

日 本 計 量 生 物 学 会

1 9 9 1 年 度 年 会

1 9 9 1 年 4 月 2 0 日 (土)

於 丸の内センタービル

層別四分表の共通オッズ比の信頼区間推定の 近似的方法と正確な方法の比較

高木 廣文 (統計数理研究所)

Approximate and Exact Methods for Computing the Confidence Limits
for the Common Odds Ratio in a Series of 2×2 Tables

Hirofumi Takagi

The Institute of Statistical Mathematics

[近似的な方法と正確な方法]

K 個の四分表があり, 第 k 層の患者数と対照者数を n_k, m_k とし, 要因をもつ例数をそれぞれ $x_k, y_k, t_k = x_k + y_k$ としよう. また, $N_k = n_k + m_k, R_k = x_k(m_k - y_k)/N_k, S_k = y_k(n_k - x_k)/N_k$ とし, $R_+ = \sum_k R_k, S_+ = \sum_k S_k$ とする. 共通オッズ比 ψ の Mantel-Haenszel(1959) 推定量は, $\hat{\psi}_{MH} = R_+/S_+$ となる. 共通オッズ比の $100(1 - \alpha)\%$ 信頼区間を近似的に求める方法は, 大きく 2 通りに分けて考えられよう. 一つは, 対数オッズ比 $\ln \psi$ を考え, その漸近分散を用いて信頼区間を求める方法であり, ここでは「対数法」と呼ぼう. Robins ら (1986) は $\ln \hat{\psi}_{MH}$ の分散として対称性のある (四分表の行, 列の入れ替え可能) 推定量,

$$V_{RBG} = \frac{1}{2} \sum_k \left[\frac{P_k R_k}{R_+^2} + \frac{Q_k R_k + P_k S_k}{R_+ S_+} + \frac{Q_k S_k}{S_+^2} \right]$$

を提唱した. ここで, $P_k = (x_k + m_k - y_k)/N_k, Q_k = (y_k + n_k - x_k)/N_k$ である. 近似的な $100(1 - \alpha)\%$ 信頼区間は, $\hat{\psi}_{MH} \exp[\pm \chi_\alpha \cdot \sqrt{V_{RBG}}]$ により求められる. ここで, χ_α^2 は自由度 1 のカイ 2 乗分布の上側 α 点である. もう一つは, 推定関数 $R_k - \psi S_k$ を考え, その期待値と分散から,

$$\frac{(|R_+ - \psi S_+| - c)^2}{\sum_k \text{var}(R_k - \psi S_k)} = \chi_\alpha^2$$

を解き, 信頼区間 (ψ_L, ψ_U) を推定する方法である. ここで, 連続修正を行なう場合, $c = (1 + \psi)/4$ とし, そうでなければ $c = 0$ とする. Sato(1990) は,

$$\text{var}_S(R_k - \psi S_k) = \psi[(Q_k + N_k^{-1})R_k + (P_k + N_k^{-1})S_k]$$

を提唱した. この他にも各種の方法が提唱されているが, 今回はこの 2 つの手法を取り上げることにする.

一方, 条件付分布である非心超幾何分布から推定を行なう, いわゆる「正確な」方法がある (Gart, 1970 and Zelen, 1971). $t_k = x_k + y_k$ が与えられているものとする. 共通オッズ比の条件付最尤推定量 $\hat{\psi}_c$ は, $\sum_k x_k = \sum_k g_k(x_k | t_k, \hat{\psi}_c)$ を満たすように定めればよい (Gart, 1970). ここで, $g_k(x_k | t_k, \hat{\psi}_c) = \sum_{u_k} u_k W_{ku_k} \hat{\psi}_c^{u_k} / \sum_{u_k} W_{ku_k} \hat{\psi}_c^{u_k}$, であり,

$$W_{ku_k} = \binom{n_k}{r_k} \binom{m_k}{t_k - r_k}, \quad L_k = \max(0, t_k - m_k) \leq u_k \leq \min(n_k, t_k) = U_k,$$

である. また, $\ln \hat{\psi}_c$ の分散の推定値は, $V_c = [\sum_k \{g_k(x_k^2 | t_k, \hat{\psi}_c) - g_k^2(x_k | t_k, \hat{\psi}_c)\}]^{-1}$ により与えられる. これらの推定値から対数法により, 近似的な信頼区間を求めることができる.

正確な $100(1-\alpha)\%$ 信頼区間 $(\hat{\psi}_L, \hat{\psi}_U)$ は, $s = \sum_k x_k$, $L = \sum_k L_k$, $U = \sum_k U_k$ とすると,

$$\sum_{z=s}^U G(z|\hat{\psi}_L) = \alpha/2, \text{ および } \sum_{z=L}^s G(z|\hat{\psi}_U) = \alpha/2,$$

を満足するように定める. ここで,

$$G(z|\hat{\psi}) = C_z \hat{\psi}^z / \sum_{r=L}^U C_r \hat{\psi}^r, \quad C_z = \sum_{R(z)} \prod_{k=1}^K W_{kr_k}$$

であり (Zelen, 1971), $R(z)$ は $\sum_k r_k = z$ となる K 個の四分表の組合せの全集合である. ただし, $s = L$ の場合, $\hat{\psi}_L = 0$ とし, $s = U$ の場合, $\hat{\psi}_U = \infty$ とする. 実際の信頼区間の計算には, Mehta ら (1985) や高木 (1990) のアルゴリズムを用いればよい.

[シミュレーションによる比較]

共通オッズ比の信頼区間を求めるための近似的な方法と正確な方法を比較するために, シミュレーションを行なった.

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, inbalanced design の場合の 3 通りについて, 患者数 n と対照者数 m および層の個数 K を数通りに変えて行なった. 共通オッズ比は, 1, 3.5, 6.5 の 3 通りとした. 患者群での要因曝露率を, 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. これらの組合せについて, 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. なお, 計算は統計数理研究所の HITAC-M682H 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.

シミュレーションの結果, 対数法を用いる場合には, 分散として V_{RBG} か V_c の違いによる差はほとんどなかった. 正確な方法は, 最も保守的な結果を与え, Sato の方法で連続修正しない場合が, 最も名義水準に近い結果を与えた. 連続修正を行なった場合には, Sato の方法と正確な方法はほぼ一致した結果を示した.

[参考文献]

- Gart, J.J. (1970). Point and interval estimation of the common odds ratio in the combination of 2×2 tables with fixed marginals. *Biometrika* 57, 471-475.
- 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.
- 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. *Journal of the American Statistical Association* 80, 969-973.
- 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 model. *Biometrics* 42, 311-323.
- Sato, T. (1990). Confidence limits for the common odds ratio based on the asymptotic distribution of the Mantel-Haenszel estimator. *Biometrics* 46, 71-80.
- 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.
- Zelen, M. (1971). The analysis of several 2×2 contingency tables. *Biometrika* 58, 129-137.