統計数理研究所彙報
第32巻 第1号（1984）

多重 2×2 分割表における共通標準差について

岡山理科大学 山 本 英 二
統計数理研究所 柳 本 武 美

（1984年4月 受付）

1. 序

多重 2×2 分割表は Case-Control study や Clinical trial などに現れるデータ形式であるがその統計的解析の一つのテーマは共通した関連性の推定・検定問題である．このモデルは離散型比例ハザードモデルで共変量が2値変量となる場合にも現れる．2×2 表における関連性の指標としては，見込比（odds ratio）と，その対数変換値，対数見込比（log odds ratio）が良く知られているが，これら以外にも Cochran (1954) によって示唆され，Fleiss (1970) によって紹介された標準差 (standardized difference) がある．彼らは数値例から対数見込比と標準差はほぼ同等の関連性の指標と考えた．Yates (1955) は共通標準差（Common standardized difference）の推定量を与え，そのゼロ仮説の χ^2 検定統計量を示唆した．Fleiss (1973) はその一様性仮説を検定するための χ^2 検定統計量を与えた．ただ，これらの χ^2 検定統計量は共通標準差としてより，むしろ共通対数見込比の χ^2 検定統計量として用いられて来た．ちなみに，Zelen (1971) が与えた対数見込比の一様性を検定するための χ^2 検定統計量は Fleiss のものと同じである．しかし，関連性の指標としては標準差は不適当であり Fleiss の χ^2 検定統計量は実用上の利点は無いとの指摘が Halperin et al. (1977) と Mantel et al. (1977) によってなされ，Fleiss も改訂版（1981）では書き直した．

ところが，より詳細に調べて見れば Fleiss が χ^2 検定統計量を導くために用いた彼の一様性の検定のための χ^2 検定法の一般理論には仮説の誤りがある．又，Yates (1955) 以後の議論では，共通標準差の母数とその推定量の間には対応関係がないため，母数での議論と推定量での議論に混乱が起っている．そこで本論文ではこれらの議論を整理し，その使用上の注意を与える．

2. 共通標準差とは

2つの2項確率変量 X_k と Y_k を K 組考える．第 k 組における X_k, Y_k の出現確率を p_k, q_k とする．X_k, Y_k からの標本の大きさを各々 n_k, m_k, 出現数を x_k, y_k とする．$s_k=x_k+y_k$, $t_k=n_k+m_k$ とする．2つの平均出現確率

$$\pi_k=\frac{p_k+q_k}{2} \tag{2.1}$$

$$\tau_k=\frac{n_k p_k+m_k q_k}{t_k} \tag{2.2}$$

を導入する．これらをまとめて表1に示す．

関連性の指標として，対数見込比（見込比）と，確率差（probability difference)，及び2つ

表1. 第k 2×2表の表記法（$k=1, 2, \cdots, K$）

確率表

	出現	非出現	計
X_k	p_k	$1-p_k$	1
Y_k	q_k	$1-q_k$	1
平均 π	π_k	$1-\pi_k$	1
平均 τ	τ_k	$1-\tau_k$	1

データ表

	出現	非出現	計
X_k	x_k	n_k-x_k	n_k
Y_k	y_k	m_k-y_k	m_k
計	s_k	t_k-s_k	t_k

の標準差を与える.

$$\beta_k=\log\frac{p_k(1-q_k)}{(1-p_k)q_k}\quad\left(\lambda_k=\frac{p_k(1-q_k)}{(1-p_k)q_k}\right) \tag{2.3}$$

$$d_k=p_k-q_k \tag{2.4}$$

$$\delta_k=\frac{d_k}{\pi_k(1-\pi_k)} \tag{2.5}$$

$$\delta_k^*=\frac{d_k}{\tau_k(1-\tau_k)} \tag{2.6}$$

確率差 d_k は Cochran (1954) が与えた. d_k が β_k と異なる指標であることは, d_k が β_k に関して

$$d_k=\frac{e^{\alpha_k}}{(1+e^{\alpha_k})^2}\beta_k+O(\beta_k^2) \tag{2.7}$$

と展開されることから明らかである. ここで便宜上母数 (p_k, q_k) を次の関係式で変換してある.

$$p_k=\frac{e^{\alpha_k+\beta_k}}{1+e^{\alpha_k+\beta_k}},\quad q_k=\frac{e^{\alpha_k}}{1+e^{\alpha_k}}, \tag{2.8}$$

と展開されることから明らかである. 元来, β_k が乗法タイプのモデルを考えているのに対し, d_k は加法タイプのモデルを考えている. 彼は, 又標準化した d_k, δ_k が対数見込比 β_k とほぼ同等の値を示すことを数値例で示した. 後に Fleiss (1970) は δ_k を標準差として紹介している. 彼の数値例でも β_k との同等性は支持されている. このことから δ_k が β_k と同等な関連性の指標として使えると考えられて来た. δ_k^* は, 後の議論の便宜上, 我々が導入した母数で, 2つの標本の大きさの比に依存している.

次に $\{\delta_k\}$ について, 3つの統計的仮説を与える.

$$H_0:\ \delta_1=\cdots=\delta_K=0$$
$$H_c(\delta):\ \delta_1=\cdots=\delta_K=\delta$$
$$H_s:\ \{\delta_k\}\text{ は任意}$$

ここで, δ を母数における共通標準差と呼ぶ. H_0 はゼロ仮説, $H_c(\delta)$ は一様性の仮説である. $H_c(\delta)$ の下での共通標準差 δ の推定問題と, 次の3つの検定問題が考えられる. H_0 を $H_c(\delta)$ ($\mathscr{H}(H_0, H_c(\delta))$) に対して検定するゼロ共通標準差の検定, $\mathscr{H}(H_c(\delta), H_s)$ に対して検定する標準差の一様性の検定, そして $\mathscr{H}(H_0, H_s)$ に対して検定する全標準差ゼロの検定である. 統計的には前2者の検定が重要なことが多い. 順に association, homogeneity, total の検定と呼ばれている (Fleiss, 1973). 他の関連性の指標に対する統計的仮説も同様に δ_k を $\beta_k(\lambda_k)$, d_k, δ_k^* に置き換えて得られる. ただし $\beta_k=0(\lambda_k=1)$, $d_k=0$, $\delta_k=0$, $\delta_k^*=0$ は全て同値であることに注意すると, H_0, H_s は全ての指標について共通であることが分かる. 一様性の仮説だけが異なり

$$H_c(\beta):\ \beta_1=\cdots=\beta_K=\beta$$
$$H_c(d):\ d_1=\cdots=d_K=d$$
$$H_c(\delta^*):\ \delta_1^*=\cdots=\delta_K^*=\delta^*$$

となる．

3. これまでの議論の整理

Cochran (1954) は $H_c(d)$ の下で，d_k の MLE $x_k/n_k-y_k/m_k$, を，重み n_km_k/t_k を使って平均を取り，共通確率差 d の推定量

(3.1) $$\tilde{d}=\sum\frac{m_kx_k-n_ky_k}{t_k}\Big/\sum\frac{n_km_k}{t_k}$$

を求め，H_0 の下での $\tilde{d}$ の分散の推定量を求めて $\mathscr{H}(H_0, H_c(d))$ に対する χ^2 検定統計量として

(3.2) $$C\chi_0^2=\sum\left(\frac{m_kx_k-n_ky_k}{t_k}\right)^2\Big/\sum\left(\frac{n_km_ks_k(t_k-s_k)}{t_k{}^3}\right)$$

を与えた．H_0 の下で $C\chi_0^2$ は漸近的に自由度 1 の χ^2 分布に従う．

Yates (1955) は δ_k の推定量として

(3.3) $$\tilde{\delta}_k=\frac{\dfrac{x_k}{n_k}-\dfrac{y_k}{m_k}}{\dfrac{s_k}{t_k}\left(1-\dfrac{s_k}{t_k}\right)}$$

を与えた．重みとしては H_0 の下での $\mathrm{Var}(\tilde{\delta}_k\,|\,H_0)$ の推定量の逆数

(3.4) $$w_k=\frac{n_km_k}{t_k}\cdot\frac{s_k}{t_k}\left(1-\frac{s_k}{t_k}\right)$$

を考えて，$H_c(\delta)$ の下での共通標準差 δ の推定量として

(3.5) $$\tilde{\delta}=\frac{\sum w_k\tilde{\delta}_k}{\sum w_k}=\sum\left(\frac{m_kx_k-n_ky_k}{t_k}\right)\Big/\sum\left(\frac{n_km_ks_k(t_k-s_k)}{t_k^3}\right)$$

を与えた．又，$\{w_k\}$ を用いた $\mathscr{H}(H_0, H_c(\delta))$ に対する χ^2 検定統計量を示唆したが，それは $C\chi_0^2$ に一致している．これは d と δ で H_0 が共通ということと，H_0 の下での分散の推定量の逆数を重みと考える方法を両者共採用しているためである．ところが β と δ の同一視から，$C\chi_0^2$ は，むしろ $\mathscr{H}(H_0, H_c(\beta))$ に対する χ^2 検定統計量として使われて来ている．

Fleiss (1973) は一様性の検定に対する彼の一般理論を Yates の議論の上に適用し，$\mathscr{H}(H_c(\delta), H_s)$ に対する χ^2 検定統計量

(3.6) $$F\chi_c^2=\sum w_k(\tilde{\delta}_k-\tilde{\delta})^2=\sum w_k\tilde{\delta}^2-C\chi_0^2$$

を提案した．この χ^2 検定統計量は，Zelen (1971) が $\mathscr{H}(H_c(\beta), H_s)$ に対する χ^2 検定統計量として提案したものと同じである．

しかし，共通対数見込比の一様性に対する χ^2 検定統計量としては，$F\chi_c^2$ は適当ではないとの批判が後になされた．Halperin et al. (1977) は $K=2$ のときに $F\chi_c^2$ が $H_c(\beta)$ の下で従う漸近分布を求め，これが必ずしも χ^2 分布でないことを示した．そして，$K=2$ として，推定値が $\hat{\beta}_1-\hat{\beta}_2$ ($\hat{\beta}_k$ は β_k の MLE, $k=1,\cdots,K$) にもかかわらず，$\tilde{\delta}_1$ と $\tilde{\delta}_2$ が大きく異なってしまう例と，その逆の例を上げた．これらのことから $\tilde{\delta}_k$ が 2 つの標本の大きさの比に大きく左右される推定量で，$\hat{\beta}_k$ とはまったく異なった関連性の推定量であることを指摘した．$F\chi_c^2$ が $\{\hat{\beta}_k\}$ の一様性でなく $\{\tilde{\delta}_k\}$ の一様性の検定のためのものであると指摘している．このことは Mantel et al. (1977) によっても同様の主旨で批判された．

母数においては δ_k と β_k は近いと言われながら，その推定量 $\tilde{\delta}_k$ と $\hat{\beta}_k$ は別物であると指摘

されたり，元来 δ に基づく χ^2 検定量が β の非零性の検定量として使われて来たり，δ についての議論は多くの混乱が見られる．この混乱は，母数と推定量が対応していないことと，どの仮説の下での分布論を行っているのかが明確に述べられていないためである．以降の節で，これらを整理して誤解されている事実を指摘する．

4.　母数と推定量の対応の誤り

$\tilde{\delta}_k$ は δ_k の推定量としては良くない．実際，$\tilde{\delta}_k$ は δ_k^* の MLE である．従って $F\chi_c^2$ は $\mathscr{H}(H_C(\delta^*), H_S)$ に対する χ^2 検定統計量と見なすのが自然である．Mantel et al. (1977), Halperin et al. (1977) の批判は当然といえる．また $C\chi_0^2$ についても，尤度方程式の線形近似の立場から，$\mathscr{H}(H_0, H_C(\beta))$ に対してより，むしろ $\mathscr{H}(H_0, H_C(\delta^*))$ に対する χ^2 検定統計量と考えるべきとの指摘を柳本・山本は行っている．

さらに δ_k は β_k とほぼ同等であると言われているが，より詳細に調べてみれば，それは $|\beta_k|$ が1より小さいときのみに成立し，一方が2以上にもなれば δ_k と β_k はかなり離れてしまう．実際，δ_k は β_k に比べてかなり小さめな値をとる．

δ_k と β_k の関係は，

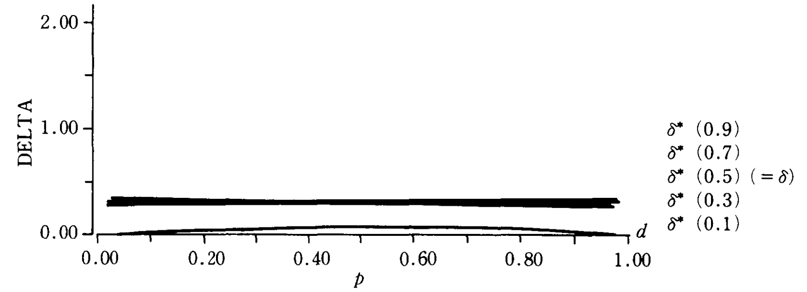

図1　β が固定されたときの δ, δ^*, p の間の関係，$\delta^*(x)$ は $x=n/t$ のときの δ^* を示す．$\beta=0.3$ の場合．

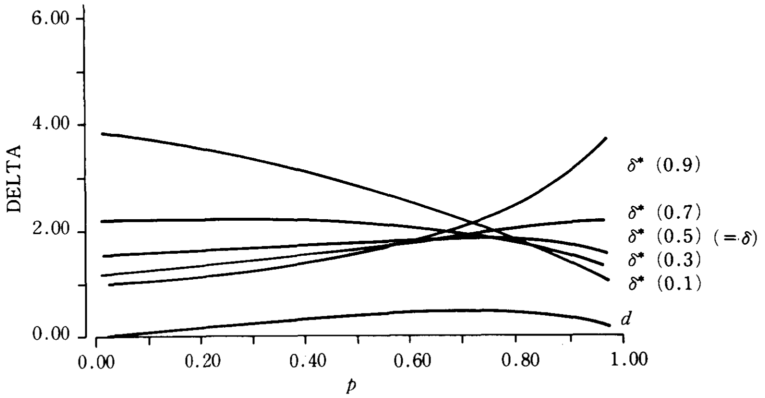

図1　(続き) $\beta=2.0$ 場合．

$$\delta_k=\beta_k-\frac{3e^{2\alpha}+3e^{\alpha}+1}{12(1+e^{\alpha})^2}\beta_k^3+O(\beta_k^4) \tag{4.1}$$

となっている．両者は 2 次のオーダーまで等しいので $|\beta_k|<1$ のときは，かなり δ_k と β_k は近いと言えるが，その外では，やはり差は開いて行く．図 1 に β_k, d_k, δ_k, $\delta_k{}^*$ の 4 者の関係を示してある．横軸に p_k をとり縦軸に 4 者の値をとっている．$|\beta_k|$ が 1 以上では，δ_k とずれること，又 d_k と δ_k^* は独自のカーブを描くことが見て取れる．

5.　推定量の分布論の誤り

次に，$F\chi_c^2$ は $\mathscr{H}(H_c(\delta^*), H_s)$ の χ^2 検定統計量としても，分布論的に誤りがある．この誤りは Fleiss（1981）による一様性のための χ^2 検定統計量を導く一般理論の誤りに由来している．そこで，まず Fleiss の一般理論の誤りを指摘し次に具体的に $F\chi_c^2$ の分布が $H_c(\delta^*)$ の下で，漸近的に χ^2 分布に従わないことを示す．

5.1.　Fleiss の一般理論における分布論の誤り

Fleiss は K 標本における平均値の一様性の検定問題での分散分析の理論を 2 項分布の関連性の指標の一様性の問題に拡張して次の 3 つの統計量を導入する．先ず関連性の指標を表わす母数 θ_k の推定量が漸近的に正規分布 $N(\theta_k, \sigma_k^2)$ に従うと仮定する．更に σ_k^2 に確率収束する σ_k^2 の推定量，$\hat{\sigma}_k^2$ が得られたと仮定する．このとき重み ω_k を $1/\hat{\sigma}_k^2$ として 3 つの統計量

$$\begin{aligned}\chi_{assoc}^2&=(\sum\omega_k\hat{\theta}_k)^2/\sum\omega_k\\ \chi_{homog}^2&=\sum\omega_k(\hat{\theta}_k-\hat{\theta})^2\\ \chi_{total}^2&=\sum\omega_k\hat{\theta}_k^2,\end{aligned} \tag{5.1}$$

ただし $\hat{\theta}=\sum\omega_k\hat{\theta}_k/\sum\omega_k$, を導入する．これらは，各々検定問題 $\mathscr{H}(H_0,\ H_c(\theta))$, $\mathscr{H}(H_c(\theta), H_s)$, $\mathscr{H}(H_0, H_s)$ の漸近的に自由度 1, $K-1$, K をもつ χ^2 検定統計量として用いられる．ここで $\chi_{total}^2=\chi_{assoc}^2+\chi_{homog}^2$ が代数的に成立することに注意する．

以上の Fleiss の議論は，どの仮説の下での分布論であるか明確に述べていないが，暗黙の内に H_0 を仮定している．しかし，χ_{homog}^2 は $H_c(\theta)$: $\theta_1=\cdots=\theta_K=\theta$ の下での漸近分布を扱かわなくてはいけない．正規母集団では $\hat{\sigma}_k^2$ として標本分散や不偏分散があり，これらは H_0 のみならず，$H_c(\theta)$ の下でも σ_k^2 に確率収束する．しかし，2 項母集団の場合は，σ_k^2 は θ_k の関数となって，H_0 の下での分散 $\sigma_k^2(0)$ の推定量 $\hat{\sigma}_k^2$ は，$H_c(\theta)$ の下では，必ずしも $\sigma_k^2(\theta)$ に確率収束しない．そのため $H_c(\theta)$ の下では χ_{homog}^2 は必ずしも漸近的に χ^2 分布に従わない．以下ではこの違いを明らかにする．

今，$\hat{\theta}_k$ が漸近正規分布 $N(\theta_k, \sigma_k^2(\theta_k))$ に従うと仮定する．更に，

$$\hat{\sigma}_k^2\xrightarrow{p}\begin{cases}\sigma_k^2(0) & under\ H_0\\ \eta_k^2(\theta)(\neq\sigma_k^2(\theta)) & under\ H_c(\theta)\end{cases} \tag{5.2}$$

と仮定する．ここで，更に以降では，記号 $\hat{\theta}\xrightarrow{p}\theta$ は母数 θ とその推定量 $\hat{\theta}$ の比 $\hat{\theta}/\theta$ が 1 に確率収束することを表わす．記号

$$\hat{\boldsymbol{\theta}}=\begin{bmatrix}\hat{\theta}_1\\ \hat{\theta}_2\\ \vdots\\ \hat{\theta}_K\end{bmatrix},\quad \mathbf{1}=\begin{bmatrix}1\\ 1\\ \vdots\\ 1\end{bmatrix},\quad \boldsymbol{\Sigma}(\theta)=\begin{bmatrix}\sigma_1^2(\theta) & & 0\\ & \ddots & \\ 0 & & \sigma_K^2(\theta)\end{bmatrix},\quad \boldsymbol{H}(\theta)=\begin{bmatrix}\eta_1^2(\theta) & & 0\\ & \ddots & \\ 0 & & \eta_K(\theta)\end{bmatrix}$$

を導入する．H_0 の下では漸近的には $\omega_k=1/\sigma_k^2(0)$ として良いので

$$\chi_{homog}^2=(\hat{\boldsymbol{\theta}}-\bar{\theta}\mathbf{1})'\boldsymbol{\Sigma}^{-1}(0)(\hat{\boldsymbol{\theta}}-\bar{\theta}\mathbf{1}) \tag{5.3}$$

ここで $\bar{\theta}=(\mathbf{1}'\boldsymbol{\Sigma}^{-1}(0)\mathbf{1})^{-1}\mathbf{1}\boldsymbol{\Sigma}^{-1}(0)\hat{\boldsymbol{\theta}}$, と表わされる．$\chi^2_{homog}$ は H_0 の下で自由度 $K-1$ の χ^2 分布に従う．一方，$H_c(\theta)$ の下では漸近的に $\omega_k=1/\eta_k^2(\theta)$ として良いので，

$$\chi^2_{homog}=(\hat{\boldsymbol{\theta}}-\bar{\theta}\mathbf{1})'\boldsymbol{H}^{-1}(\theta)(\hat{\boldsymbol{\theta}}-\bar{\theta}\mathbf{1}) \tag{5.4}$$

ここで，$\bar{\theta}=(\mathbf{1}'\boldsymbol{H}^{-1}(\theta)\mathbf{1})^{-1}\mathbf{1}\boldsymbol{H}^{-1}(\theta)\hat{\boldsymbol{\theta}}$, と表わされる．そして $H_c(\theta)$ の下で $\hat{\boldsymbol{\theta}}\sim N(\theta\mathbf{1},\boldsymbol{\Sigma}(\theta))$ に注意すれば，

$$\boldsymbol{H}^{-\frac{1}{2}}(\theta)(\hat{\boldsymbol{\theta}}-\bar{\theta}\mathbf{1})\sim N(0\mathbf{1},\boldsymbol{D}) \tag{5.5}$$

ここで $\boldsymbol{D}=(\boldsymbol{I}-\boldsymbol{H}^{-\frac{1}{2}}(\theta)\mathbf{1}(\mathbf{1}'\boldsymbol{H}^{-1}(\theta)\mathbf{1})^{-1}\mathbf{1}'\boldsymbol{H}^{-\frac{1}{2}}(\theta))\boldsymbol{H}^{-\frac{1}{2}}(\theta)\boldsymbol{\Sigma}(\theta)\boldsymbol{H}^{-\frac{1}{2}}(\theta)\cdot$
$(\boldsymbol{I}-\boldsymbol{H}^{-\frac{1}{2}}(\theta)\mathbf{1}(\mathbf{1}'\boldsymbol{H}^{-1}(\theta)\mathbf{1})^{-1}\mathbf{1}'\boldsymbol{H}^{-\frac{1}{2}}(\theta))$,

である．この分散共分散行列はランク $K-1$ の半正値行列となっているので，$\boldsymbol{D}$ の正の固有値を $d_1^2,\cdots,d_{K-1}^2$ とすると，$H_c(\theta)$ の下で

$$(\hat{\boldsymbol{\theta}}-\bar{\theta}\mathbf{1})'\boldsymbol{H}^{-1}(\theta)(\hat{\boldsymbol{\theta}}-\bar{\theta}\mathbf{1})\sim d_1^2Z_1^2+\cdots+d_{K-1}^2Z_{K-1}^2 \tag{5.6}$$

となる．ここで，$Z_1,\cdots,Z_{K-1}$ は互いに独立で $N(0,1)$ に従う確率変数である．$\boldsymbol{D}$ は一般には射影子でないから χ^2_{homog} は $H_c(\theta)$ の下では必ずしも χ^2 分布に従わない．

例として $K=2$ のときを調べる．このときは固有値が $(\sigma_1^2(\theta)+\sigma_2^2(\theta))/(\eta_1^2(\theta)+\eta_2^2(\theta))$ となるので，$H_c(\theta)$ の下で

$$\chi^2_{homog}\sim\frac{\sigma_1^2(\theta)+\sigma_2^2(\theta)}{\eta_1^2(\theta)+\eta_2^2(\theta)}\cdot Z_1^2 \tag{5.7}$$

となる．即ち $(\eta_1^2(\theta)+\eta_2^2(\theta))/(\sigma_1^2(\theta)+\sigma_2^2(\theta))\cdot\chi^2_{homog}$ が，自由度1の χ^2 分布に従う．

5.2. $F\chi_c^2$ の分布論の誤り

具体的に $F\chi_c^2$ について調べる．$H_c(\delta^*)$ の下では $\tilde{\delta}_k$ の分散は

$$\mathrm{Var}(\tilde{\delta}_k\mid H_c(\delta^*))\fallingdotseq\left(\frac{p_k(1-p_k)}{n_k}+\frac{q_k(1-q_k)}{m_k}\right)\Big/\tau_k^2(1-\tau_k)^2 \tag{5.8}$$

となる．ここで≒は左辺と右辺の比が1に収束することを示す．一方 w_k については，$H_c(\delta^*)$ の下で，

$$1/w_k\xrightarrow{p}\frac{t_k}{n_km_k}\Big/\tau_k(1-\tau_k) \tag{5.9}$$

となる．これらの右辺は H_0 の下では等しいが，$H_c(\delta^*)$ の下では等しくならない．従って $F\chi_c^2$ は $H_c(\delta^*)$ の下では必ずしも χ^2 分布に従わない．$K=2$ のときには，$H_c(\delta^*)$ の下で

$$\frac{\dfrac{t_1}{n_1m_1}\Big/\tau_1(1-\tau_1)+\dfrac{t_2}{n_2m_2}\Big/\tau_2(1-\tau_2)}{\dfrac{\frac{p_1(1-p_1)}{n_1}+\frac{q_1(1-q_1)}{m_1}}{\tau_1^2(1-\tau_1)^2}+\dfrac{\frac{p_2(1-p_2)}{n_2}+\frac{q_2(1-q_2)}{m_2}}{\tau_2^2(1-\tau_2)^2}}\cdot F\chi_c^2\sim\chi_1^2 \tag{5.10}$$

となる．したがって，δ^* の一様性に対する χ^2 検定統計量は，分布論の観点からは，(5.8)の推定量として

$$\hat{\mathrm{V}}\mathrm{ar}(\tilde{\delta}_k\mid H_c(\delta^*))=\left(\frac{x_k(n_k-x_k)}{n_k^3}+\frac{y_k(m_k-y_k)}{m_k^3}\right)\Big/\frac{s_k^2(t_k-s_k)^2}{t_k^4} \tag{5.11}$$

を考えて，その逆数を重みとした方が良い．しかし δ^* は関連性の指標としては，標本の大きさの比に依存することから実際上，ほとんど意味がないので，ここでは，δ^* の推定検定問題にこれ以上立ち入らない．

6. 議　　論

関連性の指標，δ^* については実用上の利点はない．むしろ β について種々提案されている推定量・検定量の良さの判断の道具として使えることの方が重要である．β の推定量の中には，むしろ δ^* の推定量と考えた方が妥当であり，従って β の推定量としては良くないものがある等が，明らかにされている（柳本・山本）．

δ については，その解釈は困難で，β ほどの応用上の利点はない．そして β の代用としても，1 以上になれば使えない．また β については MLE は非線形方程式を解くことになるが，δ についても MLE については同じで，計算が楽であることもない．結局，共通標準差 δ，δ^* については理論上も実用上も，ほとんどそれ自身の利点は無いといえよう．

関連度としては，むしろ β とは全く異なる指標である共通確率差 d が重要であると思われる．β が乗法モデルを考えているのに対し，d は加法モデルを考えているからである．d の推定・検定方法については Cochran によるもの以外にも，訂正を加えた Fleiss の一般理論を用いれば，共通確率差 d の推定量

$$\bar{d}=\sum w_k\left(\frac{x_k}{n_k}-\frac{y_k}{m_k}\right)\Big/\sum w_k \tag{6.1}$$

ただし，$\omega_k=1\Big/\left(\frac{x_k(n_k-x_k)}{n_k^3}+\frac{y_k(m_k-y_k)}{m_k^3}\right)$，と，ゼロ共通確率差 $\mathscr{H}(H_0, H_c(d))$ に対する χ^2 検定統計量

$$\chi_0^2=\left(\sum \omega_k\left(\frac{x_k}{n_k}-\frac{y_k}{m_k}\right)^2\right)\Big/\sum \omega_k \tag{6.2}$$

及び，一様性 $\mathscr{H}(H_c(d), H_s)$ に対する χ^2 検定統計量

$$\chi_c^2=\sum \omega_k(\hat{d}_k-\bar{d})^2 \tag{6.3}$$

ここで，$\hat{d}_k=\frac{x_k}{n_k}-\frac{y_k}{m_k}$，が新たに提案できる．

もちろん MLE による推定も非線形方程式を解くことによって求められる．

7. 数　値　例

ここでは，Halperin et al. (1977), Mantel et al. (1977) が取り上げた例を取り上げる．結果を表 2 にまとめてある．

確率変数 X_1，Y_1，X_2，Y_2 をもつ 4 個の 2 項母集団から各々標本抽出されたと見なす人工的なデータで，2 組の 2×2 表は同じ対数見込比をもつ．$F\chi_c^2$ によれば，36.13 となり一様性なしと反対の結論になってしまうとされた例である．β の一様性でなく，δ^* の一様性を問題にしているためである．その上 $F\chi_c^2$ は尤度比検定量に比べても非常に大きい．誤った χ^2 近似が用いられているからである．訂正を加えた Fleiss の方法では 5.12 となる．MLE 9.38 との差は正規分布への近似精度の差のためである．5 節により，$F\chi_c^2$ に定数をかけて χ^2 分布に従う様にすると 0.14×36.13=5.10 となり，訂正したものにほぼ等しくなる．δ については β が 3 以上もあるため，β よりかなり小さな値を示している．確率差 d も参考に計算した．

表2. Halperin et al. (1977), Mantel et al. (1977)の例, $K=2$

第1組

	出現	非出現	計
X_1	200	30	230
Y_1	10	40	50
計	210	70	280

第2組

	出現	非出現	計
X_2	10	30	40
Y_2	10	800	810
	20	830	850

関連性の指標		推定値			一様性検定量
		第1組	第2組	共通	
乗法モデル	β (MLE)	3.28	3.28	3.28	0.00
	δ^* (Yates, Fleiss)	3.57	10.36	4.26	36.13***
	δ^* (MLE)	3.57	10.36	3.79	9.38**
	δ^* (訂正Fleiss法)	3.57	10.36	3.65	5.12*
	δ (訂正Fleiss法)	2.69	2.09	2.60	3.42
加法モデル	d (訂正Fleiss法)	0.69	0.24	0.48	22.02***

参考文献

Cochran, W.G. (1954). Some methods for strengthening the common χ^2 tests, *Biometrics,* **10**, 417-451.

Fleiss, J.L. (1970). On the asserted invariance of the odds ratio, *British Journal of Preventive and Social Medicine,* **24**, 45-46.

————(1973, 1981). *Statistical Methods for Rates and Proportions* (1st, 2nd ed.), John Wiley & Sons, New York. (第1版の訳 佐久間昭：計数データの統計学, 東大出版会, 1975)

Halperin, M. et al. (1977). Testing for interaction in an $I \times J \times K$ contingency table, *Biometrika,* **64**, 271-275.

Mantel, N. et al. (1977). Tests for homegeneity of effect in an epidemiologic investigation, *American Journal of Epidemiology,* **106**, 125-129.

Yanagimoto, T. and Yamamoto, E. Simple linear approximations to the likelihood equation for combining evidence in multiple 2×2 tables: A critique of conventional procedures, *Ann. Inst. Statist. Math.,* (to appear)

Yates, F. (1955). The use of transformations and maximum likelihood in the analysis of quantal experiments involving two treatments, *Biometrika,* **42**, 382-403.

Zelen, M. (1971). The analysis of several 2×2 contingency tables, *Biometrika,* **58**, 129-137.

28645

Proceedings of the Institute of Statistical Mathematics Vol. 32, No. 1 (1984) 33

The common standardized difference in the analysis of multiple 2×2 tables

Eiji Yamamoto and Takemi Yanagimoto

(Okayama University of Science and Institute of Statistical Mathematics)

Multiple contingency tables are obtained by combining evidence from independent samples or by stratifying a sample by a background factor. Several measures of association between two binomial variables have been introduced and explored. In the analysis of multiple 2×2 tables special attention is usually paid to a common measure of association.

The standardized difference has been typically defined in two different ways. One of the two unfortunately depends on the ratio of sample sizes. The other, due to Cochran (1954), is asserted to be close to the log-odds ratio, which is a quite popular and practical measure. The main aims of this paper are to provide insights into the above two measures and to study estimators of them and test statistics. We point out erroneous results in existing literature.

Unfortunately, this discrepancy between the standardized difference and the log-odds ratio may give rise to confusion. The discrepancy may result in poor linear approximations to the likelihood equation for the log-odds ratio. The test for homogeneity of log-odds ratios by Fleiss (1973) is reviewed and criticized in detail. The exact asymptotic distribution of the test statistic under the null hypothesis of a common standardized difference is presented.