統計数理（1998）
第46巻 第1号 153-177

特集「ヘルスサイエンスと統計科学」
［総合報告］

Mantel-Haenszel の方法による複数の 2×2 表の要約

統計数理研究所　佐藤　俊哉・高木　廣文
九州大学大学院*　柳川　　堯
統計数理研究所　柳本　武美

（受付 1997年10月13日；改訂 1997年12月18日）

要　　旨

交絡要因で層別した 2×2 表から関連の指標である共通オッズ比を推定する方法として，重み付き最小二乗法，Mantel-Haenszel 法，最尤法，条件付き最尤法，Peto（一段階）法などが提案されている．本論文では，2 種類の漸近モデルを導入し，他の方法との比較を行いながら，Mantel-Haenszel の方法を中心に共通オッズ比の推定量，信頼区間の構成について，最近の発展のまとめを行う．層別した $2\times J$ 表への Mantel-Haenszel の方法の拡張，推定方程式としての Mantel-Haenszel の方法の位置づけについても議論する．

キーワード：共通オッズ比，推定関数，層別解析，Mantel-Haenszel の方法．

1. はじめに

1959年，歴史的に意義のある論文が Mantel と Haenszel によって発表された．その論文の中で Mantel と Haenszel は，原因と結果との関係をゆがめる交絡要因を調整して，帰無仮説の検定を行う方法と曝露効果の指標であるオッズ比を推定する方法を与えた（Mantel and Haenszel (1959)）．この Mantel-Haenszel の方法は計算が簡便であるだけでなく，他の競合するオッズ比の推定量である重み付き最小二乗推定量や最尤推定量と比較してもバイアスや効率の点でいい性質を持ち，1994年の終わりには 4,000 件以上も引用され，科学論文に最も引用された 200 の論文の一つとなっている（Breslow (1996)）．

Mantel-Haenszel の方法では，交絡要因のレベルによって層別した複数の 2×2 表から，それらを統合して帰無仮説の検定や各層に共通なオッズ比の推定を実施する．もともとは疫学でよく用いられるケース・コントロール研究のために提案された手法であるが，複数の 2×2 表を統合する一般的な方法論であるため，最近ではメタアナリシスにも用いられている．ケース・コントロール研究で用いる統計手法としての Mantel-Haenszel の方法の位置づけは Breslow (1996) が，また共通オッズ比の推定に用いる統計手法としての位置づけは Emerson (1994) が，それぞれ総説をまとめている．

本論文では，まず 2 節で Mantel-Haenszel の方法とそれに競合する共通オッズ比の推定方法を簡単に紹介し，複数の 2×2 表に適用される 2 つの異なった漸近モデルを導入する．続いて 3 節では，共通オッズ比の信頼区間について Mantel-Haenszel の方法の発展を述べる．4 節では

* 数理学研究科：〒812-0053 福岡市東区箱崎 6-10-1.

これら共通オッズ比の近似的な信頼区間と，拡張（非心）超幾何分布にもとづく正確な方法との比較を行う．Mantel-Haenszel の方法は現在さまざまな方向に拡張されているが，5節では，複数の $2\times J$ 表への Mantel-Haenszel の方法の拡張をまとめる．またあまり知られていないことだが，Mantel-Haenszel の方法は経時観察データ解析のための一般化推定方程式（GEE）の理論的展開のもとになっている．6節では，この推定関数としての Mantel-Haenszel の方法を論ずることにする．

2. 共通オッズ比の推定

表1に第 k 層の 2×2 表を示す．(x_k, y_k) は生起確率 (p_{1k}, p_{0k})，サイズ (n_k, m_k) の独立な二項系列とする $(k=1, \ldots, K)$．また，リスク要因への曝露状況を X（あり$=1$，なし$=0$），疾病発生状況を Y（あり$=1$，なし$=0$）とする．特に疫学的ケース・コントロール研究では，x_k は疾病を発生したケース n_k 人中曝露を受けた人数であり，y_k は疾病を発生していないコントロール m_k 人中曝露を受けた人数である．すべての層に共通なオッズ比を，

$$\psi=\frac{p_{1k}(1-p_{0k})}{p_{0k}(1-p_{1k})}, \tag{2.1}$$

と定義する．オッズ比は疾病の発生がまれな場合，曝露によって疾病発生割合が何倍になったかという解釈が可能であり，ケース・コントロール研究は疾病がまれな場合に適した研究デザインであるため，オッズ比を関連の指標として用いることが正当化される（佐藤（1993））．共通オッズ比の推定量はいくつか提案されているが，ここでは Mantel-Haenszel 推定量を含めて代表的なものを5つ紹介する．

2.1 Mantel-Haenszel 推定量

Mantel and Haenszel（1959）は共通オッズ比の推定量として，

$$\hat{\psi}_{\mathrm{MH}}=\frac{\sum_k x_k(m_k-y_k)/N_k}{\sum_k y_k(n_k-x_k)/N_k}=\frac{\sum_k R_k}{\sum_k S_k}, \tag{2.2}$$

を提案した．ただし，$R_k=x_k(m_k-y_k)/N_k$，$S_k=y_k(n_k-x_k)/N_k$ である．各層のオッズ比の推定値を $\hat{\psi}_k=x_k(m_k-y_k)/[y_k(n_k-x_k)]$ とすると，すべての層にわたって $S_k\neq 0$ であれば，Mantel-Haenszel オッズ比は各層の推定オッズ比の重み付き平均，

$$\hat{\psi}_{\mathrm{MH}}=\frac{\sum_k S_k\hat{\psi}_k}{\sum_k S_k},$$

となっている．重み S_k は帰無仮説 $H_0: \psi=1$ のもとでの推定オッズ比の漸近分散の推定量の逆数である（Tarone (1981) ; Schlesselman (1982)）．Mantel-Haenszel 推定量の重みについては6節で詳しく議論する．

表1．第 k 層のデータレイアウト．

	X=1	X=0	合　計
Y=1	x_k	n_k-x_k	n_k
Y=0	y_k	m_k-y_k	m_k
合　計	t_k	N_k-t_k	N_k

2.2 重み付き最小二乗推定量

各層の推定オッズ比の漸近分散の逆数を重みとして平均をとるのであるが，オッズ比は正の値しかとらないので対数変換した後で，

$$\log \hat{\psi}_{\mathrm{WLS}}=\frac{\sum_k \log \hat{\psi}_k/w_k}{\sum_k w_k^{-1}}, \tag{2.3}$$

とすることを Woolf (1955) が提案した．このため，経験的ロジット推定量ともよばれる．実際の重み w_k^{-1} としては $\log \hat{\psi}_k$ の漸近分散の推定量，

$$w_k=\frac{1}{x_k}+\frac{1}{n_k-x_k}+\frac{1}{y_k}+\frac{1}{m_k-y_k},$$

の逆数を用いる．重み付き最小二乗推定量はゼロセルが存在すると計算できないので，Gart (1966, 1971) はすべてのセルに 1/2 を加えた Haldane-Anscombe のバイアス修正 (Haldane (1956); Anscombe (1956)) を行って $\log \hat{\psi}_{\mathrm{WLS}}$ を求めることを提案した．

2.3 最尤推定量

共通オッズ比を推定する際のスコア関数は，

$$\sum_k x_k=\sum_k E(x_k|\psi),$$

であるが，積二項分布モデルのもとでの無条件の最尤推定量 $\hat{\psi}_{\mathrm{U}}$ では，右辺は漸近期待値でよく，

$$\sum_k x_k=\sum_k E^A(x_k|\hat{\psi}_{\mathrm{U}}), \tag{2.4}$$

の解として求められる．漸近期待値は，$E_k^A=E^A(x_k|\psi)$ として，次の二次方程式，

$$E_k^A(m_k-t_k+E_k^A)=\psi(n_k-E_k^A)(t_k-E_k^A), \quad k=1, \ldots, K,$$

の適切な根である．(2.4)式は，表 1 で (n_k, m_k, t_k) を固定した拡張（非心）超幾何分布のもとで，すべての層について $N_k\to\infty$ の場合の近似として提案されたが (Gart (1970))，Breslow (1976) は積二項分布モデルにもとづく無条件のスコア関数であることを示した．Breslow の結果から，ロジスティックモデル，

$$\log\frac{\Pr(Y=1|X, k)}{1-\Pr(Y=1|X, k)}=\alpha_k+\beta X,$$

において β の最尤解を b とすると，$\hat{\psi}_{\mathrm{U}}=e^b$ が得られる．

積拡張超幾何分布モデルにもとづく条件付き最尤推定量 $\hat{\psi}_{\mathrm{C}}$ では，拡張超幾何分布の正確な期待値を用いて，

$$\sum_k x_k=\sum_k E(x_k|t_k, \hat{\psi}_{\mathrm{C}}), \tag{2.5}$$

を解く必要がある (Birch (1964))．拡張超幾何分布の正確な期待値は，

$$E(x_k|t_k, \psi)=\sum_u u\Pr(u|t_k, \psi),$$

ただし，$u\in[\max(0, t_k-m_k), \min(n_k, t_k)]$，$\Pr(x|t_k, \psi)$ は拡張超幾何分布の確率関数，

$$\Pr(x|t_k, \psi)=\frac{\binom{n_k}{x}\binom{m_k}{t_k-x}\psi^x}{\sum_u\binom{n_k}{u}\binom{m_k}{t_k-u}\psi^u},$$

である．

2.4 Peto 推定量

$\log\psi=\beta$ として β の条件付き最尤推定を考えよう．Fisher スコア法による反復推定を行うと，i 回目の推定値から $i+1$ 回目の推定値は，

$$\hat{\beta}^{i+1}=\hat{\beta}^{i}+\frac{\sum_k x_k-\sum_k E(x_k|t_k,\hat{\psi}^i)}{\sum_k \mathrm{var}(x_k|t_k,\hat{\psi}^i)},$$

と改良できる．Peto (一段階) 推定量 $\hat{\psi}_{\mathrm{P}}$ は，初期値 $\beta=0$ から Fisher スコア法を一回だけ反復した，

$$\log\hat{\psi}_{\mathrm{P}}=\hat{\beta}^{1}=\frac{\sum_k x_k-\sum_k E(x_k|t_k,\psi=1)}{\sum_k \mathrm{var}(x_k|t_k,\psi=1)}, \tag{2.6}$$

として得られる（Yusuf et al.（1985））．ただし，$E(x_k|t_k,\psi=1)$ と $\mathrm{var}(x_k|t_k,\psi=1)$ は超幾何分布の平均と分散，

$$E(x_k|t_k,\psi=1)=\frac{n_k t_k}{N_k},\quad \mathrm{var}(x_k|t_k,\psi=1)=\frac{n_k m_k t_k(N_k-t_k)}{N_k^2(N_k-1)},$$

である．

2.5 2つの極限モデル

疫学研究などでは，数個の年齢階級にデータを層別したり，あるいは年齢階級ごとに層別サンプリングを行って複数の 2×2 表を得る．この場合に適切な極限モデルは，層（表）の数Kは高々数個であり，各層の標本数 n_k，m_k が無限に増加するというものである．ただし，n_k/N_k および N_k/N_+ は有界であるとする（$N_+=\sum_k N_k$）．これを極限モデル I とよぶ．

これに対し，層別要因が性，年齢，喫煙状況，肥満度などと多岐にわたり，層の数が非常に多くなってしまったり，あるいは疾病を発生したケース1名に，様々なリスク要因をマッチさせた疾病を発生していないコントロールをサンプリングすることもよく行われる．この場合に適切な極限モデルは，層の数Kは無限に増加するが，n_k，m_k の組み合わせは有限であり，かつ N_+/K は有界，というものである（Breslow（1981））．これを極限モデルIIとよぶ．

Mantel and Haenszel（1959）は層別した 2×2 表で，リスク要因への曝露と疾病の発生に関連はないという帰無仮説を検定するための統計量として，

$$X^2_{\mathrm{MH}}=\frac{\left(\left|\sum_k x_k-\sum_k\frac{n_k t_k}{N_k}\right|-\frac{1}{2}\right)^2}{\sum_k\frac{n_k m_k t_k(N_k-t_k)}{N_k^2(N_k-1)}}, \tag{2.7}$$

を提案した．X^2_{MH} は極限モデル I，極限モデルIIどちらのもとでも近似的に自由度1のカイ二乗分布にしたがう．このためどちらの極限モデルのもとでも妥当な検定となり，いまでは Mantel-Haenszel 検定としてよくしられている．Mantel-Haenszel 検定は，積拡張超幾何分布モデルにもとづいた，$\psi=1$ の条件付きスコア検定として導くこともできる．

5節で述べるように Mantel-Haenszel 推定量は Mantel-Haenszel 検定と密接な関係にあり，Mantel-Haenszel 推定量はどちらの極限モデルのもとでも共通オッズ比の一致推定量である（Hauck（1979）; Breslow（1981））．Mantel-Haenszel 推定量は漸近有効な推定量ではないが，その効率は一般にパラメータの広い範囲で90%以上となることが報告されている（Breslow（1981）; Donner and Hauck（1986）; Hauck and Donner（1988））．Peto 推定量も Mantel-Haenszel 検定にもとづいている．しかし Peto 推定量は $\psi=1$ の近傍での一回近似であるので，$\psi=1$ 以外では共通オッズ比の一致推定量とはならず，また一般に帰無仮説 $\psi=1$ の方向にバイアスがかかる．このバイアスは共通オッズ比が1から離れるにつれ大きくなるが，それでも Mantel-Haenszel 検定にもとづいた推定量であるため，共通オッズ比の値がそれほど大きくな

表 2.　共通オッズ比の推定量の性質.

共通オッズ比の推定量	極限モデル I		極限モデル II	
	一致性	有効性	一致性	有効性
Mantel-Haenszel 推定量	○	△*	○	△*
重み付き最小二乗推定量	○	○	×	−
無条件の最尤推定量	○	○	×	−
条件付き最尤推定量	○	○	○	○
Peto 推定量	×	−	×	−

*帰無仮説 H_0：$\psi=1$ のもとでのみ漸近有効

い場合にはどちらの極限モデルのもとでもいい性質を持つことが期待できる．ただし，n_k と m_k かつ t_k と N_k-t_k が大きく食い違う場合には Peto 推定量のバイアスは深刻であり，しかも帰無仮説から離れる方向へのバイアスも起こることが報告されている（Greenland and Salvan (1990)）.

重み付き最小二乗推定量と無条件の最尤推定量は，ともに極限モデル I では共通オッズ比の一致推定量でかつ漸近有効であるが，極限モデル II ではもはや一致推定量ではない．極限モデル II では，(n_k, m_k) が中程度のサイズであっても重み付き最小二乗推定量にはかなり保守的な漸近バイアスが存在し，無条件の最尤推定量には (n_k, m_k) が小さくなるにしたがって大きい正のバイアスが存在する．$(n_k,\ m_k)\equiv(1,\ 1)$ の場合，無条件の最尤推定量は ψ^2 に確率収束する (Breslow (1981))．条件付き最尤推定量はどちらの極限モデルのもとでも共通オッズ比の一致推定量であり，かつ漸近有効である．$(n_k, m_k)\equiv(1, 1)$ の場合，条件付き最尤推定量と Mantel-Haenszel 推定量は一致する．このことも Mantel-Haenszel 推定量が好まれる理由のひとつとなっている．

表 2 に，各推定量の上記の性質をまとめておく．

3.　共通オッズ比の信頼区間

本節でははじめに重み付き最小二乗推定量と Peto 推定量の漸近分散の推定について述べ，次に Mantel-Haenszel 推定量の分散推定の歴史を概観する．これらにもとづく共通オッズ比の信頼区間はすべて漸近的な方法である．

3.1　重み付き最小二乗推定量と Peto 推定量

極限モデル I のもとでの重み付き最小二乗推定量の漸近分散の推定量は，

$$V(\log\hat{\psi}_{\mathrm{WLS}})=\frac{1}{\sum_k w_k^{-1}}, \tag{3.1}$$

である（$V(\cdot)$ は漸近分散の推定量とする）．これより，共通オッズ比の近似両側 $1-\alpha$ 信頼区間は真数をとって，

$$\hat{\psi}_{\mathrm{WLS}}\exp\left[\pm z_{\alpha/2}/\sqrt{\textstyle\sum_k w_k^{-1}}\right],$$

を得る．ただし，$z_{\alpha/2}$ は標準正規分布の上側 $100(\alpha/2)$％点である．この信頼区間は極限モデル I でのみ妥当であるが，シミュレーションの結果から $\psi=1$ から離れると挙動が悪くなることが報告されている（Mehta et al. (1985)）.

Peto 推定量の漸近分散の推定量は,

$$V(\log \hat{\psi}_{\mathrm{P}})=\frac{1}{\sum_k \frac{n_k m_k t_k(N_k-t_k)}{N_k^2(N_k-1)}}, \tag{3.2}$$

であり，共通オッズ比の近似 $1-\alpha$ 信頼区間は,

$$\hat{\psi}_{\mathrm{P}} \exp\left[\pm z_{\alpha/2} \Big/ \sqrt{\sum_k \frac{n_k m_k t_k(N_k-t_k)}{N_k^2(N_k-1)}}\ \right],$$

で与えられる．この信頼区間はオッズ比が 1 に近い場合，極限モデル I，II どちらでも用いることができるが，ランダム化臨床試験のように (t_k, N_k-t_k) または (n_k, m_k) がほぼ等しいデザインでのみ用いるべきである (Greenland and Salvan (1990)).

無条件の最尤推定量と条件付き最尤推定量の漸近分散は，Fisher 情報行列の逆行列から推定することができる．他の推定量と同様に対数スケールで信頼区間を構成した後，オッズ比のスケールに戻すのが一般的である．無条件の最尤推定量による信頼区間は漸近モデル I のもとでのみ妥当であるが，条件付き最尤推定量による信頼区間は，漸近モデル I，II どちらのもとでも妥当である．無条件の最尤推定量による信頼区間はロジスティック回帰を実行するプログラムパッケージを用いて，また条件付き最尤推定量による信頼区間は SAS PROC PHREG を用いて計算することができる.

3.2 Mantel-Haenszel 推定量

Mantel-Haenszel 推定量は条件付き最尤推定量以外で唯一，2 つの漸近モデルのもとで一致推定量となるので，簡単に計算できる漸近分散の推定量が長い間望まれていた．Mantel-Haenszel 推定量の分布も歪んでいるので，通常は対数スケールで信頼区間を構成する．$\log \hat{\psi}_{\mathrm{MH}}$ の漸近分散の推定量を V_{MH} とすると，共通オッズ比の近似 $1-\alpha$ 信頼区間は,

$$\hat{\psi}_{\mathrm{MH}} \exp[\pm z_{\alpha/2}\sqrt{V_{\mathrm{MH}}}\,],$$

から求められる．現在までに多くの V_{MH} が提案されており，以下では順を追って整理しよう．(各推定量は，提案者の頭文字を V の添字として区別する.)

Miettinen (1976) は次のような「検定にもとづく方法 (test-based method)」を提案した．ある母数を θ，その推定量を $\hat{\theta}$ とし，$g(\hat{\theta})$ を正規化かつ分散安定化変換とする．大標本のもとでの θ の近似信頼区間は,

$$g^{-1}\left\{g(\hat{\theta})\pm z_{\alpha/2}\sqrt{\mathrm{var}[g(\hat{\theta})]}\ \right\},$$

で与えられる．一般に $\mathrm{var}[g(\hat{\theta})]$ は簡単に得られなくとも，仮説 $\theta=\theta_0$ の有意性検定は簡単にできる場合がある．仮説 $\theta=\theta_0$ のもとでは，$[g(\hat{\theta})-g(\theta)]^2/\mathrm{var}[g(\hat{\theta})]$ は近似的に自由度 1 のカイ二乗分布にしたがうので，漸近的にこの式と同等である検定が存在するならば，その検定統計量を上式と等しいとおいて $\mathrm{var}[g(\hat{\theta})]$ を求める．Mantel-Haenszel 推定量の場合，g は対数変換であり，帰無仮説 $\psi=1$ の検定統計量は (2.7) 式から簡単に得られるので,

$$V_{\mathrm{M}}=\frac{(\log \hat{\psi}_{\mathrm{MH}})^2}{X_{\mathrm{MH}}^2}, \tag{3.3}$$

を得る．この近似は計算が簡単であり，Mantel-Haenszel 検定・推定量ともに漸近モデル I，II どちらでも用いることができるため，疫学の分野で非常によく用いられ，現在でも PC SAS Version 6.12 の PROC FREQ では CMH オプションを指定するとこの方法による近似信頼区間を計算する．しかし，この近似は明らかに帰無仮説 $\psi=1$ のもとでしか成り立たず (しかも対数変換は正規化変換でも分散安定化変換でもない)，オッズ比が 1 から離れるにつれ実質的な信

頼係数は $1-\alpha$ よりもずっと小さくなる（Halperin（1977); Greenland（1984））.

Hauck (1979) は, Mantel-Haenszel 推定量が S_k を重みとした各層のオッズ比の重み付き平均と書けることを利用して漸近モデル I のもとでの分散の推定量を導いた.

$$\mathrm{var}^A(\hat{\psi}_{\mathrm{MH}})=\frac{\sum_k[E^A(S_k)]^2\mathrm{var}^A(\hat{\psi}_k)}{[\sum_k E^A(S_k)]^2},$$

であり, $\mathrm{var}^A(\hat{\psi}_k)=\psi^2E^A(w_k)$ また $\mathrm{var}^A(\log\hat{\psi}_{\mathrm{MH}})=\mathrm{var}^A(\hat{\psi}_{\mathrm{MH}})/\psi^2$ であるから,

(3.4) $$V_{\mathrm{H}}=\frac{\sum_k S_k^2 w_k}{(\sum_k S_k)^2},$$

を示した. w_k として 1/2 を加えたバイアス修正を用いることを Breslow and Liang（1982), Schlesselman（1982）が提案した.

すべての層で行 $X=1$ と $X=0$（あるいは列 $Y=1$ と $Y=0$）を入れ換えても $\log\hat{\psi}_{\mathrm{MH}}$ は符号が変わるだけで分散は変わらない. しかし, 行または列を交換した場合の V_{H} を $V_{\mathrm{H}}^*=(\sum_k R_k^2 w_k)/(\sum_k R_k)^2$ とすると, 一般に $V_{\mathrm{H}}\neq V_{\mathrm{H}}^*$ であり, V_{H} は対称性を持たない (Ury (1982)). V_{H} に対称性を持たせるため, V_{H} と V_{H}^* の幾何平均をとった,

(3.5) $$V_{\mathrm{HS}}=\sqrt{V_{\mathrm{H}}\times V_{\mathrm{H}}^*},$$

が Breslow and Liang（1982）によって提案された.

漸近モデル II のもとでの $\log\hat{\psi}_{\mathrm{MH}}$ 分散の推定量は, Breslow（1981）により条件付き推定量,

(3.6) $$V_{\mathrm{BC}}=\frac{\sum_k\mathrm{var}(R_k-\psi S_k|t_k,\ \hat{\psi}_{\mathrm{MH}})}{[\sum_k E(R_k|t_k,\ \hat{\psi}_{\mathrm{MH}})]^2},$$

と, 経験的推定量,

(3.7) $$V_{\mathrm{BE}}=\frac{\sum_k(R_k-\hat{\psi}_{\mathrm{MH}}S_k)^2}{(\sum_k R_k)^2},$$

が提案された.

V_{H}, V_{BE} は仮定した極限モデルのもとでは, $\log\hat{\psi}_{\mathrm{MH}}$ の漸近分散の一致推定量となるが, ことなった極限モデルのもとでは一致推定量とはならない. もちろんサンプリングモデルに応じて使い分ければいいのだが, 計算が簡単で, どちらの極限モデルのもとでも $\log\hat{\psi}_{\mathrm{MH}}$ の漸近分散の一致推定量となるものがあれば便利である. また, 漸近モデル I と II は両極端のモデルであるため, 両方のモデルでいい性質を示す推定量は, その中間の状況（各層のサイズも層の数も中程度）でもいい性質を持つかもしれない. 条件付き推定量 V_{BC} はどちらの極限モデルのもとでも一致推定量となるが, 漸近モデル II が仮定できない場合, 拡張超幾何分布の期待値と分散の計算はやっかいである（Robins et al.（1986））.

Breslow and Liang（1982）は, どちらの極限モデルのもとでも一致推定量となる統合推定量,

(3.8) $$V_{\mathrm{BL}}=\frac{N_+V_{\mathrm{HS}}+K^2V_{\mathrm{BE}}}{N_++K^2},$$

を提案した. 極限モデル I では $N_+\to\infty$ であるから V_{HS} が優位となり, 極限モデル II では $K\to\infty$ かつ $N_+/K\to$ 有界であるから V_{BE} が優位となる.

Flanders（1985）は, 漸近モデル II の特別な場合である $(n_k,\ m_k)\equiv(1,\ m_k)$ のもとで Hauck（1979）の提案した V_{H} が一致推定量となるように修正を行った,

(3.9) $$V_{\mathrm{F}}=\frac{\sum_k S_k\left[\frac{1}{\hat{\psi}_{\mathrm{MH}}}\left(P_k+\frac{1}{N_k}\right)+Q_k-\frac{1}{N_k}\right]}{(\sum_k S_k)^2},$$

を提案した. ただし, $P_k=(x_k+m_k-y_k)/N_k$, $Q_k=(n_k-x_k+y_k)/N_k$ である. シミュレーション

の結果から，漸近モデルＩに近い状況と$(n_k, m_k)\equiv(1, m)$で良好であるばかりでなく，もっと広い範囲にわたって名義的な信頼水準に近い値を与えた（Flanders (1985)）．

Robins et al. (1986) は，$\log\hat{\psi}_{\mathrm{MH}}$の漸近分散が，

$$\mathrm{var}^A(\log\hat{\psi}_{\mathrm{MH}})=\frac{\sum_k \mathrm{var}(R_k-\psi S_k)}{[E(\sum_k R_k)]^2}, \tag{3.10}$$

であり，拡張超幾何分布にもとづく正確な分散と期待値の不偏推定量を用いるか，二項分布にもとづく分散と期待値の不偏推定量を用いれば，漸近モデルＩおよびIIのもとでの一致推定量が得られることを示した．Hauck (1979) のV_{H}は，漸近モデルＩのもとでの$R_k-\psi S_k$の漸近分散として，

$$\mathrm{var}^A(R_k-\psi S_k)=\psi^2\left[\frac{n_k m_k}{N_k}p_{0k}(1-p_{1k})\right]^2\left[\frac{1}{n_k p_{1k}(1-p_{1k})}+\frac{1}{m_k p_{0k}(1-p_{0k})}\right],$$

を用いた自然なFisher一致推定量と解釈できるが，Robins et al. (1986) はこの漸近分散のFisher一致推定量のクラスの中で$\mathrm{var}(R_k-\psi S_k)$の不偏推定量となる，

$$\widehat{\mathrm{var}}(R_k-\psi S_k)=\psi S_k\frac{\psi y_k+(m_k-y_k)}{N_k}+R_k\frac{x_k+\psi(n_k-x_k)}{N_k}, \tag{3.11}$$

を導いた．$\mathrm{var}(R_k-\psi S_k)$は，行($Y=1$, $Y=0$)または列($X=1$, $X=0$)の交換に関して$\mathrm{var}(S_k-R_k/\psi)=\mathrm{var}(R_k-\psi S_k)/\psi^2$という意味での対称性を持つが，この推定量(3.11)式は列の交換について対称ではなく，別な推定量，

$$\widehat{\mathrm{var}}(R_k-\psi S_k)=\psi S_k\frac{x_k+\psi(n_k-x_k)}{N_k}+R_k\frac{\psi y_k+(m_k-y_k)}{N_k}, \tag{3.12}$$

が得られてしまう．このため，Robins et al. は(3.11)式と(3.12)式の平均をとった，

$$\widehat{\mathrm{var}}(R_k-\psi S_k)=\frac{1}{2}(R_k+\psi S_k)(P_k+\psi Q_k), \tag{3.13}$$

を用いて，ψには一致推定量である$\hat{\psi}_{\mathrm{MH}}$を代入した$\log\hat{\psi}_{\mathrm{MH}}$の漸近分散の推定量，

$$V_{\mathrm{RBG}}=\frac{1}{2}\left[\frac{\sum_k P_k R_k}{(\sum_k R_k)^2}+\frac{\sum_k(Q_k R_k+P_k S_k)}{(\sum_k R_k)(\sum_k S_k)}+\frac{\sum_k Q_k S_k}{(\sum_k S_k)^2}\right], \tag{3.14}$$

を提案した．

Phillips and Holland (1987) は，拡張超幾何分布モデルから，

$$\widehat{\mathrm{var}}(R_k-\psi S_k|t_k, \psi)=\psi S_k\left[\left(P_k+\frac{1}{N_k}\right)+\psi\left(Q_k-\frac{1}{N_k}\right)\right], \tag{3.15}$$

を導き，V_{F}はこの推定量にもとづいていることを明らかにした．Flanders (1985) のシミュレーション結果が広い範囲にわたって良好であったのも，V_{F}が特別な場合だけでなく漸近モデルII一般で一致推定量であるからであった．Phillips and Holland は(3.15)式が行または列の交換に対して対称ではないことから，

$$\widehat{\mathrm{var}}(R_k-\psi S_k|t_k, \psi)=R_k\left[\left(P_k-\frac{1}{N_k}\right)+\psi\left(Q_k+\frac{1}{N_k}\right)\right], \tag{3.16}$$

を導き，(3.15)式と(3.16)式の平均が(3.13)式となることから，V_{F}とV_{RBG}との関係を明らかにした．

$\mathrm{var}(R_k-\psi S_k)$の行または列の交換に対する対称性を保った自然な推定量は，二項分布モデルにもとづいてSato (1990) が，

$$\widehat{\mathrm{var}}(R_k-\psi S_k)=\psi\left[\left(Q_k+\frac{1}{N_k}\right)R_k+\left(P_k+\frac{1}{N_k}\right)S_k\right], \tag{3.17}$$

を与えた．Sato (1990) はさらに，$\log\hat{\psi}_{\mathrm{MH}}$の分散ではなく，Mantel-Haenszel推定量が，

$$\textstyle\sum_k R_k - \psi \sum_k S_k = 0\,,$$

という不偏な推定方程式の解であること（Davis (1985)）に着目して，

$$(3.18)\qquad \frac{\left(\left|\sum_k R_k - \psi \sum_k S_k\right| - c\right)^2}{\psi \sum_k \left[\left(Q_k + \dfrac{1}{N_k}\right) R_k + \left(P_k + \dfrac{1}{N_k}\right) S_k\right]} = z_{\alpha/2}^2\,,$$

の 2 根を信頼限界とすることを提案した．ただし，c は連続修正を行う場合 $(1+\psi)/4$，そうでなければゼロである．$\log \hat{\psi}_{\mathrm{MH}}$ にもとづく信頼区間では，$\sum_k R_k$ または $\sum_k S_k$ のどちらかがゼロの場合不定となってしまうが，(3.18)式の推定方程式にもとづく方法では片側信頼区間を求めることができる．また，(1, 1)マッチングのとき Mantel-Haenszel 推定量は条件付き最尤推定量に一致するが，(3.18)式も拡張超幾何分布の正規近似にもとづく信頼区間(Breslow and Day (1980)) に一致する．

Yanagimoto (1992) は対称性だけではなく，$\psi=1$ の場合に，

$$\mathrm{var}(R_k - S_k | t_k,\ \psi=1) = \mathrm{var}(x_k - n_k t_k / N_k | t_k,\ \psi=1)\,,$$

であることから，

$$(3.19)\qquad \widehat{\mathrm{var}}(R_k - \psi S_k) = \psi \frac{n_k m_k t_k (N_k - t_k)}{N_k^2 (N_k - 1)} - (\psi - 1)^2 \frac{R_k S_k}{N_k - 1}\,,$$

を(3.18)式の分母に用いることを提案した．Clayton and Hills (1993) は $\psi=1$ の近傍では $\log \hat{\psi}_{\mathrm{MH}}$ の分散が，

$$V_{\mathrm{CH}} = \frac{\sum_k \dfrac{n_k m_k t_k (N_k - t_k)}{N_k^2 (N_k - 1)}}{\left(\sum_k R_k\right)\left(\sum_k S_k\right)}\,,$$

で近似できると証明ぬきで述べたが，これは(3.19)式の右辺第 1 項のみをとることに相当する．

4.　共通オッズ比の信頼区間の比較

3 節で紹介した共通オッズ比の信頼区間は，漸近モデル I，II どちらかの大標本のもとでの近似信頼区間であった．本節では，これらの信頼区間の大標本および小標本での挙動について，シミュレーションの結果をまとめる．特に小標本の場合について，積拡張超幾何分布にもとづく正確な信頼区間との比較を行う．最後に実際の数値例により，いくつかの方法を比較する．

4.1　正確な信頼区間

積拡張超幾何分布モデルからは，共通オッズ比の正確な信頼区間を求めることができる．$x_+ = \sum_k x_k$, $L_k = \max(0,\ t_k - m_k)$, $U_k = \min(n_k,\ t_k)$, $L_+ = \sum_k L_k$, $U_+ = \sum_k U_k$, $W(k,\ x) = \binom{n_k}{x}\binom{m_k}{t_k - x}$ とする．x_+ は共通オッズ比 ψ 推定の十分統計量であり，その条件付き分布は，

$$\Pr(x_+ | \psi) = \frac{C_{x_+} \psi^{x_+}}{\sum_{b=L_+}^{U_+} C_b \psi^b}\,,$$

となる (Zelen (1971))．ただし，$C_a = \sum_{x \in a} \prod_k W(k, x)$, $x \in a$ は a を K 個の層に分割するすべての可能な組み合わせを表す．これより，共通オッズ比の $1-\alpha$ 信頼区間（ψ_L, ψ_U）は，

$$(4.1)\qquad \sum_{a=x_+}^{U_+} \Pr(a | \psi_L) = \alpha/2\,, \quad \sum_{a=L_+}^{x_+} \Pr(a | \psi_U) = \alpha/2\,,$$

の解である．ただし，$x_+=L_+$ の場合 $\psi_L=0$，$x_+=U_+$ の場合 $\psi_U=\infty$ とする．(4.1)式を解くためのプログラムは Thomas(1975)が与えたが，より高速なネットワークアルゴリズムを Mehta et al. (1985)，Vollset et al. (1991) が発表している．以下では，Takagi (1990) にしたがって，簡単に説明する．

まずあらかじめ，各層ごとに，$W(k,x)$，$L_k \le x \le U_k$ を求めておく．係数 C_a は，以下の手順で求められる．

［1］ $C(1, x_1)=W(1, x_1)$，$L_1 \le x_1 \le U_1$ とおく．

［2］ $SL_k=\sum_{i=1}^{k} L_i$，$SU_k=\sum_{i=1}^{k} U_i$，$SL_k \le z_k \le SU_k$，$L_{k+1} \le x_{k+1} \le U_{k+1}$ とし，

$$C(k+1, z_{k+1})=\sum_{z_{k+1}=z_k+x_{k+1}} C(k, z_k)\times W(k+1, x_{k+1}),$$

を求める．右辺は，$z_{k+1}=z_k+x_{k+1}$ を満たすすべての和を意味する．

［3］ ［2］のステップを $k=K$ まで繰り返す．

最終的に $C_a=C(K, z_K)$ として求められる（ここで，$a=z_K$ である）．

このアルゴリズムによる最大の計算量は，$n_{\max}=\max[n_1,\ldots, n_K]$ とすると，

$$\sum_{k=2}^{K}\left(\sum_{i=1}^{k-1} n_i+1\right)(n_k+1) \le (K-1)(n_{\max}+1)+\frac{K(K-1)}{2}n_{\max}(n_{\max}+1),$$

となり，$(nK)^2$ のオーダーとなる．実際の計算では，計算途中で数値オーバーフローを起こす可能性が高く，その対策も必要となる（オーバーフロー対策の詳細は，Takagi (1990) 参照）．

簡単な数値例により上記のアルゴリズムの説明を行う．次のような3個の 2×2 表を考えよう．$n_k=m_k=t_k=2$ とする．

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$

簡単な計算から，$W(k, 0)=1$，$W(k, 1)=4$，$W(k, 2)=1$，$k=1, 2, 3$ となる．

左から2つの表についての手順を，表3.1に示した．表3.1の左側の z_1 は1番目の 2×2 表の左上のセルが (0, 1, 2) となる場合を示す（$z_1=x_1$ である）．その右側の $C(1, z_1)$ は，その場合の数を示す．最上段の x_2 は2番目の 2×2 表の左上のセルが (0, 1, 2) となる場合を示す．その下の $W(2, x_2)$ は，その 2×2 表での場合の数を示す．2つの 2×2 表を組み合わせた場合の計算は，x_2 の値の下の欄に，$C(1, x_1)$ の値を1行ずつ表のようにずらして記入する．その数値に対応する $W(2, x_2)$ を掛け算する．掛け算の右横に，2つの 2×2 表での左上のセルに入る数値の和を記入する（$z_2=z_1+x_2$）．その右横に，各行で計算した数値の和を記入する．ここでは，左上のセルの値が (0, 1, 2, 3, 4) になる場合は，1, 8, 18, 8, 1 通りずつであることが分かる．

表3.1． 2つの 2×2 表での係数 C_a の計算手順．

$z_1=x_1$	$C(1, z_1)$	x_2	0	1	2	$z_2=z_1+x_2$	$C(2, z_2)$
		$W(2, x_2)$	1	4	1		
0	1		1×1	—	—	0	1
1	4		4×1	1×4	—	1	8
2	1		1×1	4×4	1×1	2	18
—	—		—	1×4	4×1	3	8
—	—		—	—	1×1	4	1

表 3.2.　3 番目の 2×2 表を加えた場合の C_a の計算手順.

z_2	$C(2, z_2)$	x_3	0	1	2	$z_3=z_2+x_3$	$C(3, z_3)$
		$W(3, x_3)$	1	4	1		
0	1		1×1	—	—	0	1
1	8		8×1	1×4	—	1	12
2	18		18×1	8×4	1×1	2	51
3	8		8×1	18×4	8×1	3	88
4	1		1×1	8×4	18×1	4	51
—	—		—	1×4	8×1	5	12
—	—		—	—	1×1	6	1

表 3.2 は，表 3.1 で求めた結果を利用して，さらに 1 つ 2×2 表を加えた場合について計算を繰り返している．すなわち，z_2 と $C(2, z_2)$ を表の左側に記し，表頭の 2 行には x_3 と $W(3, x_3)$ の値を記入する．計算の方法は，表 3.1 についてと全く同様である．$C(2, z_2)$ の値を x_3 の値の各列に 1 行ずつずらして記入し，x_3 の値に対応した $W(3, x_3)$ の値を掛け算する．計算結果を各行ごとに合計した値が，$z_3(=z_2+x_3)$ の各値に対応した $C(3, z_3)$ となる．この例では，$C(3, z_3)$, $z_3=0, 1, 2, 3, 4, 5, 6$ が正確な確率の計算に必要な係数を与えることになる．仮に，2×2 表がもう 1 つ追加されても，同様の方法で簡単に係数の計算が行える．

表3.2から，$\psi=1$とした場合，$x_+\geq 3$となるFisherの正確な確率は(4.1)式から以下のように，

$$\sum_{a=3}^{6}\Pr(a|\psi=1)=\frac{\sum_{a=3}^{6}C_a}{\sum_{b=0}^{6}C_b}=\frac{88+51+12+1}{1+12+51+88+51+12+1}=0.7037,$$

と簡単に計算できる．

4.2　シミュレーションと実際例による比較

まず，共通オッズ比の信頼区間の構成について，近似的な方法と正確な方法を比較するためのシミュレーションの結果（Takagi (1991)）について説明する．

シミュレーションの設定条件として，matched design の場合，balanced design の場合，imbalanced design の場合の 3 通りについて，ケース数 n とコントロール数 m および層の数 K を数通りに変えて行った．共通オッズ比は，$\psi=1, 3.5, 6.5$ の 3 通りとした．ケース群での要因曝露割合を，matched design の場合には，$0.3+0.5(k-1)/(K-1)$ とし，その他の場合には，$0.05+0.8(k-1)/(K-1)$ とした．これらの組み合わせについて，各 1 万回の試行を行い，設定した共通オッズ比が，求めた 95%信頼区間に含まれない回数を求めた．計算は統計数理研究所の大型計算機を用い，乱数は物理乱数発生装置による一様乱数を使用した．

表 4 に共通オッズ比 $\psi=1$ の場合，表 5 に $\psi=3.5$ の場合，表 6 に $\psi=6.5$ の場合の結果をそれぞれ示した．表中で「正確な方法」のうち「Exact-P」は(4.1)式による通常の正確な信頼区間を示す．「Mid-P」は正確な方法が保守的な結果を与えることから，観察された層別 2×2 表の確率を 1/2 倍して計算する方法であり（Lancaster (1961); Vollset and Hirji (1991)），

$$
\frac{1}{2}\Pr(x_+|\psi_L)+\sum_{a=x_++1}^{U_+}\Pr(a|\psi_L)=\alpha/2\,, \tag{4.2}
$$
$$
\frac{1}{2}\Pr(x_+|\psi_U)+\sum_{a=L_+}^{x_+-1}\Pr(a|\psi_U)=\alpha/2\,,
$$

となるように信頼区間を構成する.

表中の「CMLE」は(2.5)式の条件付き最尤推定値を対数変換後その分散を用いる方法であり,「RBG」は(2.2)式のMantel-Haenszel推定量を対数変換後(3.18)式による分散を用いる方法である. (3.18)式の推定方程式による方法では,「修正あり」は連続修正をした場合,「修正なし」は連続修正を行わない場合を示す.

表4ではmatched designの場合, (3.18)式の修正ありが最も保守的な結果を与え, exact-Pも同様に保守的な結果を与えた. Mid-Pと(3.18)式の修正なしではほぼ名義水準通りの結果を

表4. 95%信頼区間が母共通オッズ比を含まない割合(%): $\psi=1$ の場合.

			正確な信頼区間				対数法				(3.18)式			
			Exact-P		Mid-P		CMLE		RBG		修正なし		修正あり	
n	m	K	下側	上側	下側	上側	下側	上側	下側	上側	下側	上側	下側	上側
1	1	200	2.0	2.0	2.5	2.5	2.4	2.4	2.4	2.4	2.5	2.5	1.9	1.9
		100	1.8	1.8	2.5	2.6	2.2	2.3	2.2	2.3	2.5	2.6	1.7	1.7
1	2	100	1.8	1.9	2.4	2.4	2.3	2.4	2.3	2.4	2.4	2.5	1.8	1.8
		50	1.7	1.7	2.4	2.4	2.3	2.4	2.3	2.3	2.5	2.6	1.6	1.6
1	4	50	1.7	1.8	2.4	2.5	2.3	2.5	2.2	2.5	2.4	2.6	1.6	1.8
		25	1.5	1.5	2.4	2.3	2.2	2.3	2.1	2.2	2.5	2.5	1.4	1.4
1	8	50	1.8	1.7	2.4	2.4	2.3	2.4	2.3	2.4	2.4	2.5	1.7	1.7
		25	1.4	1.5	2.3	2.3	2.1	2.3	2.0	2.3	2.4	2.6	1.3	1.5
5	5	20	1.7	1.7	2.4	2.3	2.4	2.3	2.4	2.3	2.6	2.5	1.8	1.7
		10	1.4	1.5	2.3	2.3	2.3	2.3	2.3	2.3	2.6	2.6	1.4	1.5
15	15	20	2.1	2.0	2.5	2.5	2.5	2.5	2.5	2.5	2.6	2.5	2.1	2.1
		10	1.9	1.8	2.5	2.4	2.5	2.4	2.5	2.4	2.6	2.5	1.9	1.8
		5	1.6	1.6	2.3	2.4	2.3	2.4	2.4	2.4	2.6	2.7	1.6	1.6
30	30	20	2.1	2.1	2.5	2.4	2.5	2.4	2.5	2.4	2.5	2.5	2.1	2.1
		10	1.9	2.0	2.4	2.5	2.4	2.5	2.4	2.5	2.5	2.5	1.9	2.0
		5	1.8	1.7	2.4	2.3	2.4	2.4	2.5	2.4	2.6	2.5	1.9	1.8
5	10	20	1.8	1.8	2.4	2.4	2.4	2.4	2.4	2.4	2.5	2.5	1.8	1.8
		10	1.6	1.6	2.3	2.4	2.3	2.4	2.3	2.4	2.6	2.6	1.6	1.6
15	30	20	2.0	2.1	2.4	2.6	2.4	2.5	2.4	2.6	2.5	2.6	2.0	2.1
		10	1.9	1.9	2.5	2.4	2.5	2.4	2.5	2.4	2.6	2.5	1.9	1.9
		5	1.7	1.7	2.4	2.4	2.4	2.4	2.4	2.5	2.6	2.6	1.7	1.8
30	60	10	2.2	2.0	2.5	2.4	2.5	2.4	2.6	2.4	2.6	2.4	2.2	2.0
		5	1.9	1.9	2.4	2.5	2.4	2.5	2.4	2.5	2.5	2.6	1.9	1.9

与えた．対数法はそれらの中間的な結果であり，CMLE と RBG の結果にほとんど差はなかった．Balanced design の場合，exact-P と (3.18) 式の修正ありが保守的な結果を与えた．他の 4 方法は，ほぼ名義水準通りの結果であったが，(3.18) 式の修正なしでは，やや liberal な結果となった．Imbalanced design の場合，exact-P と (3.18) 式の修正ありがほぼ同様な保守的な結果を与えた．他の 4 方法は，ほぼ名義水準通りの結果であったが，(3.18) 式の修正なしでは，やや liberal な結果となった．

表 5 の共通オッズ比が 3.5 の場合，matched design では，(3.18) 式の修正ありが最も保守的であり，とくに下限は極めて保守的であった．Exact-P も保守的な結果を与えたが，mid-P は最も名義水準に近い結果を与えた．他の 3 方法は下限は保守的，上限は liberal であり，(3.18) 式の修正なしはとくに上限は liberal であった．Balanced design も matched design と大差ない結果であり，mid-P がほぼ名義水準通りの結果を与えた．n，m，K の値が大きい場合には，

表 5．95%信頼区間が母共通オッズ比を含まない割合（%）：ψ=3.5 の場合．

			正確な信頼区間				対数法				(3.18) 式			
n	m	K	Exact-P		Mid-P		CMLE		RBG		修正なし		修正あり	
			下側	上側	下側	上側	下側	上側	下側	上側	下側	上側	下側	上側
1	1	200	1.8	2.0	2.4	2.4	1.9	2.7	1.9	2.7	2.0	2.8	1.5	2.5
		100	1.5	1.7	2.4	2.3	1.4	2.9	1.4	2.9	2.1	2.9	1.2	2.1
1	2	100	1.7	1.8	2.4	2.4	1.9	2.6	2.0	2.6	2.1	2.8	1.3	2.0
		50	1.5	1.7	2.4	2.4	1.6	2.7	1.7	2.7	2.0	3.1	1.0	1.9
1	4	50	1.6	1.7	2.3	2.4	2.0	2.5	2.0	2.6	2.0	2.9	1.1	1.8
		25	1.3	1.5	2.2	2.3	1.4	2.5	1.5	2.5	1.7	3.1	0.6	1.7
1	8	50	1.7	1.8	2.4	2.4	2.2	2.5	2.1	2.5	2.1	2.8	1.2	1.8
		25	1.5	1.4	2.4	2.2	2.0	2.3	2.0	2.3	2.0	2.8	0.8	1.5
5	5	20	1.6	1.7	2.3	2.4	2.0	2.7	2.1	2.6	2.2	2.8	1.3	1.9
		10	1.3	1.5	2.1	2.3	1.5	2.7	1.6	2.6	2.0	3.1	0.8	1.8
15	15	20	1.9	2.0	2.4	2.4	2.3	2.6	2.3	2.5	2.3	2.6	1.7	2.0
		10	1.7	1.9	2.4	2.4	2.2	2.6	2.3	2.6	2.4	2.7	1.5	1.9
		5	1.4	1.6	2.2	2.4	1.9	2.6	2.0	2.6	2.3	2.8	1.1	1.6
30	30	20	2.1	2.1	2.5	2.5	2.4	2.6	2.5	2.5	2.5	2.6	2.0	2.1
		10	1.9	2.0	2.4	2.5	2.3	2.6	2.4	2.6	2.4	2.6	1.8	2.0
		5	1.7	1.8	2.4	2.4	2.2	2.6	2.2	2.6	2.4	2.7	1.5	1.8
5	10	20	1.7	1.8	2.4	2.4	2.2	2.6	2.3	2.5	2.4	2.7	1.5	1.8
		10	1.5	1.6	2.3	2.3	2.0	2.6	2.1	2.5	2.4	2.8	1.1	1.6
15	30	20	2.1	2.2	2.5	2.6	2.4	2.7	2.5	2.6	2.5	2.7	1.9	2.1
		10	1.8	1.9	2.4	2.4	2.3	2.5	2.4	2.5	2.4	2.6	1.6	1.9
		5	1.5	1.7	2.3	2.4	2.0	2.6	2.1	2.5	2.3	2.7	1.2	1.7
30	60	10	2.0	2.0	2.5	2.4	2.4	2.5	2.4	2.5	2.5	2.5	1.9	2.0
		5	1.8	1.9	2.5	2.4	2.3	2.6	2.4	2.5	2.5	2.6	1.6	1.9

CMLE, RBG, (3.18)式の修正なしはほぼ名義水準通りの結果を与えたが，上限がやや liberal になった．Imbalanced design の場合は，exact-P と(3.18)式の修正ありは保守的な結果であり，とくに(3.18)式で下限は保守的であった．Mid-P ではほぼ名義水準通りであった．(3.18)式の修正なしでもほぼ名義水準通りであったが，上限がやや liberal になっていた．CMLE と RBG では，n, m, K が小さい場合，下限がやや保守的になった．n, m, K が大きい場合，ほぼ名義水準通りの結果であったが，CMLE の上限はやや liberal であった．

表6の共通オッズ比が6.5の場合，matched design では，(3.18)式の修正ありが最も保守的な結果を与えたが，$n:m$ が 1:1 の場合の上限は名義水準に近かった．対数法の CMLE と RBG では下限は保守的であり，とくに層の個数が小さい場合には極端に保守的になった．2つの対数法での上限は liberal な結果を与えた．(3.18)式の修正なしでは，下限は保守的，上限は liberal であった．Mid-P での下限はやや保守的な傾向があるが，ほぼ名義水準通りの結果であっ

表6．95%信頼区間が母共通オッズ比を含まない割合（%）：ψ=6.5 の場合．

			正確な信頼区間				対数法				(3.18) 式			
n	m	K	Exact-P		Mid-P		CMLE		RBG		修正なし		修正あり	
			下側	上側	下側	上側	下側	上側	下側	上側	下側	上側	下側	上側
1	1	200	1.6	2.0	2.3	2.4	1.5	2.9	1.5	2.9	1.7	3.2	1.1	2.4
		100	1.2	1.7	2.1	2.3	0.4	3.0	0.4	3.0	1.4	3.2	0.6	2.2
1	2	100	1.5	1.9	2.3	2.5	1.5	2.9	1.6	2.9	1.7	3.2	0.9	2.1
		50	1.2	1.7	2.1	2.4	0.4	2.9	0.5	3.0	1.4	3.4	0.4	2.0
1	4	50	1.5	1.7	2.4	2.4	1.7	2.7	1.7	2.7	1.7	3.1	0.7	1.8
		25	0.8	1.4	1.8	2.2	0.3	2.6	0.3	2.7	1.1	3.4	0.1	1.5
1	8	50	1.6	1.7	2.4	2.4	2.1	2.5	2.0	2.6	1.9	2.9	0.8	1.6
		25	1.4	1.5	2.2	2.3	1.6	2.5	1.5	2.6	1.6	3.2	0.4	1.5
5	5	20	1.4	1.6	2.3	2.4	1.7	2.8	1.7	2.7	1.9	3.0	0.8	1.8
		10	0.8	1.4	1.7	2.3	0.6	2.8	0.7	2.7	1.3	3.2	0.2	1.6
15	15	20	1.9	1.9	2.4	2.4	2.2	2.6	2.3	2.6	2.3	2.6	1.5	1.9
		10	1.7	1.7	2.4	2.4	2.0	2.7	2.1	2.6	2.2	2.7	1.2	1.8
		5	1.3	1.5	2.2	2.3	1.5	2.8	1.6	2.7	1.9	3.0	0.7	1.6
30	30	20	2.1	2.2	2.5	2.6	2.3	2.7	2.3	2.7	2.3	2.7	1.8	2.1
		10	2.0	1.9	2.6	2.5	2.3	2.7	2.4	2.6	2.5	2.7	1.6	1.9
		5	1.6	1.7	2.3	2.3	2.0	2.7	2.0	2.5	2.2	2.7	1.1	1.7
5	10	20	1.7	1.8	2.4	2.4	2.1	2.7	2.1	2.7	2.2	2.9	1.1	1.8
		10	1.3	1.6	2.2	2.3	1.6	2.7	1.8	2.6	2.1	2.9	0.6	1.5
15	30	20	2.0	2.1	2.5	2.5	2.4	2.6	2.4	2.6	2.5	2.6	1.7	2.0
		10	1.8	1.8	2.4	2.3	2.2	2.6	2.3	2.5	2.4	2.6	1.4	1.6
		5	1.5	1.7	2.3	2.4	1.9	2.8	2.0	2.7	2.2	2.9	1.0	1.6
30	60	10	2.1	2.0	2.5	2.4	2.4	2.6	2.4	2.5	2.4	2.6	1.7	1.9
		5	1.8	1.9	2.4	2.4	2.1	2.7	2.2	2.6	2.3	2.7	1.3	1.8

た．Balanced design の場合，exact-P と(3.18)式の修正ありが保守的な結果を与え，(3.18)式での下限はより保守的であった．Mid-P ではほぼ名義水準通りであったが，n, m が小さく層の数が少ないと下限が保守的になっていた．CMLE と RBG はほぼ同様の結果であり，下限がやや保守的，上限がやや liberal になっていた．とくに，層の数が少ない場合に，保守的になる傾向が強かった．(3.18)式の修正なしでは，下限は対数法よりも名義水準に近く，上限はより liberal になっていた．Imbalanced design の場合，(3.18)式の修正ありが最も保守的な結果を与え，とくに下限がより保守的であった．Exact-P も保守的な結果を与えたが，mid-P ではほぼ名義水準通りの結果であった．CMLE と RBG では層の数が少ない場合には下限はやや保守的になり，層の数が多ければ名義水準に近くなっていたが，上限はやや liberal となっていた．(3.18)式の修正なしでは，対数法よりも下限はより名義水準に近くなっていたが，上限はより liberal となっていた．

シミュレーション全体を通して，対数法を用いる場合には，条件付き最尤推定量にもとづいても，Mantel-Haenszel 推定量にもとづいても，差はほとんどなかった．Exact-P と(3.18)式の連続修正を行う方法は，最も保守的な結果を与えた．Mid-P はすべての場合，ほぼ名義水準通りの結果を与えた．また，(3.18)式で連続修正しない場合も，信頼区間の下限については近似的な方法として最も名義水準に近い結果を与えた．連続修正を行った場合には，(3.18)式と exact-P 法はほぼ一致した結果を示したが，保守的な結果を与える傾向があった．

表 7 は，肝炎治療にアルファインターフェロンを使用し，肝炎マーカーの一つである HBeAg 消失を調べた臨床試験のメタアナリシスデータである（Emerson (1994)）．このデータについて，各手法による共通オッズ比の推定と信頼区間を比較した結果を表 8 に示した．表中，「Yanagimoto」は(3.18)式に Yanagimoto (1992) の分散(3.19)式を用いたものである．表 7 から，Peto 法と重み付き最小二乗法では，共通オッズ比の推定値，信頼区間ともに他の方法に比べ非常に保守的となった（共通オッズ比は他の方法のほぼ 1/2）．一方，Mantel-Haenszel 推定値は 9.96，（無条件の）最尤推定値は 11.28，条件付き最尤推定値は 10.31 と，最尤推定値が

表 7．アルファインターフェロンのメタアナリシス（Emerson (1994)）．

試験	アルファインターフェロン			コントロール		
	HBeAg の消失		合計	HBeAg の消失		合計
	あり	なし		あり	なし	
	x_k	n_k-x_k	n_k	y_k	m_k-y_k	m_k
1	4	10	14	0	5	5
2	1	11	12	1	11	12
3	2	8	10	0	10	10
4	6	26	32	0	9	9
5	9	45	54	0	18	18
6	7	11	18	0	18	18
7	5	13	18	0	6	6
8	5	5	10	1	9	10
9	10	0	10	2	8	10
10	3	13	16	2	12	14
11	7	8	15	2	12	14

表8. 共通オッズ比と信頼区間の比較.

		点推定値	95%信頼区間	
			下限	上限
Mantel-Haenszel 推定量		9.96		
RBG			3.99	24.85
(3.18)式	修正なし		4.09	24.25
	修正あり		3.77	29.18
Yanagimoto	修正なし		3.91	22.80
	修正あり		3.61	26.69
Peto 法		5.42	3.02	9.72
重み付き最小二乗法*		5.31	2.42	11.66
最尤推定法		11.28	4.43	28.71
条件付き最尤推定法		10.31	4.13	25.74
Exact-P			4.05	31.70
Mid-P			4.32	32.56

*ゼロセルには½を加えた

若干大きいが，ほぼ同じ値であった．どの値も Peto 法の 95%信頼上限よりも大きくなっている．

最尤推定法による信頼区間は mid-P とほぼ同じであった．また，シミュレーション結果で示したように，RBG，(3.18)式(連続修正なし)，条件付き最尤推定法による信頼区間はほぼ等しく，exact-P と(3.18)式(連続修正あり)もほぼ等しかった．Yanagimoto の方法は(3.18)式と似た結果を与えるが，上限がやや小さくなる傾向にあった．

5. 複数の $2\times J$ 表への拡張

多くの疫学研究では，リスク要因への曝露状況は「あり，なし」と2値で測定されるだけではなく，いくつかの用量レベル（一日喫煙本数など）ごとに測定される．そのような場合，リスク要因の用量に応じて疾病発生も上昇しているかどうか，いわゆる用量・反応関係が問題となる．用量・反応関係を調べるためには，ロジスティック回帰などの線形回帰型のモデルが多用されているが，回帰係数の推定には無条件の最尤推定が適用されていることが多く，漸近モデルIIの下で推定量の一致性が成り立たない．また，特に反応が線形でない場合など，適用されるモデルによって推定値が大きく変化することもある．このようなとき，Mantel-Haenszel タイプの推定法が有効である．本節では，複数のオッズ比の推定問題に対し Mantel-Haenszel の方法の拡張を述べる．

5.1 Mantel-Haenszel 推定量の単純な適用の非整合性

表9に第 k 層の $2\times J$ 表を示す．$\{(x_{1k},\ldots,x_{Jk});(y_{1k},\ldots,y_{Jk})\}$ は生起確率 $\{(p_{11k},\ldots,p_{1Jk});(p_{01k},\ldots,p_{0Jk})\}$，サイズ (n_k, m_k) の独立な多項系列とする $(k=1,\ldots,K)$．リスク要因の曝露状況を X の J 個のレベル(なし$=1$，少しあり$=2,\ldots$，顕著にあり$=J$)，疾病発生状況を Y(あり$=1$，なし$=0$)で表している．

表 9． 第 k 層の $2\times J$ 表．

	X=1	X=2	$\cdots$	X=J	合計
Y=1	x_{1k}	x_{2k}	$\cdots$	x_{Jk}	n_k
Y=0	y_{1k}	y_{2k}	$\cdots$	y_{Jk}	m_k
合計	t_{1k}	t_{2k}	$\cdots$	t_{Jk}	N_k

表 9 で，曝露レベル $X=j$ をベースラインとし $X=j'$ に注目して得られる，すべての層に共通なオッズ比を，

$$\psi_{j'}^{(j)}=\frac{p_{1j'k}p_{0jk}}{p_{0j'k}p_{1jk}},$$

で表す．このような組 (j, j') の取り方は $J(J-1)/2$ 個ある．しかし関係式，

$$\psi_{j'}^{(j)}=\frac{\psi_{j'}^{(1)}}{\psi_{j}^{(1)}}, \tag{5.1}$$

が成り立つので $J-1$ 個の $\psi_{j}^{(1)}$ $(j=2,\dots,J)$ さえ推定しておけば，もし推定量が上の関係式を満たすという整合性をもっているなら，この関係式から $\psi_{j'}^{(j)}$ の推定値が得られる．

残念なことに，Mantel-Haenszel 推定量の単純な適用はこの整合性をもたない．つまり，曝露レベル $X=j$ と $X=j'$ だけを抜き出したときの 2×2 表から得られる Mantel-Haenszel 推定量と，曝露レベル $X=1$ と $X=j$，および曝露レベル $X=1$ と $X=j'$ だけを抜き出したときの二つの 2×2 表から得られる Mantel-Haenszel 推定量の間には上の関係式が成り立たない．このことは，推定の手順によって異なる値の推定値が得られることを意味する．これではまずいので Greenland (1989)，Yanagawa and Fujii (1990, 1995) は上の関係式を満たすように Mantel-Haenszel 推定量を一般化した．以下に，その紹介を行う．

5.2 Greenland-Yanagawa-Fujii 推定量

Mickey and Elashoff (1985) は，層別された $2\times J$ 表に対数線形モデル，

$$\log E(x_{jk})=\lambda_0+\lambda^{Y}+\lambda_j^{X}+\lambda_k^{C}+\lambda_{1j}^{YX}+\lambda_{1k}^{YC}+\lambda_{jk}^{XC},$$
$$\log E(y_{jk})=\lambda_0-\lambda^{Y}+\lambda_j^{X}+\lambda_k^{C}-\lambda_{1j}^{YX}-\lambda_{1k}^{YC}+\lambda_{jk}^{XC},$$

を適用して要因 Y と要因 X の関連性を解析することを考えた．ただし，C は層別に用いた交絡要因を表し，

$$\textstyle\sum_j\lambda_j^{X}=\sum_j\lambda_{1j}^{YX}=\sum_j\lambda_{jk}^{XC}=0,$$
$$\textstyle\sum_k\lambda_k^{C}=\sum_k\lambda_{1k}^{YC}=\sum_k\lambda_{jk}^{XC}=0,$$

である．このとき，

$$\log\psi_{j'}^{(j)}=\log\frac{E(x_{j'k})E(y_{jk})}{E(y_{j'k})E(x_{jk})}=2(\lambda_{1j'}^{YX}-\lambda_{1j}^{YX}), \tag{5.2}$$

が成り立ち，要因 Y と要因 X の関連性を示すパラメータ $\lambda_{1j'}^{YX}$ が，

$$\lambda_{1j'}^{YX}=\frac{1}{2J}\textstyle\sum_j\log\psi_{j'}^{(j)}, \tag{5.3}$$

で表される．

さて，通常の対数線形モデルによる解析では積ポアソンモデルを用いて $\lambda_{1j'}^{YX}$ の無条件の最尤推定量が算出される．ところが，2.5 節で指摘したように，無条件の最尤推定量は漸近モデル II のもとでは一致推定量ではない．そこで Mickey and Elashoff (1985) は (5.3) 式の右辺の $\psi_{j'}^{(j)}$

に Mantel-Haenszel 推定量 $\hat{\phi}^{(j)}_{\mathrm{MH},j'}$ を代入した，

$$\hat{\lambda}^{YX}_{1j'}=\frac{1}{2J}\sum_j \log \hat{\phi}^{(j)}_{\mathrm{MH},j'},$$

を $\lambda^{YX}_{1j'}$ の推定量とした．ただし Mantel-Haenszel 推定量は，

$$\hat{\phi}^{(j)}_{\mathrm{MH},j'}=\frac{\sum_k x_{j'k}y_{jk}/N_k}{\sum_k y_{j'k}x_{jk}/N_k},$$

で与えられる．$\hat{\lambda}^{YX}_{1j'}$ は漸近モデル I，およびIIの下で $\lambda^{YX}_{1j'}$ の一致推定量となる．

Greenland (1989) は，この推定量を再び(5.2)式に代入して得られる関係式，

$$\log \hat{\phi}^{(j)}_{\mathrm{G},j'}=\frac{1}{J}\sum_{s=1}^{J}(\log \hat{\phi}^{(s)}_{\mathrm{MH},j'}-\log \hat{\phi}^{(s)}_{\mathrm{MH},j}),$$

から $\hat{\phi}^{(j)}_{\mathrm{G},j'}$ を算出して $\phi^{(j)}_{j'}$ の推定量とすることを提案した．$\hat{\phi}^{(j)}_{\mathrm{G},j'}$ は(5.1)式を満たすので，実際には $\hat{\phi}^{(1)}_{\mathrm{G},j'}$ $(j'=1,\dots,J)$ を求めておけば $\phi^{(j)}_{j'}$ の推定量は $\hat{\phi}^{(j)}_{\mathrm{G},j'}=\hat{\phi}^{(1)}_{\mathrm{G},j'}/\hat{\phi}^{(1)}_{\mathrm{G},j}$ から求めることができる．また，この推定量は，Mantel-Haenszel 推定量の連続な関数であることから，漸近モデル I，およびIIの下で一致推定量であることも自明である．Yanagawa and Fujii (1990) は，$2\times J$ 表のオッズ比の層間均一性の検定を開発する途中で，Greenland (1989) と独立に同一の推定量を与えており，この推定量は Greenland-Yanagawa-Fujii (GYF) の一般化 Mantel-Haenszel 推定量とよばれている．

簡単のため $b^{(j)}_{j'}=\log \hat{\phi}^{(j)}_{\mathrm{MH},j'}$ とおくと，$\log \hat{\phi}^{(j)}_{\mathrm{G},j'}$ の漸近分散は，

$$\begin{aligned}\mathrm{var}^A(\log \hat{\phi}^{(j)}_{\mathrm{G},j'})=\frac{1}{J^2}\Bigg\{&\sum_{b=1}^{J}[\mathrm{var}^A(b^{(s)}_{j'})+\mathrm{var}^A(b^{(s)}_{j})-2\,\mathrm{cov}^A(b^{(s)}_{j'},\ b^{(s)}_{j})]\\&+\sum_{s\neq t}^{J}[\mathrm{cov}^A(b^{(s)}_{j'},\ b^{(t)}_{j'})-2\,\mathrm{cov}^A(b^{(s)}_{j'},\ b^{(t)}_{j})+\mathrm{cov}^A(b^{(s)}_{j},\ b^{(t)}_{j})]\Bigg\},\end{aligned}$$

で与えられる．いま，$\mathrm{var}^A(b^{(i)}_h)$ の推定量を U_{hii}，$\mathrm{cov}^A(b^{(i)}_h,\ b^{(j)}_h)$ の推定量を U_{hij} で表す．Greenland (1989) は U_{hii}，U_{hij} としてつぎの統計量，

$$U_{hii}=\frac{1}{2}\sum_k\left(\frac{r_{hik}}{R_{hi}}+\frac{r_{ihk}}{R_{ih}}\right)\left(\frac{u_{hik}}{R_{hi}}+\frac{u_{ihk}}{R_{ih}}\right),$$

$$U_{hij}=\frac{1}{3}\sum_k\frac{1}{N_k^2}\left(\frac{x_{hk}y_{ik}y_{jk}}{R_{hi}R_{hj}}+\frac{t_{hk}x_{jk}y_{ik}}{R_{hi}R_{jh}}+\frac{t_{hk}x_{ik}y_{jk}}{R_{ih}R_{hj}}+\frac{x_{ik}x_{jk}y_{hk}}{R_{ih}R_{jh}}\right),$$

ただし，

$$r_{stk}=\frac{x_{sk}y_{tk}}{N_k},\quad u_{stk}=\frac{x_{sk}+y_{tk}}{N_k},\quad R_{st}=\sum_k\frac{x_{sk}y_{tk}}{N_k},$$

を提案し，上式の右辺を対応する U_{hij} で置き換えて得られる分散の推定量が，漸近モデル I，およびIIの下で，$\log \hat{\phi}^{(j)}_{\mathrm{G},j'}$ の分散の一致推定量であることを示した．($s,\ t,\ i,\ j$ がすべて異なるときは $\mathrm{cov}^A(b^{(i)}_s,\ b_t^{(j)})$ の推定値はゼロとする.)

5.3 一般化射影推定量

Yanagawa and Fujii (1995) は $\theta_j=\log \phi^{(1)}_j$ とおき，

$$Q=\sum_{j<j'}^{J}W_{jj'}[b^{(j)}_{j'}-(\theta_{j'}-\theta_j)]^2,$$

を最小にする $\theta_j=\hat{\theta}_j$ を用いて，

$$\hat{\psi}^{(j)}_{\mathrm{YF},j'}=\exp(\hat{\theta}_{j'}-\hat{\theta}_j),$$

で $\psi^{(j')}_j$ の推定を行うことを提案した．ただし $W_{jj'}$ は $W_{jj'}=W_{j'j}$ を満たす $\{\theta_j\}$ に依存しない所与のウエイトである．この推定量を一般化 Mantel-Haenszel 射影推定量という．いま，$\boldsymbol{\theta}=(\theta_1,\ldots,\theta_J)$ とおき，$\boldsymbol{A}=(a_{st})$ を $(s,\ t)$ 要素が，

$$a_{st}=\begin{cases}\sum_j W_{s+1,j}-W_{s+1,s+1}, & s=t \quad (s=1,\ldots,J-1)\\ -W_{s+1,t+1}, & s\neq t \quad (s,\ t=1,\ldots,J-1),\end{cases}$$

の $(J-1)\times(J-1)$ 行列，$\boldsymbol{T}$ をその第 s 要素が

$$T_s=\textstyle\sum_j W_{j,s+1}b^{(j)}_{s+1},$$

で与えられる $J-1$ 次元ベクトルとするとき，一般化 Mantel-Haenszel 射影推定量は具体的に，

$$\hat{\boldsymbol{\theta}}=\boldsymbol{A}^{-1}\boldsymbol{T},$$

で与えられ，特に，すべての $j<j'$ に対して $W_{jj'}=1$ のときは GYF 推定量と一致する．Yanagawa and Fujii (1995) は，$W_{jj'}=(R_{jj'}R_{j'j})/(R_{jj'}+R_{j'j})$ とした場合の一般化射影推定量の分散は GYF 推定量の分散より若干小さいことを示している．また，要因Xに割り付けられたすべての層に共通なスコア $C_j\ (j=1,\ldots,J)$ に対して $\theta_j=\beta C_j$ と表される場合，β の一般化 Mantel-Haenszel 射影推定量，

$$\hat{\beta}=\frac{1}{\sum_{j<j'}W_{jj'}(C_j-C_{j'})^2}\sum_{j<j'}W_{jj'}b^{(j')}_j(C_j-C_{j'}), \tag{5.4}$$

も与えている．

一般化 Mantel-Haenszel 射影推定量 $\hat{\boldsymbol{\theta}}$ の分散共分散行列は，

$$\mathrm{var}^A(\hat{\boldsymbol{\theta}})=\boldsymbol{A}^{-1}\,\mathrm{var}^A(\boldsymbol{T})\boldsymbol{A}^{-1},$$

で与えられ，さらに $\mathrm{var}^A(\boldsymbol{T})$ の (s,t) 要素 σ_{st} は，

$$\sigma_{st}=\begin{cases}\sum_j W^2_{j,s+1}\,\mathrm{var}^A(b^{(j)}_{s+1})+2\sum^{J}_{j<j'}W_{j,s+1}\ W_{j',s+1}\,\mathrm{cov}^A(b^{(j)}_{s+1},\ b^{(j')}_{s+1}), & s=t\\ \sum_j W_{j,s+1}W_{j,t+1}\,\mathrm{cov}^A(b^{(j)}_{s+1},\ b^{(j)}_{t+1}), & s\neq t\end{cases}$$

で与えられる．このことから $\hat{\boldsymbol{\theta}}$ の分散共分散行列の推定量は，上述の Greenland の場合と同様に上式の右辺を対応する U_{hij} で置き換えることによって得られる．Yanagawa and Fujii (1995) は，U_{hii} は Greenland と同一であるが，U_{hij} については Greenland より簡略化されたつぎの U^*_{hij} を用いることを提案している．

$$U^*_{hij}=\frac{\sum_k r_{ijk}t_{hk}}{2R_{ih}R_{hj}}+\frac{\sum_k r_{jik}t_{hk}}{2R_{hi}R_{jh}}.$$

これから得られる分散の推定量もまた，漸近モデル I，および II の下で，$\hat{\boldsymbol{\theta}}$ の分散の一致推定量である．

5.4　Mantel 検定

最後に Mantel-Haenszel 検定の一般化について述べる．要因Xに，第 k 層で，適当なスコア $C_{jk}\ (j=1,\ldots,J)$ が割り付けられているとする．このとき Mantel (1963) は，要因Yと要因Xの関連性を検定するために Mantel-Haenszel 検定を次のように一般化した．

$$X_{\mathrm{M}}^2=\frac{(\sum_k\sum_j C_{jk}x_{jk}-e)^2}{V},$$

ここに，

$$e=\sum_k\sum_j C_{jk}\frac{t_{jk}}{N_k},$$

$$V=\sum_k\frac{n_k m_k}{N_k^2(N_k-1)}[N_k\sum_j C_{jk}^2 t_{jk}-(\sum_j C_{jk}t_{jk})^2],$$

である．統計量 X_{M}^2 は，近似的に自由度1のカイ二乗分布に従う．この統計量に基づく検定をMantel検定，あるいは拡張Mantel-Haenszel検定という．柳川 (1986) は，$\psi_j^{(1)}$ に対応する第 k 層のオッズ比 $\psi_{jk}^{(1)}$ を，

$$\log\psi_{jk}^{(1)}=\beta(C_{jk}-C_{1k}),$$

と表し，帰無仮説 H_0：$\beta=0$ の対立仮説 H_1：$\beta\neq0$ に対する漸近的一様最強力相似検定を導いて，これがMantel検定と一致することを示した．この定式化から明らかなように，Mantel検定は層別されたデータに関する各層の用量・反応関係の要約検定である．スコア C_{jk} としては，対象とされている要因 X に応じて $C_{jk}=j$ としたり，レベル $X=j$ の代表値を割り付けたり，あるいは以下のようなウィルコクスン・スコア，

$$C_{1k}=\frac{t_{1k}+1}{2},\ C_{jk}=\sum_{s=1}^{j-1}t_{sk}+\frac{t_{jk}+1}{2}\quad(j\geq2),$$

が割り付けられる．

6. 推定方程式と Mantel-Haenszel の方法

Mantel-Haenszelの方法が扱う層別解析のモデルでは共通オッズ比 ψ 以外に，K 個の局外母数があるので（無条件の）最尤法の性質は良くない．推定方程式はこの最尤法の欠点を取り除くために提案された．推定方程式の方法は，一般化線形モデル(McCullagh and Nelder (1989))における擬似尤度推定法と，不偏な推定量方程式から良い推定方式を求めようとする推定方程式の理論（Godambe (1991)）に大別される．

第 k 層のみのデータを考えた場合，オッズ比は推定方程式

$$g_k(\psi\,;\,x_k,\,n_k,\,y_k,\,m_k)=x_k(m_k-y_k)-\psi y_k(n_k-x_k)=0,$$

の解として得られる．$E[g_k(\psi\,;\,x_k,\,n_k,\,y_k,\,m_k)]=E[g_k(\psi\,;\,x_k,\,n_k,\,y_k,\,m_k)|t_k,\,\psi]=0$ であるから，推定方程式 g_k は不偏である．共通オッズ比を推定する問題は K 個の層から得られる推定方程式を結合することになるから，重み ω_k を用いて次の推定方程式，

$$\sum_k\omega_k g_k(\psi\,;\,x_k,\,n_k,\,y_k,\,m_k)=0, \tag{6.1}$$

を考える．この方程式ももちろん不偏である．プロファイルスコア方程式(2.4)および条件付最尤推定量のための方程式(2.5)に比べてシンプルであることから，その結果得られる推定量が頑健になることが期待できる．よい重み ω_k を論じるために，一般の推定関数を $l(\theta,\{x_k\})=\sum_k\omega_k l_k(\theta,x_k)$ と書こう．Godambe (1960) は感受性の規準 $E^2[l_\theta(\theta,\{x_k\})]/E[l^2(\theta,\{x_k\})]$ を最小にする重みを用いることを提案した．ただし，$l_\theta(\theta,\{x_k\})=dl(\theta,\{x_k\})/d\theta$ である．ところが実際にこの規準で求めた最適な重みは局外母数にも依存してしまう．そこで $\psi=1$ とおいた局所的に最適な重みを求めると，局外母数には依存しないで $\omega_k=1/N_k$ になる．この局所最適な重みを(6.1)式に代入すると，Mantel-Haenszel推定量が導かれる（Yanagimoto (1990)）．

2.1 節では，漸近モデル I のもとで Mantel-Haenszel 推定量の重みを特徴づけたが，「Mantel-Haenszel 推定方程式」を考えることで漸近モデルとは無関係に，帰無仮説 H_0：$\psi=1$ のもとでの局所最適な重みとして解釈できる．局所的に最適な重みを適用した例として，x_k(および y_k) が独立な事象の集まりではなく相関がある場合 (Liang (1985))，Hardy-Weinberg 法則のモデル (Olson (1993)) などがある．

推定方程式は，推定量を求めるだけでなく，信頼区間あるいは検定量を構成するためにも用いられる．このためには推定関数がスコア関数と似ていることが望ましい．スコア関数の良い性質として不偏性と情報不偏性があり，推定方程式も $E[l(\theta,\{x_k\})]=0$(不偏性)，$E[l^2(\theta;\{x_k\})]=-E[l_\theta(\theta;\{x_k\})]$ (情報不偏性) であることが望ましい．方程式(6.1)は $\omega_k=1/N_k$ とおくと，$\psi=1$ で情報不偏になる．また，$dg_k(\psi;x_k,n_k,y_k,m_k)/d\psi\geq 0$ であるから，推定方程式(6.1)の左辺は擬似スコア関数と見なされる．その結果，3.2 節での信頼区間(3.18)も擬似スコア関数にもとづく方法と解釈できる．

Liang は，Mantel-Haenszel 推定方程式を二つの方向に拡張した．ケース 1 名から数名に対しコントロール数名をマッチした疫学的ケース・コントロール研究の解析には，条件付きロジスティック回帰が用いられる (Breslow and Day (1980)) が，この計算は共通オッズ比の条件付き最尤推定量の計算以上にやっかいである．Liang (1987) は，ロジスティックモデルの回帰係数推定に Mantel-Haenszel 推定方程式の拡張を行った．この方法は，5 節で述べた複数の 2×J 表へ応用することができる (Sato (1991))．また，Mantel-Haenszel 推定方程式は相関のあるデータに対しても不偏であり，Mantel-Haenszel オッズ比はこの場合でも共通オッズ比の一致推定量となる (Liang (1985))．(独立性を仮定している最尤推定量は，条件付きでも無条件でも，もはや一致性を持たない．) この結果から，個人内で結果を繰り返し観察した経時観察データや，家族などを単位としたクラスターサンプリングを行った層内で相関のあるデータに対しても，Mantel-Haenszel 推定方程式を拡張することができて，その方法論は今日では GEE (Generalized Estimating Equations) としてよくしられている (Liang and Zeger (1986); Diggle et al. (1994))．

7. おわりに

Mantel-Haenszel の方法は ad hoc な方法として導入され，経験的にその有効性が論じられてきたが，今日では理論的にも美しい結果に基づく方法であることが分かってきた．その結果 Mantel-Haenszel の方法は，層別された離散型モデルでの共通母数の推測法の輝かしい古典として理解される．

しかし，ロジスティックモデルなどの回帰分析がルーチンのデータ解析として主流になっている今日でも，Mantel-Haenszel の方法が時代遅れの古典になってしまったわけではない．一般にはあまり認識されていないが，ロジスティックモデルには次に示すいくつかの制約がある (Greenland (1991))．

1）交絡要因のレベルが異なっても，曝露の効果は等しい
2）曝露や交絡要因の効果のオッズは掛け算で影響する
3）曝露効果のオッズは指数関数的に増加する
4）曝露以外の交絡要因のオッズも指数関数的に増加する

これらの制約は通常データからは証明できないので，ロジスティックモデルを用いた解析だけから結論を導くことは好ましくなく，もっと制約の緩い解析方法により感度分析を行って，結論が食い違わないことを確認しておくべきである (Rothman (1986))．Mantel-Haenszel の方

法でも，共通オッズ比を推定しているのであるから，ロジスティックモデルと同様に上記 1)から 3)の仮定を必要とするが，交絡要因に関してパラメトリックなモデルを仮定しているわけではないので 4)は必要ない．したがって，Mantel-Haenszel の方法などの層別解析は，簡単な感度分析の手段として今後とも用いられ続けるであろう．

参考文献

Anscombe, F. J. (1956). On estimating binomial response relations, *Biometrika*, **43**, 461-464.

Birch, M. W. (1964). The detection of partial association, I. The 2×2 case, *J. Roy. Statist. Soc. Ser. B*, **26**, 313-324.

Breslow, N. E. (1976). Regression analysis of the log odds ratio: a method for retrospective studies, *Biometrics*, **32**, 409-416.

Breslow, N. E. (1981). Odds ratio estimators when the data are sparse, *Biometrika*, **68**, 73-84.

Breslow, N. E. (1996). Statistics in epidemiology: the case-control study, *J. Amer. Statist. Assoc.*, **91**, 14-28.

Breslow, N. E. and Day, N. E. (1980). *Statistical Methods in Cancer Research, Vol. 1: The Analysis of Case-Control Studies*, Oxford University Press, New York.

Breslow, N. E. and Liang, K.-Y. (1982). The variance of the Mantel-Haenszel estimator, *Biometrics*, **38**, 942-952.

Clayton, D. G. and Hills, M. (1993). *Statistical Models in Epidemiology*, Oxford University Press, Oxford.

Davis, L. J. (1985). Generalization of the Mantel-Haenszel estimator to nonconstant odds ratio, *Biometrics*, **41**, 487-495.

Diggle, P. J., Liang, K.-Y. and Zeger, S. L. (1994). *Analysis of Longitudinal Data*, Oxford University Press, New York.

Donner, A. and Hauck, W. W. (1986). The large-sample relative efficiency of the Mantel-Haenszel estimator in the fixed strata case, *Biometrics*, **42**, 537-545.

Emerson, J. D. (1994). Combining estimates of the odds ratio: the state of the art, *Statistical Methods in Medical Research*, **3**, 157-178.

Flanders, D. W. (1985). A new variance estimator for the Mantel-Haenszel odds ratio, *Biometrics*, **41**, 637-642.

Gart, J. J. (1966). Alternative analysis of contingency tables, *J. Roy. Statist. Soc. Ser. B*, **28**, 164-179.

Gart, J. J. (1970). Point and interval estimation of the common odds ratio in the combination of 2×2 tables with fixed margins, *Biometrikà*, **57**, 471-475.

Gart, J. J. (1971). The comparison of proportions: a review of significance tests, confidence intervals and adjustments for stratification, *Review of International Statistical Institute*, **39**, 148-169.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation, *Ann. Math. Statist.*, **31**, 1208-1211.

Godambe, V. P. (ed.) (1991). *Estimating Functions*, Oxford University Press, New York.

Greenland, S. (1984). A counterexample to the test-based principle of setting confidence limits, *American Journal of Epidemiology*, **120**, 4-7.

Greenland, S. (1989). Generalized Mantel-Haenszel estimators for $K\,2\times J$ tables, *Biometrics*, **45**, 183-191.

Greenland, S. (1991). *Validity Concepts in Epidemiologic Research*, UCLA School of Public Health (unpublished course text).

Greenland, S. and Salvan, A. (1990). Bias in the one-step method for pooling study results, *Statistics in Medicine*, **9**, 247-252.

Haldane, J. B. S. (1956). The estimation and significance of the logarithm of a ratio of frequencies, *Annals of Human Genetics*, **20**, 309-311.

Halperin, M. (1977). Re: "Estimability and estimation in case-referent studies", Letter to the Editor, *American Journal of Epidemiology*, **105**, 496-498.

Hauck, W. W. (1979). The large sample variance of the Mantel-Haenszel estimator of a common odds ratio, *Biometrics*, **35**, 817-819.

Hauck, W. W. and Donner, A. (1988). The asymptotic relative efficiency of the Mantel-Haenszel estimator in the increasing-number-of-strata case, *Biometrics*, **44**, 379-384.

Lancaster, H. O. (1961). Significance tests in discrete distributions, *J. Amer. Statist. Assoc.*, **56**, 223-234.

Liang, K.-Y. (1985). Odds ratio inference with dependent data, *Biometrika*, **72**, 678-682.

Liang, K.-Y. (1987). Extended Mantel-Haenszel estimating procedure for multivariate logistic regression models, *Biometrics*, **43**, 289-299.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models, *Biometrika*, **73**, 13-22.

Mantel, N. (1963). Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure, *J. Amer. Statist. Assoc.*, **58**, 690-700.

Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease, *Journal of the National Cancer Institute*, **22**, 719-748.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed., Chapman and Hall, London.

Mehta, C. R., Patel, N. R. and Gray, R. (1985). Computing an exact confidence interval for the common odds ratio in several 2×2 contingency tables, *J. Amer. Statist. Assoc.*, **80**, 969-973.

Mickey, R. M. and Elashoff, R. M. (1985). A generalization of the Mantel-Haenszel estimator of partial association for $2\times J\times K$ tables, *Biometrics*, **41**, 623-635.

Miettinen, O. S. (1976). Estimability and estimation in case-referent studies, *American Journal of Epidemiology*, **103**, 498-502.

Olson, J. M. (1993). Testing the Hardy-Weinberg low across strata, *Annals of Human Genetics*, **57**, 219-295.

Phillips, A. and Holland, P. W. (1987). Estimators of the variance of the Mantel-Haenszel log-odds-ratio estimate, *Biometrics*, **43**, 425-431.

Robins, J. M., Breslow, N. E. and Greenland, S. (1986). Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models, *Biometrics*, **42**, 311-323.

Rothman, K. J. (1986). *Modern Epidemiology*, Little-Brown, Boston.

Sato, T. (1990). Confidence limits for the common odds ratio based on the asymptotic distribution of the Mantel-Haenszel estimator, *Biometrics*, **46**, 71-80.

Sato, T. (1991). An estimating equation approach for the analysis of case-control studies with exposure measured at several levels, *Statistics in Medicine*, **10**, 1037-1042.

佐藤俊哉 (1993). 疫学研究における生物統計手法, 日本統計学会誌, **22**, 増刊号, 493-513.

Schlesselman, J. J. (1982). *Case-Control Study : Design, Conduct, Analysis*, Oxford University Press, New York.

Takagi, H. (1990). A simple recursive algorithm for the exact confidence limits for the common odds ratio in a series of 2×2 tables, Research Memo., No. 385, The Institute of Statistical Mathematics, Tokyo.

Takagi, H. (1991). Approximate and exact methods for computing the confidence limits for the common odds ratio in a series of 2×2 tables, *Bulletin of the Biometric Society of Japan*, **12**, 55-65.

Tarone, R. E. (1981). On summary estimators of relative risk, *Journal of Chronic Diseases*, **34**, 463-468.

Thomas, D. G. (1975). Exact and asymptotic methods for the combination of 2×2 tables, *Computers and Biomedical Research*, 8, 423-446.

Ury, H. K. (1982). Hauck's approximate large-sample variance of the Mantel-Haenszel estimator, *Biometrics*, **38**, 1094-1095.

Vollset, S. E. and Hirji, K. F.(1991). A microcomputer program for exact and asymptotic analysis of several 2×2 tables, *Epidemiology*, **2**, 217-220.

Vollset, S. E., Hirji, K. F. and Elashoff, R. M. (1991). Fast computation of exact confidence limits for the common odds ratio in a series of 2×2 tables, *J. Amer. Statist. Assoc.*, **86**, 404-409.

Woolf, B. (1955). On estimating the relation between blood groups and disease, *Annals of Human Genetics*, **19**, 251-253.

柳川 堯 (1986). 『離散多変量データの解析』, 共立出版, 東京.

Yanagawa, T. and Fujii, Y. (1990). Homogeneity test with a generalized Mantel-Haenszel estimator for $L\,2\times J$ tables, *J. Amer. Statist. Assoc.*, **85**, 744-748.

Yanagawa, T. and Fujii, Y. (1995). Projection-method Mantel-Haenszel estimator for $K\,2\times J$ tables, *J.*

Amer. Statist. Assoc., **90**, 649-656.

Yanagimoto, T. (1990). Combining moment estimators of a parameter common through strata, *J. Statist. Plann. Inference*, **25**, 187-198.

Yanagimoto, T. (1992). The Mantel-Haenszel statistics for the extended odds ratio in the negative binomial distribution, *J. Japan Statist. Soc.*, **22**, 7-17.

Yusuf, S., Peto, R., Lewis, J., Collins, R. and Sleight, P. (1985). Beta blockade during and after myocardial infarction: an overview of the randomized trials, *Progress in Cardiovascular Diseases*, **27**, 335-371.

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

The Mantel-Haenszel Method for Stratified Analysis of 2×2 Tables

Tosiya Sato and Hirofumi Takagi
(The Institute of Statistical Mathematics)

Takashi Yanagawa
(Department of Mathematics, Kyushu University)

Takemi Yanagimoto
(The Institute of Statistical Mathematics)

Several methods have been proposed for estimating a common odds ratio in stratified 2×2 tables, such as the weighted least squares, the Mantel-Haenszel, the unconditional maximum likelihood, the conditional maximum likelihood, and the Peto one-step methods. In this paper, we review the development and history of the Mantel-Haenszel method of point and interval estimations of the common odds ratio. Introducing two different asymptotics, i.e. sparse data and large-strata, we compare the Mantel-Haenszel method with other methods including the exact method. The extension of the Mantel-Haenszel method to stratified $2 \times J$ tables cases and the relationship between the Mantel-Haenszel method and the estimating equation approach are discussed.

Key words: Common odds ratio, estimating function, Mantel-Haenszel method, stratified analysis.