

Vol. 37 No. 10

特別附録 平成2年10月15日発行

October 1990

JAPANESE JOURNAL OF PUBLIC HEALTH

日本公衆衛生雑誌

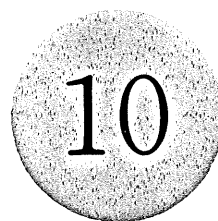
第37巻・第10号 特別附録 平成2年10月

第49回日本公衆衛生学会総会抄録集 III

徳 島

日本公衛誌
J. J. P. H

日 本 公 衆 衛 生 学 会
Japanese Society of Public Health



特別附録

層別四分表の共通オッズ比の正確な推定値と信頼区間について

たかぎ ひろふみ

○高木 廣文 柳本 武美 (統計数理研究所)

患者-対照研究等で層別四分表の共通オッズ比の推定値及び信頼区間を求める場合、条件付分布に基づく「正確な」方法がある。これまで計算上の困難が強調されてきたが、計算アルゴリズムの開発により、実用的な手法となった。Mantel-Haenszel 推定量に基づく方法に関しては、各近似的方法の結果も良いけれども、必ずしも満足すべきものではない。共通オッズ比に関する推測は、基本的には条件付分布に基づいて行なうべきである。

[近似的な方法と正確な方法] 患者-対照研究では、ある疾患に対する特定の要因の有無による影響を検討する場合、交絡因子の影響を除くために標本を層別し、各四分表に共通するオッズ比を求めることがよく行なわれる。いま K 個の四分表があり、その第 k 層の患者数と対照数を n_k, m_k 、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 推定量 (MHE) は、 $\hat{\psi}_{MH} = R_+/S_+$ となる (Mantel and Haenszel, 1959)。共通オッズ比の近似的な $100(1-\alpha)\%$ 信頼区間を構成する方法は、大きく 2 通りに分けて考えられる。一つは、オッズ比の分布は右に歪んでいるので、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。 $\ln \hat{\psi}_{MH}$ 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、 $V_{RBG} = \frac{1}{2} \sum_k \left[\frac{P_k R_k}{R_k^2} + \frac{Q_k R_k + P_k S_k}{R_k S_+} + \frac{Q_k S_k}{S_+^2} \right]$ を Robins ら (1986) は提唱している。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、 χ^2_α を自由度 1 のカイ 2 乗分布の上側 α 点とすると、 $\hat{\psi}_{MH} \exp[\pm \chi^2_\alpha \cdot \sqrt{V_{RBG}}]$ により求められる。もう一つは、検定統計量を基にする方法であり、 $\frac{(|R_+ - \psi S_+| - c)^2}{\sum_k \text{var}(R_k - \psi S_k)} = \chi^2_\alpha$ を解き、信頼区間 (ψ_L, ψ_U) を推定する。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、Yanagimoto (1990) は、 $\text{var}_Y(R_k - \psi S_k) = [\psi n_k m_k t_k (N_k - t_k)/N_k^2 - (\psi - 1)^2 R_k S_k]/(N_k - 1)$ を提唱している。

一方、条件付分布である非超幾何分布から推定を行なう「正確な」方法がある (Gart, 1970 and Zelen, 1971)。いま、 $t_k = x_k + y_k$ が与えられているものとする。共通オッズ比の条件付最大推定量 $\hat{\psi}_c$ 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、 $g_k(x_k | t_k, \hat{\psi}_c) = \sum_{i=0}^{t_k} u_i W_{k,i} \hat{\psi}_c^i / \sum_{i=0}^{t_k} W_{k,i} \hat{\psi}_c^i$ であり、 $W_{k,i} = \binom{n_k}{i} \binom{m_k}{t_k - i} L_k = \max(0, i_k - m_k) \leq u_k \leq \min(n_k, i_k) = U_k$ である。また、 $\ln \hat{\psi}_c$ の分散の推定値は、 $V_c = [\sum_k \{g_k(x_k | t_k, \hat{\psi}_c) - g_k^2(x_k | t_k, \hat{\psi}_c)\}]^{-1}$ により与えられる。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、 $\sum_{i=L}^s G(z | \hat{\psi}_L) = \alpha/2, \sum_{i=L}^s G(z | \hat{\psi}_U) = \alpha/2$ を満足するように定める。ここで、 $G(z | \hat{\psi}) = C_s \hat{\psi}^z / \sum_{i=L}^U C_i \hat{\psi}^i, C_s = \sum_{i=L}^U \prod_{k=1}^K W_{k,i}$ であり (Zelen, 1971)、 $R(z) = \sum_k r_k = z$ 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。従来、条件付分布に基づく推定は標本数が小さい場合にしか実用的ではなかったが、現在では比較的大標本でも計算が可能であり (Satten and Kupper, 1990, and Takagi, 1990)、信頼区間も Mehta ら (1985) や高木 (1990) のアルゴリズムを用いれば計算は容易である。

[応用例による比較] 近似的な方法を正確な方法と比較するために、オッズ比が広範な範囲に渡るように Robins ら (1986) や Sato (1990) が用いた 4 例と、柳川 (1986) の表 3.10 (p.76) のデータを用いて解析を行い、表に示した。対数法は、matched data に関しては、良い下限を与えるが、上限は良いとは言えず、正確な信頼区間にくらべて狭い範囲を与える傾向にある。検定統計量による方法は、連続修正を行なった場合の方が概して良好な結果を与えるが、近似は不十分である。今後、各種の場合を想定してシミュレーションを行い、より詳細な比較を行なう予定である。

表 正確な方法と近似的方法による共通オッズ比の推定値と 95% 信頼区間

資料	Leisure World ¹⁾ matched-pairs K=63, N=126		Leisure World ²⁾ 1:4 matching K=63, N=315		Ille-et Vilaine ³⁾ K=6, N=975		Avadex ⁴⁾ K=4, N=317		肺癌と造船 ⁵⁾ K=6, N=304	
	$\hat{\psi}_{MH}$	$\hat{\psi}_c$	$\hat{\psi}_{MH}$	$\hat{\psi}_c$	$\hat{\psi}_{MH}$	$\hat{\psi}_c$	$\hat{\psi}_{MH}$	$\hat{\psi}_c$	$\hat{\psi}_{MH}$	$\hat{\psi}_c$
推定値	9.67	9.67	8.46	7.95	5.16	5.25	3.08	3.05	1.61	1.61
信頼区間	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
正確な方法	3.00	49.59	3.43	21.55	3.57	7.76	1.24	7.13	.741	3.49
対数法: V_c	2.94	31.73	3.49	18.15	3.63	7.60	1.38	6.72	.791	3.27
対数法: V_{RBG}	2.94	31.73	3.41	20.99	3.56	7.47	1.39	6.82	.791	3.29
Sato: 連続修正なし	3.13	29.86	3.54	20.25	3.58	7.43	1.41	6.74	.795	3.28
Sato: 連続修正あり	2.82	39.77	3.29	23.53	3.50	7.66	1.28	7.81	.746	3.53
Yanagimoto: 連続修正なし	3.13	29.86	3.54	20.25	3.55	7.32	1.40	6.56	.798	3.21
Yanagimoto: 連続修正あり	2.82	39.77	3.29	23.53	3.46	7.53	1.28	7.55	.750	3.45

1) Breslow and Day (1980, p.167). 2) Ibid (p.174). 3) Gart (1971, Table 2). 4) 柳川 (1986, 表 3.10).

[参考文献] Breslow, N.E. and Day, N.E. (1980). *Statistical Methods in Cancer Research. Vol.1: The Analysis of Case-Control Studies*. Lyon: International Agency for Research on Cancer. /Gart, J.J. (1970). *Biometrika* 57, 471-475. /Gart, J.J. (1971). *Review of the International Statistical Institute* 39, 148-169. /Mantel, N. and Haenszel, W. (1959). *Journal of the National Cancer Institute* 22, 719-748. /Mehta, C.R., Patel, N.R., and Gray, R. (1985). *Journal of the American Statistical Association* 80, 969-973. /Robins, J.M., Breslow, N.E., and Greenland, S. (1986). *Biometrics* 42, 311-323. /Sato, T. (1990). *Biometrics* 46, 71-80. /Satten, G. A. and Kupper L.L. (1990). *Biometrics* 46, 217-223. /Takagi, H. (1990). *Computer Science Monographs*, submitted. /高木廣文 (1990). 第 18 回日本行動計量学会大会発表論文抄録集. /柳川 堯 (1986). 離散多変量データの解析, 共立出版. /Yanagimoto, T. (1990). *Journal of Statistical Planning and Inference* 25, 187-198. /Zelen, M. (1971). *Biometrika* 58, 129-137.