する問題とシンプソンのパラドックス

2つの層 $(h=1,2)$ の2×2分割表（x は要因，y は目当ての変数を示す）を考える。これは，$H\times I\times J$ $(H=I=J=2)$ の3元分割表とみることもできる（[2]）。

層 h	$y=1$	$y=0$	例数
$x=1$	a_h	b_h	$M_{1,h}$
$x=0$	c_h	d_h	$M_{0,h}$
計	$N_{1,h}$	$N_{0,h}$	T_h

これら2つの層のデータは均質である，と想定して併合し，各度数の合計をとれば，1つの2×2分割表に縮約される。

併合	$y=1$	$y=0$	例数
$x=1$	a_1+a_2	b_1+b_2	$M_{1,1}+M_{1,2}$
$x=0$	c_1+c_2	d_1+d_2	$M_{0,1}+M_{0,2}$
計	$N_{1,1}+N_{1,2}$	$N_{0,1}+N_{0,2}$	T_1+T_2

層ごとの分割表に関する有意差検定において，層別解析の χ^2 統計量（χ_1^2 とかく）と，それらを併合した分割表から生じる通常の χ^2 統計量（χ_2^2 とかく）とを比較検討し，2つの数値が大きくくい違う状況（シンプソンのパラドックスといわれる）を数学的に解明すること，これがこの論文の研究テーマに他ならない。

初めに，一般的な2×2分割表の表記に沿って，Mantel-Haenszelによる有意差検定の考え方を説明する（[1],[2]参照）。

一般	$y=1$	$y=0$	例数
$x=1$	$a\,(p_1)$	$b\,(1-p_1)$	M_1
$x=0$	$c\,(p_0)$	$d\,(1-p_0)$	M_0
計	N_1	N_0	T

周辺度数を固定すると，自由度は $\upsilon=1$ であるから，1つの度数 a の変動に着目すれば十分である。帰無仮説 $H_0: p_1=p_0$（$=p$ とおく）が成り立つ状況においては，$x=1$ と $x=0$ の2群を一緒にして，計欄にある N_1 個の $y=1$ と N_0 個の $y=0$，合わせて T 個の y の値全部を1種の母集団と考える。そこから，例数欄にある M_1 個のサンプルを非復元抽出して，$y=1$ の個数 a が求まる。共通の p の推定量は $\hat{p}=N_1/T$ となるので，確率変数 a は平均 $M_1\hat{p}$，分散 $M_1M_0\hat{p}\,(1-\hat{p})/(T-1)$ をもつ超幾何分布に従う（[3]）。T が十分大きいとき，a の標準化 $z=\dfrac{a-M_1N_1/T}{\sqrt{M_1M_0N_1N_0/T^2(T-1)}}$ の確率分布は，標準正規分布 $N(0,1)$ で近似される。さらに，整数値から連続量へ移行する際には，1/2の補正を行うのが適切である。こうして，一般的な2×2分割表における有意差検定では，z^2 に1/2の補正を施した次の統計量が用いられる。

(1-0)　$\chi^2 = \left\{\left|a - M_1N_1/T\right| - 1/2\right\}^2 T^2(T-1)/M_1M_0N_1N_0$

次に，層別解析で用いられる χ^2 統計量に移る。平均からの偏差 $a_h - M_{1,h}N_{1,h}/T_h$ が，層 h によらずに，同じ向きにそろうという均質性を期待すれば，互いに独立な2変量の (a_1, a_2) から，1変量の和 $a_1 + a_2$ に縮約して調べるのが有効となる。よく知られた公式 $V\left(\sum a_h\right) = \sum V(a_h)$ を適用すると，

(1-1)　$$\chi_1^2 = \frac{\left\{\left|\sum(a_h - M_{1,h}N_{1,h}/T_h)\right| - 1/2\right\}^2}{\sum\left\{M_{1,h}M_{0,h}N_{1,h}N_{0,h}/T_h^2(T_h - 1)\right\}}$$

に到達する。($\sum$は $h = 1,2$ にわたる和を示す)。

最後に，2つの分割表を併合して，各度数の合計をとった場合を扱う。このとき得られる2×2分割表に関する χ^2 統計量は，(1-0)と同じ形で記述される。

(1-2)　$$\chi_2^2 = \frac{\left\{\left|(a_1 + a_2) - (M_{1,1} + M_{1,2})(N_{1,1} + N_{1,2})/(T_1 + T_2)\right| - 1/2\right\}^2 \left(T_1 + T_2\right)^2 \left(T_1 + T_2 - 1\right)}{\left(M_{1,1} + M_{1,2}\right)\left(M_{0,1} + M_{0,2}\right)\left(N_{1,1} + N_{1,2}\right)\left(N_{0,1} + N_{0,2}\right)}$$

われわれの研究目的は，こうして生じる2種類の統計量 χ_1^2 と χ_2^2 を比較することである。特に[2] p.160に記されている広津先生の主張を精密化したい。この本に例示されているシンプソンのパラドックスは $a_1 = 120, b_1 = 30, c_1 = 40, d_1 = 10$ および $a_2 = 10, b_2 = 40, c_2 = 30, d_2 = 120$ のケースであり，$\chi_1^2 = 0,\ \chi_2^2 = 34.723$, という風に，2つの数値は大きくくい違う。そして，2因子交互作用の $H \times I, H \times J$ 分割表に関する χ^2 統計量は，それぞれ 97.765, 141.256 の値をもち，両者とも一様性の帰無仮説 H_0 を棄却する結果となる。両者のうち，少なくとも1つがを H_0 を採択するならば，層別解析の χ_1^2 に基づく有意差検定は，簡便な χ_2^2 で代用することができる。これが広津先生の教えるところである。

次節では，層のサンプル･サイズが同一 $(T_1 = T_2 = T)$ の場合に限定して考察する。周辺度数の T に対する割合をそれぞれ，$M_{1,h}/T = r_h, N_{1,h}/T = s_h\ (h = 1,2)$ と記し，

$$\bar{r} = \left(r_1 + r_2\right)/2,\ \bar{s} = \left(s_1 + s_2\right)/2,\ \rho = 2\sqrt{\bar{r}(1 - \bar{r})},\ \sigma = 2\sqrt{\bar{s}(1 - \bar{s})}$$

を導入する。われわれが設定する十分条件は，次の3つである。

(1)　$\mathrm{Min}\left(\left|r_1 - r_2\right|/\rho, \left|s_1 - s_2\right|/\sigma\right) \le D_1$

(2)　$\mathrm{Max}\left(\left|r_1 - r_2\right|/\rho, \left|s_1 - s_2\right|/\sigma\right) \le D_2$

(3)　$\mathrm{Min}\left(\left|r_1 - r_2\right|/\rho, \left|s_1 - s_2\right|/\sigma\right) \ge D_3$

これらの条件下で，2種類の χ^2 検定を比較して得られる研究結果は，それぞれ定理1, 2, 3に結実する。現論文では，$\mathrm{Max}\ (\rho, \sigma) = 1$ (つまり，$\bar{r}$ と $\bar{s}$ のうち少なくとも1つは1/2に等しい)が成り立つという仮定をおいて，3つの定理を証明する。§2の終わりに，上記の限界値 D_1, D_2, D_3 の数値例を付記する。

こうして，広津先生の教えは定理1として精密化され，定理2はその逆命題を扱う。そしてシンプソンのパラドックスは，われわれの定理3によって具体化されることになる。一方，$\mathrm{Max}\ (\rho, \sigma) < 1$ (および $T_1 \ne T_2$)の場合には，取り組まねばならぬ数式が複雑で，証明にはかなりの時間が必要と思われる。後日，同じタイトルの(2)として発表することを期したい。

§2.　3つの定理と証明

層のサンプル･サイズは同一で，$T_1 = T_2 = T$ は十分大きいと仮定する。各度数と T との比を考えて，$a_h / T = r_h s_h + t_h\ (h = 1,2)$ とおくと，t_h は平均からの偏差を表す。$\bar{t} = (t_1 + t_2)/2$ とし，$\delta = r_1 - \bar{r}, \varepsilon = s_1 - \bar{s}$ を導入すれば，δ, ε は $H \times I, H \times J$ 分割表における平均からの偏差にそれぞれ対応している。層 h の分割表は，T 倍を除けば次の形をとる。

層 h	$y = 1$	$y = 0$	例数
$x = 1$	$r_h s_h + t_h$	$r_h(1 - s_h) - t_h$	r_h
$x = 0$	$(1 - r_h)\, s_h - t_h$	$(1 - r_h)(1 - s_h) + t_h$	$1 - r_h$
計	s_h	$1 - s_h$	1

まず，4つの度数はすべて非負であることから，

$$-\mathrm{Min}\left(\bar{r}\bar{s}, (1-\bar{r})(1-\bar{s})\right) \le \bar{t} + \delta\varepsilon \le \mathrm{Min}\left(\bar{r}(1-\bar{s}), (1-\bar{r})\bar{s}\right)$$

を得る。次式に出て来る複号の土は，$h = 1$ のとき＋，$h = 2$ のときーに対応する。

$$\begin{cases} r_h(1-r_h) = \left((\rho/2)^2 - \delta^2\right) \pm (1-2\bar{r})\delta \\ s_h(1-s_h) = \left((\sigma/2)^2 - \varepsilon^2\right) \pm (1-2\bar{s})\varepsilon \end{cases}$$

これらの積和を計算して，次式に至る。

$$\text{(2-0)} \qquad \sum r_h(1-r_h)s_h(1-s_h) = 2\left\{\left((\rho/2)^2 - \delta^2\right)\left((\sigma/2)^2 - \varepsilon^2\right) + (1-2\bar{r})(1-2\bar{s})\delta\varepsilon\right\}$$

$\bar{r}$ か $\bar{s}$ のうち，少なくとも1つは1/2に等しいと仮定したので，(2-0)式の第2項が消えて，χ_1^2 の式(1-1)は単純な形に書き直される。

$$\text{(2-1)} \qquad \chi_1^2 = \frac{\left\{4|\bar{t}| - T^{-1}\right\}^2 (2T-2)}{\left(\rho^2 - (2\delta)^2\right)\left(\sigma^2 - (2\varepsilon)^2\right)}$$

次に，2つの層を併合して，度数を合計したときは，$\sum(r_h s_h + t_h) = 2(\bar{r}\bar{s} + \bar{t} + \delta\varepsilon)$ となるので，T 倍を除いた分割表は次のようにかける。

併合	$y = 1$	$y = 0$	例数
$x = 1$	$2(\bar{r}\bar{s} + \bar{t} + \delta\varepsilon)$	$2\left(\bar{r}(1-\bar{s}) - \bar{t} - \delta\varepsilon\right)$	$2\bar{r}$
$x = 0$	$2\left((1-\bar{r})\,\bar{s} - \bar{t} - \delta\varepsilon\right)$	$2\left((1-\bar{r})(1-\bar{s}) + \bar{t} + \delta\varepsilon\right)$	$2(1-\bar{r})$
計	$2\bar{s}$	$2(1-\bar{s})$	1

こうして，(1-2)式の χ_2^2 を書き直して，次式を得るのは，容易な仕事である。

$$\text{(2-2)} \qquad \chi_2^2 = \frac{\left\{4|\bar{t} + \delta\varepsilon| - T^{-1}\right\}(2T-1)}{\rho^2\sigma^2}$$

さて，$2|\delta|/\rho = u, 2|\varepsilon|/\sigma = v$ と変数変換すると，座標平面上の点 (u,v) は長方形

$$0 \le u < 2\mathrm{Min}(\bar{r}, 1-\bar{r})/\rho = \rho/(1+\sqrt{1-\rho^2}) = u_0,\ 0 \le v < 2\mathrm{Min}(\bar{s}, 1-\bar{s})/\sigma = \sigma\ \left(1+\sqrt{1-\sigma^2}\right) = v_0$$

の内部を動く。ここで，Max $(u_0, v_0) = 1$ である。自由度 $\upsilon = 1$ の χ^2 分布のパーセント点 χ^2_α は，$N(0,1)$ の上側確率 $Q(\bullet)$ を用いて，$\chi_\alpha = Q^{-1}(\alpha/2)$ と表される。$Q_1 = \chi_{\alpha_1} / \sqrt{2T-2}$，$Q_2 = \chi_{\alpha_2} / \sqrt{2T-1}$ とおくと，2つの χ^2 検定の棄却域はそれぞれ，$\bar{t}$ が以下の(2-3), (2-4)を満たすという条件で定まる。すなわち，有意水準 α_1 で，$\chi_1 \geq \chi_{\alpha_1}$ となって，層別解析の帰無仮説 H_0 を棄却できるのは，

(2-3)　$4|\bar{t}| \geq Q_1\rho\sigma\sqrt{(1-u^2)(1-v^2)} + T^{-1}$

が成り立つとき。また，有意水準 α_2 で，$\chi_2 \geq \chi_{\alpha_2}$ となって，併合した分割表の帰無仮説 H_0 が棄却される条件は次の通り。

(2-4)　$4|\bar{t} + \delta\varepsilon| \geq Q_2\rho\sigma + T^{-1}$

われわれの結果を述べる用意が，これで整ったことになる。

定理1.　$Q_1 < Q_2 < 1$ が成り立つような $0 < \alpha_2 < \alpha_1 \leq 5\%$ を選び，$D_1 = D_1(T, \alpha_2, \alpha_1) = \sqrt{\left(Q_2^2 - Q_1^2\right)/\left(1 - Q_1^2\right)}$ と定める。このとき，Min $(u, v) \leq D_1$ ならば，$\chi_2 \geq \chi_{\alpha_2}$ から $\chi_1 \geq \chi_{\alpha_1}$ が従う。

定理2.　$0 < \alpha_1 < \alpha_2 \leq 5\%$ に対し，$D_2 = D_2(T, \alpha_1, \alpha_2) = \sqrt{\left(Q_1 - Q_2\right)/\left(Q_1 + 1\right)}$ と定める。このとき，Max $(u, v) \leq D_2$ ならば，$\chi_1 \geq \chi_{\alpha_1}$ から $\chi_2 \geq \chi_{\alpha_2}$ が従う。

定理3.

(i) $\bar{t}$ と $\varepsilon\delta$ は同符号とする。$Q_1 < Q_2 < 1$ が成り立つような $0 < \alpha_2 < 5\% < \alpha_1 < 1$ を選び，$D_3 = D_3(T, \alpha_2, \alpha_1) = \mathrm{Max}\left(Q_2, \sqrt{\left(Q_2 - Q_1\right)/\left(1 - Q_1\right)}\right)$ と定める。$D_3 <$ Min (u_0, v_0) のとき，Min $(u, v) \geq D_3$ ならば，$\chi_2 \leq \chi_{\alpha_2}$ から $\chi_1 \leq \chi_{\alpha_1}$ が従う。

(ii) $\bar{t}$ と $\varepsilon\delta$ は異符号で，$|\bar{t}| > |\varepsilon\delta|$ が成り立つとする。$0 < \alpha_1 < 5\% < \alpha_2 < 1$ に対し，$D_2 = \sqrt{\left(Q_1 - Q_2\right)/\left(Q_1 + 1\right)}$ と定める。$D_2 <$ Min (u_0, v_0) のとき，Min $(u, v) \geq D_2$ ならば，$\chi_1 \leq \chi_{\alpha_1}$ から $\chi_2 \leq \chi_{\alpha_2}$ が従う。

定理3は，シンプソンのパラドックスの具体化である。すなわち，(i)は $\chi_2 = \chi_{\alpha_2}$（$\alpha_2 < 5\%$ を小さくとる）の値は大きいけれども，$\chi_1 \leq \chi_{\alpha_1}$（$\alpha_1 > 5\%$ は大きくとる）の値は小さいケース。(ii)は反対に，$\chi_1 = \chi_{\alpha_1}$（$\alpha_1 < 5\%$ を小さくとる）の値は大きいけれども，$\chi_2 \leq \chi_{\alpha_2}$（$\alpha_2 > 5\%$ は大きくとる）の値は小さいケースを与えている。u, v がともに限界値 D_3, D_2 以上という条件下で，それぞれ成り立つ現象である。

証明　証明の容易な定理2から始めよう。(2-3)を使うと，$u, v \leq D_2$ ならば，

$$4|\bar{t} + \delta\varepsilon| \geq 4\left(|\bar{t}| - |\delta\varepsilon|\right) \geq Q_1\rho\sigma\sqrt{(1-u^2)(1-v^2)} - \rho\sigma uv + T^{-1}$$

$$\geq \rho\sigma\left\{Q_1(1 - D_2^2) - D_2^2\right\} + T^{-1} = \rho\sigma Q_2 + T^{-1}$$

次に定理3(ii)に移る。(2-3)の逆向き不等式を使うと，$\bar{t}$ と $\varepsilon\delta$ が異符号なので，$u, v \geq D_2$ ならば，$4|\bar{t} + \delta\varepsilon| = 4\left(|\bar{t}| - |\delta\varepsilon|\right) \leq \rho\sigma\left\{Q_1\sqrt{(1-u^2)(1-v^2)} - uv\right\} + T^{-1} \leq \rho\sigma Q_2 + T^{-1}$ となり，$\chi_2 \leq \chi_{\alpha_2}$ が示される。

定理3(i)では，(2-4)の逆向き不等式を使う。$\bar{t}$ と $\varepsilon\delta$ が同符号なので，
$4|\bar{t}| = 4\left(|\bar{t}+\delta\varepsilon| - |\delta\varepsilon|\right) \le \rho\sigma\left(Q_2 - uv\right) + T^{-1}$ を得る。$D_3 \le u < u_0, D_3 \le v < v_0$ のとき，
$Q_2 \le uv + Q_1\sqrt{(1-u^2)(1-v^2)}$ を示すのがわれわれの問題である。まず双曲線 $uv = k$ 上に制限して，
$k + Q_1\sqrt{1+k^2-(u^2+v^2)} = f(u,v)$ の最小値を求める。

（イ） $D_3^2 \le k \le D_3$ の範囲のとき：u^2+v^2 の最大値は，$\{u,v\} = \{D_3, k/D_3\}$ のとき達成されるので，
$f(u,v) \ge k + Q_1\left\{1+k^2-D_3^2-k^2/D_3^2\right\}^{1/2} = D_3 t + Q_1\sqrt{1-D_3^2}\sqrt{1-t^2}$ と評価できる。ここで，$t = k/D_3$ は区間 $D_3 \le t \le 1$ を動く。上記 t の関数は，区間の端点において最小となることがわかり，
$f(u,v) \ge \mathrm{Min}\left(D_3, D_3^2 + Q_1(1-D_3^2)\right) = Q_2$ が示される。

（ロ） $D_3 < k < u_0 v_0$ の範囲のとき：u^2+v^2 の上限は $\{u,v\}$ が $\{1,k\}$ に近づくとき達成されるので，
$f(u,v) > k > D_3 \ge Q_2$ を得る。以上で定理3(i)の証明が終わる。

最後に定理1を証明する。(2-4)を使えば，$4|\bar{t}| \ge 4\left(|\bar{t}+\delta\varepsilon| - |\delta\varepsilon|\right) \ge \rho\sigma\left(Q_2 - uv\right) + T^{-1}$ を得る。(u,v) の動く範囲は，長方形 $\left[0,u_0\right)\times\left[0,v_0\right)$ から小長方形 $\left(D_1,u_0\right)\times\left(D_1,v_0\right)$ を取り除いた領域 Γ である。（ただし，$u_0 < D_1$ または $v_0 < D_1$ の場合には，Γ は長方形となり，以下の議論を少し修正して証明することができる。）

われわれの示すべきことは，$(u,v) \in \Gamma$ に対し，$\dfrac{Q_2 - uv}{\sqrt{\left(1-u^2\right)\left(1-v^2\right)}} \ge Q_1$ が成り立つこと。まず，双曲線 $uv = k$ 上に制限して，$\dfrac{Q_2 - k}{\sqrt{1+k^2-\left(u^2+v^2\right)}} = g(u,v)$ の最小値を求める。

（イ） $0 \le k \le D_1^2$ の範囲のとき：u^2+v^2 の最小値は $u = v = \sqrt{k}$ のとき達成されるので，

$g(u,v) \ge \dfrac{Q_2 - k}{1-k} \ge \dfrac{Q_2 - D_1^2}{1-D_1^2} = \dfrac{Q_2 + Q_1^2}{1+Q_2} > Q_1$ が示される。

（ロ） $D_1^2 < k < D_1$ の範囲のとき：u^2+v^2 の最小値は，$\{u,v\} = \{D_1, k/D_1\}$ のとき達成されるので，
$t = k/D_1$ とおくと，$g(u,v) \ge \dfrac{Q_2 - D_1 t}{\sqrt{1-D_1^2}\sqrt{1-t^2}}$ を得る。この t の関数は，$D_1 < Q_2$ に注意すると，
$t = D_1/Q_2$ のとき最小値 $\dfrac{Q_2^2 - D_1^2}{\sqrt{1-D_1^2}\sqrt{Q_2^2-D_1^2}} = \sqrt{\dfrac{Q_2^2-D_1^2}{1-D_1^2}} = Q_1$ をとることがわかり，定理1の証明が完了する。

証明の中で追求した極値問題については，定理1のみ，(u,v) の領域 Γ の境界線上の2点 $(D_1, D_1/Q_2)$, $(D_1/Q_2, D_1)$ において最小値をとり，他のケースはすべて，長方形領域の頂点で極値をとる。これに関連して，次の事実を付記する。定理1の十分条件を定理2のタイプ $(\mathrm{Max}\,(u,v) \le D_4)$ に変更すると，定理1の主張を成り立たせる限界値は，$D_4 = \sqrt{(Q_2 - Q_1)/(1-Q_1)}$ となり，D_1 よりもずっと大きくとれる。

註　$H \times I$ と $H \times J$，2つの分割表に関する χ^2 統計量は，

$$\chi_3^2 = (2|\delta| - T^{-1})^2(2T-1)/\rho^2,\ \chi_4^2 = (2|\varepsilon| - T^{-1})^2(2T-1)/\sigma^2$$

とかける。有意水準5%で帰無仮説 H_0 を採択できる状況を考える。ある $\alpha_0 > 5\%$ に対して，上記 χ^2 統計量が $\chi^2_{\alpha_0}$ 以下となるのは，$Q_0 = \chi_{\alpha_0} / \sqrt{2T-1}$ として，$u \le Q_0 + (\rho T)^{-1}, v \le Q_0 + (\sigma T)^{-1}$ がそれぞれ成り立つときである。今 Max $(\rho,\sigma) = 1$ なので，$Q_0 + T^{-1}$ と定理1の D_1（数値例参照），2つの数値を比べれば，われわれの結果は§1で述べた広津先生の主張（[2]）の精密化になっていると言えよう。

最後に，D_1, D_2, D_3 の数値例をいくつか付記して，この節を閉じることにする。
簡略に Min (α_1, α_2) = min, Max (α_1, α_2) = max とかき，表中の数値は D_i $(i = 1, 2, 3)$ の値を表す。与えられた T の値と組 (min, max) に対して，$D_1 < D_2$ と $D_2 < D_3$ が常に成り立つ。

（1）$T = 100$ のとき：

(D_1, D_2)	max = 1%	max = 2.5%	max = 5%
min = 0.5%	0.0794, 0.1187	0.1208, 0.1840	0.1435, 0.2247
min = 1%		0.0904, 0.1429	0.1192, 0.1931
min = 2.5%			0.0772, 0.1325

(D_3, D_2)	max = 10%	max = 20%	max = 30%
min = 0.5%	0.3049, 0.2629	0.3446, 0.3010	0.3678, 0.3241
min = 1%	0.2728, 0.2370	0.3173, 0.2792	0.3429, 0.3044
min = 2.5%	0.2181, 0.1919	0.2731, 0.2430	0.3033, 0.2721

（2）$T = 200$ のとき：

(D_1, D_2)	max = 1%	max = 2.5%	max = 5%
min = 0.5%	0.0559, 0.1015	0.0849, 0.1580	0.1010, 0.1932
min = 1%		0.0637, 0.1224	0.0839, 0.1657
min = 2.5%			0.0545, 0.1131

(D_3, D_2)	max = 10%	max = 20%	max = 30%
min = 0.5%	0.2516, 0.2262	0.2855, 0.2590	0.3057, 0.2790
min = 1%	0.2251, 0.2035	0.2630, 0.2399	0.2850, 0.2615
min = 2.5%	0.1801, 0.1642	0.2264, 0.2081	0.2521, 0.2331

謝辞

丁寧に原稿を読んで，有益な助言を与えて下さった査読者に感謝申し上げます。また，資料の整理とともに，原稿の清書をして下さった鴨藤江利子さんに，深く感謝致します。

参考文献

[1] 丹後俊郎：メタ・アナリシス入門．朝倉書店，2002.
[2] 広津千尋：医学・薬学データの統計解析．東京大学出版会，2004.
[3] 広津千尋：離散データ解析．教育出版，1982.
[4] 山内二郎(編)：統計数値表JSA-1972．日本規格協会，1972.