Ann. Inst. Statist. Math.
37 (1985), Part A, 37-49

SIMPLE LINEAR APPROXIMATIONS TO THE LIKELIHOOD EQUATION FOR COMBINING EVIDENCE IN MULTIPLE 2×2 Tables: A CRITIQUE OF CONVENTIONAL PROCEDURES

TAKEMI YANAGIMOTO AND EIJI YAMAMOTO

(Received Nov. 18, 1983; revised Jan. 31, 1984)

Summary

The conventional procedures for a common odds ratio in multiple 2×2 tables are explored and critiqued. Three types of linear approximation to the likelihood equations under some models of common measures of association are used to derive the popular conventional estimators and test statistics. Some of them are derived using the model of the common standardized difference which is an unacceptable measure. The derivation provides us with some characteristics of the procedures. The advantages of procedures based on the conditional and unconditional likelihoods are discussed.

1. Introduction

Several estimators and test procedures in multiple 2×2 tables have been devised when the possibility of a common odds ratio is assumed. Woolf [27] presented an estimator and a test procedure and Mantel-Haenszel [21] introduced a well known estimator and test procedure. Birch [3] defined another estimator and test statistic, which are close to an estimator by Yates [29] and a test statistic by Cochran [6]. These estimators and test statistics are usually referred to by the author's name. Other estimators were defined, such as the unconditional maximum likelihood (UML) estimator, after adding a constant to an entry in each cell. Gart [12] recommended a half integer 0.5 as the constant and Hitchcock [19] suggested a quarter-integer 0.25. In this decade the UML and the conditional maximum likelihood (CML) procedures have attracted the researcher's attention in the analysis of the contingency table.

Fleiss [11] presented a fine overview on this confusing problem.

Key words and phrases: Common odds ratio, conditional likelihood, half integer correction, Mantel-Haenszel estimator, quarter integer correction.

He recommended, in the earlier edition of his monograph, the Yates estimator and the extensive use of the Cochran test. But Halperin et al. [15] criticized the extensive use of the Cochran test for the homogeneity under some conditions, and he agreed with the criticism.

Many simulation studies were conducted to compare the estimators. Studies by Mckinlay [23], Farewell and Prentice [9], Lubin [20], Hauck et al. [18] and Breslow [5] compared the bias of the estimators and suggested that the CML estimator, Mantel-Haenszel and the quarter integer correction are less biased.

The main result of this paper is to show that the Yates, Mantel-Haenszel and Woolf estimators can be regarded as the root of linear equations which approximate the likelihood equation and their test statistics can be regarded as approximations of those of the conditional and unconditional likelihood procedures. This derivation follows some fundamental properties of them. We show that the Birch estimator as well as the Yates can be regarded as an estimator of the common standardized difference (Fleiss [10]) rather than that of the common log-odds ratio. This corresponds to the criticism by Halperin et al. [15]. Similar criticisms extend to the Cochran and Mantel-Haenszel test statistics.

The half- and quarter-integer correction of the UML estimator is examined in relation to the CML estimator. The quarter-integer correction provides the better approximation to the CML than the half integer. Following the above approximations, the test procedures are also discussed.

2. Models and likelihood equation

Consider a $2K$ set of two binomial variates X_k and Y_k, $k=1,\cdots,K$ with their incidence probabilities p_k and q_k, respectively. Let n_k and m_k be sample sizes from X_k and Y_k, respectively, and x_k and y_k be the numbers of individuals with positive occurence, respectively. We will write x_k+y_k as s_k and n_k+m_k as t_k. Following the usual notion the subscript '.' denotes the summation, thus $x.=\sum_{k=1}^{K} x_k$.

Consider a measure of association, R_k, between X_k and Y_k. We are often interested in a model with common R_k's, R, for $k=1,\cdots,K$, while p_k and q_k depend on k. The problem of combining the common association of multiple 2×2 tables is to estimate a common measure of association R and to test for $R=0$ against $R\neq 0$, where $R=0$ means $p_k=q_k$. Interest has been largely focused on the log-odds ratio, that is,

$$\beta_k=\log(p_k(1-q_k)/q_k(1-p_k)),$$

or, equivalently, on the odds ratio

$$\gamma_k = p_k(1-q_k)/q_k(1-p_k) .$$

Write $p_k = \exp(\alpha_k+\beta_k)/(1+\exp(\alpha_k+\beta_k))$, then $q_k = \exp(\alpha_k)/(1+\exp(\alpha_k))$. The models used in the conventional problem of combining the common effect of multiple 2×2 tables are

$$M_0 : \ \beta_k = 0 \qquad \text{for any } k$$

$$M_C : \ \beta_k = \beta \qquad \text{for any } k$$

and

$$M_S : \ \beta_k\text{'s are arbitrary} .$$

Many procedures were proposed to estimate a common measure β in the model M_C, to test for the null hypothesis M_0 against the alternative M_C and to test for M_0 against M_S.

The UML estimator under the model M_C is given as the root of the equations, if they have a root,

$$\text{(2.1)} \qquad \sum_k x_k - \sum_k n_k \exp(\alpha_k+\beta)/(1+\exp(\alpha_k+\beta)) = 0$$

$$\text{(2.2)} \qquad s_k - n_k \exp(\alpha_k+\beta)/(1+\exp(\alpha_k+\beta)) - m_k \exp(\alpha_k)/(1+\exp(\alpha_k)) = 0 \quad \text{for } k=1,\cdots,K .$$

Since the equation (2.2) has a unique root, if it has a root, we denote it by $\alpha_k(\beta)$. Put

$$\mathrm{UL}_k(\beta) = x_k - n_k \exp(\alpha_k(\beta)+\beta)/(1+\exp(\alpha_k(\beta)+\beta)) .$$

Then the UML estimator of β is the root of the equation $\sum \mathrm{UL}_k(\beta)=0$. It is easy to show that $\mathrm{UL}_k(\beta)$ is strictly decreasing in β. Following Yanagimoto-Kamakura the power expansion of $\mathrm{UL}_k(\beta)$ up to the second order is expressed as

$$\mathrm{UL}_k(\beta) \fallingdotseq x_k - [n_k s_k/t_k + \{s_k(t_k-s_k)n_k m_k/t_k^3\}\beta + 1/2\{s_k(t_k-s_k)(t_k-2s_k)n_k m_k(m_k-n_k)/t_k^5\}\beta^2] .$$

Parallel derivation is available for other measures of association. The UML estimator of the common odds ratio is the root of the equation, $\sum \mathrm{UO}_k(\gamma)=0$ with each term

$$\mathrm{UO}_k(\gamma) = x_k/\gamma - n_k \exp(\alpha_k(\gamma))/(1+\gamma\exp(\alpha_k(\gamma))) ,$$

where $\alpha_k(\gamma)$ is again the root of the equation (2.2) by replacing $\exp(\beta)$ by γ.

Next we consider the standardized difference δ_k, which is expressed by

$$\delta_k = (p_k-q_k)/((n_k p_k + m_k q_k)(t_k - n_k p_k - m_k q_k)/t_k^2) .$$

The standardized difference depends on the sample size in the kth stratum, as well as the incidence probabilities, p_k and q_k. Thus it may be a measure to be avoided, but as shown later it is a key measure for deriving the conventional estimators and test statistics. The UML estimator of the common measure, δ, is given by the root of the following equation, if a root exists:

$$\mathrm{UD}(\delta)=\sum\left\{\left(\frac{x_k}{p_k}-\frac{n_k-x_k}{1-p_k}\right)\frac{\partial}{\partial\delta}p_k\right\}=0\ ,$$

where $p_k(\delta_k, q_k)$ is a root of

$$\delta_k=(p_k(\delta_k, q_k)-q_k)/[\{n_k p_k(\delta_k, q_k)+m_k q_k\}\cdot\{n_k(1-p_k(\delta_k, q_k))+m_k(1-q_k)\}/t_k^2]\ ,$$

and $q_k(\delta)$ is a root of

$$\left(\frac{x_k}{p_k}-\frac{n_k-x_k}{1-p_k}\right)\frac{\partial}{\partial q_k}p_k+\frac{y_k}{q_k}-\frac{m_k-y_k}{1-q_k}=0\ .$$

It is shown that the roots of the above equations are unique, if they exist.

3. Linear approximation

We will show that linear approximations to the three log-likelihood functions result in the conventional estimators. For convenience we denote any one of the three functions, $\mathrm{UL}(x)$, $\mathrm{UO}(x)$ and $\mathrm{UD}(x)$ by $U(x)$, and write $U(x)=\sum U_k(x)$.

The iteration procedure is required to obtain the root of the equation $U(x)=0$, since the function $U(x)$ is nonlinear. A linear approximation permits us to get an approximated root, though the simple approximation may follow serious bias. We discuss three types of linear approximations $U(x)$, which can be used to derive the conventional estimators and test statistics.

Recall that the values of the functions and their derivatives are easily obtained. In fact it follows that

$$\mathrm{UL}_k(0)=\mathrm{UO}_k(1)=\mathrm{UD}_k(0)=x_k-n_k s_k/t_k\ .$$

Suppose that $\hat\beta_k$ is a (unique) root of the equation, $\mathrm{UL}_k(\beta)=0$, if it exists. The values of the derivatives of $\mathrm{UL}_k(\beta)$ at $\beta=0$ and $\hat\beta_k$ are: $\mathrm{UL}_k'(0)=-s_k(t_k-s_k)n_k m_k/t_k^3$ and $\mathrm{UL}_k'(\hat\beta_k)=-\{x_k(n_k-x_k)y_k(m_k-y_k)/n_k m_k\}/\{x_k(n_k-x_k)/n_k+y_k(m_k-y_k)/m_k\}$, if $\hat\beta_k$ exists.

Using these easily available values three types of linear approximations to a function $U(x)$ are introduced.

DEFINITION 1. For a function $U(x)$ three types of the linear approximation are defined.

1) (Tangent line at a null point; NL) Let x_0 be a null point. $U(x)$ is approximated by $N\circ U(x)=U(x_0)+U'(x_0)(x-x_0)$.
2) (Tangent line at a root; RL) Let x_r be a unique root of $U(x)=0$. $U(x)$ is approximated by $R\circ U(x)=U'(x_r)(x-x_r)$, if x_r exists.
3) (Connection line between a null point and a root; CL) Let x_0 and x_r be defined in 1) and 2). An approximation function $C\circ U(x)$ is defined as follows;
   i) If x_r exists and $x_r\neq x_0$, $C\circ U(x)=[(U(x_r)-U(x_0))/(x_r-x_0)](x-x_r)$ $=[-U(x_0)/(x_r-x_0)](x-x_r)$
   ii) If x_r exists and $x_r=x_0$, $C\circ U(x)=U'(x_r)(x-x_r)$
   iii) If x_r does not exist, $C\circ U(x)=U(x_0)$.

Remark 1. Though the definitions of $R\circ U(x)$ and $C\circ U(x)$ are incomplete, they are good enough for our purposes.

The conventional estimators in study are:

1) (Mantel-Haenszel estimator). The estimator is for the odds ratio γ, which is given by

$$\hat{\gamma}_{MH}=\sum\{x_k(m_k-y_k)/t_k\}/\sum\{y_k(n_k-x_k)/t_k\}\ .$$

This estimator is usually not defined, when $\sum y_k(n_k-x_k)/t_k=0$. $\mathrm{Log}\,(\hat{\gamma}_{MH})$ is regarded as an estimator of the log-odds ratio β.

2) (Woolf estimator). The estimator is for β, which is given by

$$\hat{\beta}_W=\sum\{w_k\log x_k(m_k-y_k)/y_k(n_k-x_k)\}/\sum w_k\ ,$$

where $w_k=[x_k(n_k-x_k)y_k(m_k-y_k)/n_km_k]/[x_k(n_k-x_k)/n_k+y_k(m_k-y_k)/m_k]$. This estimator is usually not defined, when it holds that $x_k(m_k-y_k)y_k\cdot(n_k-x_k)=0$ for some k.

3) (Birch estimator). The estimator is given by

$$\hat{\beta}_B=\sum\frac{m_kx_k-n_ky_k}{t_k}\Big/\sum\frac{m_kn_ks_k(t_k-s_k)}{t_k^2(t_k-1)}\ .$$

4) (Yates estimator). The estimator was originally for the standardized difference. But it is quite close to the Birch estimator, which is given by

$$\hat{\beta}_Y=\sum\frac{m_kx_k-n_ky_k}{t_k}\Big/\sum\frac{m_kn_ks_k(t_k-s_k)}{t_k^3}\ .$$

Among the above four, the Birch and Yates estimators have been given less attention. But they are closely related to the commonly used test statistics by Cochran [6] and Mantel-Haenszel [21]. The above estimators except for the Birch estimator are derived through linear approximation functions to $\mathrm{UL}_k(\beta)$, $\mathrm{UO}_k(\gamma)$ and $\mathrm{UD}_k(\delta)$, which are stated explicitly in the following propositions.

PROPOSITION 1. *Let* $C\circ\mathrm{UO}_k(\gamma)$ *be the CL approximation to* $\mathrm{UO}_k(\gamma)$. *Then the root of the equation* $C\circ\mathrm{UL}(\gamma)=C\circ\mathrm{UO}_k(\gamma)=0$ *is the Mantel-Haenszel estimator.*

PROPOSITION 2. *Let* $R\circ\mathrm{UL}_k(\beta)$ *be the RL approximation to* $\mathrm{UL}_k(\beta)$. *Then the root of the equation* $R\circ\mathrm{UL}(\beta)=R\circ\mathrm{UL}_k(\beta)=0$ *is the Woolf estimator.*

PROPOSITION 3. (i) *Let* $C\circ\mathrm{UD}_k(\delta)$ *be the CL approximation to* $\mathrm{UD}_k(\delta)$. *Then the root of the equation* $C\circ\mathrm{UD}(\delta)=0$ *is the Yates estimator.*
(ii) *Let* $N\circ\mathrm{UL}_k(\beta)$ *be the NL approximation to* $\mathrm{UL}_k(\beta)$. *Then the root of the equation* $N\circ\mathrm{UL}(\beta)=0$ *is again the Yates estimator.*

We remark that the Yates estimator can be derived in two different ways. It is not a consistent estimator of the common log-odds ratio but that of the common standardized difference, even when n_k and m_k of each stratum tend to infinity with a fixed ratio of n_k to m_k and a fixed K. The inconsistency comes from the fact that $N\circ\mathrm{UL}(\beta_0)$ is in general not equal to $\mathrm{UL}(\beta_0)$ for a true β_0 which is an important point for consistency. The CL approximation is preferable to the NL, since the approximated function by the CL approximation is quite close around both the points $\beta=0$ and β_0. The fine behavior of the Mantel-Haenszel estimator is supported partly by a characterization in Proposition 1. The Yates estimator is regarded as an estimator of δ rather than β. This, together with the sample size dependency of δ, results in the limited use of the Yates estimator. The above criticism to the Yates estimator can be extended to that to the Birch estimator and the Cochran test statistic, which will be discussed later.

4. Conditional likelihood

Consider the CML estimator, $\hat{\beta}_C$, under the model M_C, which is the root of the equation, if it exists,

$$\sum x_k-\sum\frac{\sum_u\binom{n_k}{u}\binom{m_k}{s_k-u}u\exp(\beta u)}{\sum_u\binom{n_k}{u}\binom{m_k}{s_k-u}\exp(\beta u)}=0\,.$$

Denote the left-hand side by $\mathrm{CL}(\beta)=\sum\mathrm{CL}_k(\beta)$. The conditional likelihood ratio test statistic for M_0 against M_C is given by $2\int_0^{\hat{\beta}_C} \mathrm{CL}_k(\beta)d\beta$. Andersen [1] proves that under certain regularity conditions $\hat{\beta}_C$ is a consistent estimator of β in M_C when K tends to infinity. Harkness [16] shows that $(\mathrm{CL}_k(\beta)-\mathrm{UL}_k(\beta))\beta<0$ unless either $\beta=0$ or $s_k(t_k-s_k)=0$ holds. Coupling these two results, we suspect that the UML estimator of β, $\hat{\beta}_U$, is upwardly biased and is inconsistent when K tends to infinity. The fact of upward bias is observed in existing simulation results. Thus, the CML estimator is considered to be preferable to the UML. But computation difficulties are much more serious for the CML estimator.

The Taylor expansion of $CL_k(\beta)$ follows from (5.10) in Cox [7], which is expressed as

$$(4.1)\quad \mathrm{CL}_k(\beta)=x_k-[n_k s_k/t_k+\{s_k(t_k-s_k)n_k m_k/t_k^2(t_k-1)\}\beta \\ +1/2\{s_k(t_k-s_k)(t_k-2s_k)n_k m_k(m_k-n_k)/t_k^3(t_k-1)(t_k-2)\}\beta^2] .$$

The difference between the UML and CML estimators is partly interpreted by the coefficients of the second order in the power expansion of $\mathrm{UL}_k(\beta)$ and $\mathrm{CL}_k(\beta)$, $n_k m_k s_k(t_k-s_k)/t_k^3$ and $n_k m_k s_k(t_k-s_k)/t_k^2(t_k-1)$. The ratio of the former to the latter is $(t_k-1)/t_k=a_k$. The difference reflects that between the Yates and Birch estimators.

PROPOSITION 4. (*Birch* [3]). *The Birch estimator is the root of the equation* $N\circ\mathrm{CL}(\beta)=0$.

Proposition 4 derives the inconsistency of the Birch estimator when K tends to infinity.

We consider two adjustments of $\mathrm{UL}_k(\beta)$ to $\mathrm{CL}_k(\beta)$. A simple adjustment is obtained by introducing a function $\mathrm{AL}_k(\beta)=\mathrm{UL}_k(\beta/a_k)$. $\mathrm{AL}_k(\beta)$ well approximates $\mathrm{CL}_k(\beta)$ (Breslow [4] and Yanagimoto-Kamakura [28]). We can obtain the approximated estimate, $\hat{\beta}_A$, by replacing $\mathrm{UL}_k(\beta)$ by $\mathrm{AL}_k(\beta)$. Consider a specific condition that $t_k=t$ for any k. Then the estimator β_A is $(t-1)\beta_U/t$.

Another adjustment is derived by replacing entries in cells. In place of n_k, x_k, m_k and y_k we use n_k/a_k, $x_k+n_k s_k/t_k(t_k-1)$, m_k/a_k and $y_k+m_k s_k/t_k(t_k-1)$, respectively. The formal application of the UML method to data after this replacement leads to another adjusted estimator. The Taylor expansion of the function corresponding to $\mathrm{UL}(\beta)$ is given by

$$\sum\left\{x_k-\frac{n_k s_k}{t_k}-\frac{s_k(t_k-s_k)n_k m_k}{t_k^2(t_k-1)}\beta-\frac{1}{2}\,\frac{s_k(t_k-s_k)(t_k-2s_k)n_k m_k(m_k-n_k)}{t_k^4(t_k-1)}\beta^2\right\} .$$

The idea of changing entries for bias reduction of the estimator of

β were proposed by Gart [12] and Hitchcock [19]. Gart recommended adding to each entry the half integer 0.5, which had been proposed by Haldane [14] and Anscombe [2] get a less biased estimator of the logit of a binomial distribution. Hitchcock suggested the use of the quarter-integer 0.25 rather than 0.5 based on her numerical experiments in some situations. Her suggestion was supported by Hauck et al. [18] from their simulation study. It should be remarked that our correction terms are close to Hitchcock's 0.25. In fact the total sum of the correction terms of entries in the kth 2×2 table is $t_k/(t_k-1)$, which locates between 1 and 2 but is close to 1 for a large number t_k. Especially when it holds that $n_k=m_k=s_k$, the correction terms to entries in the kth table become to be a common constant $t_k/4(t_k-1)$, which is close to 0.25.

5. Diagnosis of the linear approximation

To obtain some characteristics of the conventional estimator compared with the UML and CML estimators, we explore the behavior of the function UL(β). We assume $s_k(t_k-s_k)\neq0$ in this section. Then each term of UL(β), UL$_k(\beta)$, has the following properties.

PROPOSITION 5.
i) $\lim_{\beta\to\infty} UL_k(\beta)=x_k-\mathrm{Min}\,(n_k, s_k)$,
$\lim_{\beta\to-\infty} UL_k(\beta)=x_k-\mathrm{Max}\,(0, s_k-m_k)$.
ii) *$UL_k(\beta)$ has a unique point of reflexion.*

PROOF. Statement i) is obvious. For simplicity we will omit the subscript k. The condition UL$''(\beta)=0$ is expressed by

$$(\alpha'+1)^2(1-e^{\alpha+\beta})+\alpha''(1+e^{\alpha+\beta})=0\ , \tag{5.1}$$

where α is given by (2.2), and α' and α'' are defined by

$$ne^{\alpha+\beta}(\alpha'+1)/(1+e^{\alpha+\beta})^2+me^{\alpha}\alpha'/(1+e^{\alpha})^2=0 \tag{5.2}$$

$$\alpha'^2(1-e^{\alpha})+\alpha''(1+e^{\alpha})=0\ . \tag{5.3}$$

Eliminating α' and α'' from (5.1)-(5.3), we get $(1-2p)/n^2p^2(1-p)^2=(1-2q)/m^2q^2(1-q)^2$. Recall that (2.2) can be written by $s=np+mq$. Since $(1-2p)/n^2p^2(1-p)^2$ is strictly decreasing in p, the roots p and q are unique, if they exist. It is easy to show UL$''(\beta)=0$ has a root.

As Proposition 5 shows, the behavior of UL(β) is quite different from that of a straight line. The approximated lines do not satisfy condition (i). The linear approximation can not be expected to be well fitted for a wide range of the parameter. The lack of fit of the linear approximation is extended to that of the approximation by the polynomial function. Goodman [13] introduced an estimator based on a second order polynomial, but it is unsatisfactory as seen in Mckinlay [24].

The behavior of $\mathrm{UL}''(\beta)$ explains partly characteristics of the conventional estimators. Since $\mathrm{UL}_k''(0)=-n_k m_k s_k(t_k-s_k)(t_k-2s_k)(m_k-n_k)/t_k^5$, the point of reflexion is positive if $C_k>0$ and negative if $C_k<0$, where $C_k=(t_k-2s_k)(n_k-m_k)$. Suppose for simplicity that $C_k=0$, that is, $t_k=2s_k$ or $n_k=m_k$ for $k=1,\cdots,K$. Then $\mathrm{UL}_k(\beta)$ is concave for $\beta<0$ and convex for $\beta>0$. This implies that $\beta(N\circ\mathrm{UL}_k(\beta)-\mathrm{UL}_k(\beta))<0$ unless $\beta=0$. Summing up over k it follows that $\beta(N\circ\mathrm{UL}(\beta)-\mathrm{UL}(\beta))<0$. This implies that $\mathrm{sgn}(\hat\beta_Y)=\mathrm{sgn}(\hat\beta_U)$ and $|\hat\beta_Y|\leqq|\hat\beta_U|$, where the equality holds only when $\hat\beta_Y=0$.

Next we assume in addition that $x_k y_k(n_k-x_k)(m_k-y_k)\neq 0$ and $x_k\geqq n_k s_k/t_k$ for any k. It follows that $R\circ\mathrm{UL}_k(\beta)<\mathrm{UL}_k(\beta)$ for $\beta>0$ and $\beta\neq\hat\beta_k$, which derives that $0<\hat\beta_W<\hat\beta_U$ unless $\hat\beta_W=0$ or $\hat\beta_k=\hat\beta_U$ for any k. The behavior of $\hat\beta_W$ is not clear when we delete the additional latter assumption on x_k. The assumption $n_k=m_k$ is of practical importance and is usualy employed in existing simulation studies.

6. Test statistic

The unconditional likelihood ratio test statistics for M_0 against M_C is derived from

$$\chi_U^2=2\int_0^{\hat\beta_U}\mathrm{UL}(\beta)d\beta=2\int_0^{\hat\gamma_U}\mathrm{UO}(\gamma)d\gamma\ .$$

Using linear approximation functions in place of $\mathrm{UL}(\beta)$ or $\mathrm{UO}(\gamma)$, we obtain a series of test statistics. The use of $R\circ\mathrm{UL}(\beta)$ leads to

$$\chi_W^2=\hat\beta_W\cdot R\circ\mathrm{UL}(0)=\frac{\{\sum w_k\log x_k(m_k-y_k)/y_k(n_k-x_k)\}^2}{\sum w_k}\ .$$

The use of $N\circ\mathrm{UL}(\beta)$ and $C\circ\mathrm{UD}(\delta)$ leads to the same test statistics because of Proposition 3 and (3.1), which is expressed by

$$\chi_{CO}^2=\left(\sum\frac{m_k x_k-n_k y_k}{t_k}\right)^2\Big/\left(\sum\frac{n_k m_k s_k(t_k-s_k)}{t_k^3}\right)\ ,$$

which was introduced by Cochran [6]. The above derivation of χ_{CO}^2 by the use of $N\circ\mathrm{UL}(\beta)$ is equivalent to that of the logit score test by Day and Byar [8]. The statistic χ_{CO}^2 has a simple form and is intuitively appealing. Radhakrishna [26] showed that χ_{CO}^2 is a locally most powerful test statistic for M_0 against M_C. However recall that $N\circ$

UL (β) may be ill-approximated to UL (β), while $C\circ$UD (δ) is considered to approximate UD (δ) comparatively better. This suggests that χ^2_{CO} is closer to the likelihood ratio test statistic for M_0 against the common standardized difference model, than to χ^2_U.

The use of $N\circ$CL (β) or equivalently $N\circ$AL (β) leads to

$$\chi^2=\left(\sum \frac{m_k x_k-n_k y_k}{t_k}\right)^2\Big/\sum\left(\frac{n_k m_k s_k(t_k-s_k)}{t_k^2(t_k-1)}\right).$$

Mantel and Haenszel [21] recommended the test statistic,

$$\chi^2_{MH}=\left(\sum\left|\frac{m_k x_k-n_k y_k}{t_k}\right|-0.5\right)^2\Big/\sum \frac{n_k m_k s_k(t_k-s_k)}{t_k^2(t_k-1)}.$$

This statistic is apparently close to χ^2_{CO}. The term (t_k-1) in place of t_k is interpreted as adjusting to the conditional from the unconditional likelihood ratio test statistic. Thus χ^2_{MH} is close to the likelihood ratio test statistic for M_0 against the common standardized difference.

Next we move to the test statistic for M_0 against M_S. Summing up each chi square statistic over all strata, two statistics are derived;

$$\chi^2_{WS}=\sum w_k(\log x_k(m_k-y_k)/y_k(n_k-x_k))^2$$

and

$$\chi^2_{COS}=\sum \frac{t_k(m_k x_k-n_k y_k)^2}{n_k m_k s_k(t_k-s_k)}.$$

The test statistics for M_C against M_S are reasonably defined by $\chi^2_{WS}-\chi^2_W$ and $\chi^2_{COS}-\chi^2_{CO}$. Unfortunately the latter statistic is unacceptable under some conditions which Mantel et al. [22] presented explicitly. They assert that $\chi^2_{COS}-\chi^2_{CO}$ is based on the standardized difference which is an unacceptable measure because of its dependence on the sample sizes, n_k and m_k. Their assertation agrees with ours that χ^2_{CO} can be regarded rather as a test statistic for M_0 against the common standardized difference model.

Example. We emphasize the χ^2_{CO} is close to the likelihood ratio test statistic for M_0 against the common standardized difference. Consider the two 2×2 tables in Table 1. It follows that χ^2_{CO} and $\chi^2_{COS}-\chi^2_C$ are 3.766 and .200, respectively. On the other hand we get $\chi^2_U=$

Table 1. Working example

	Stratum 1		Stratum 2	
	With factor	Without factor	With factor	Without factor
Diseased	5	50	50	5
Not diseased	50	980	500	98

3.905, and the likelihood ratio test statistic for the homogeneity of odds ratios is 0 since the common odds ratio of both tables is 1.96=exp (.673). Numerical calculation of the likelihood ratio test statistics for M_0 against the common standardized difference and for the homogeneity of standardized differences results in 3.748 and .157, respectively. These are close to χ^2_{CO} and $\chi^2_{COS}-\chi^2_{CO}$.

The good approximation of χ^2_{CO} to the likelihood ratio test statistic for M_0 against the common odds ratio is observed when the estimated common standardized difference is small. In the above example the estimated value is .594. When the estimated common standardized difference is large, say 2, and the total sample size is moderately large, the test based on any test statistic should be highly significant.

7. Discussion

Common features of the conventional procedures are the understandability and the ease of the computation. Though the former advantage is important indeed, we should keep in mind the fact that the intuitive reasoning in the analysis of multiple 2×2 tables is often misleading as seen in Fleiss [11]. The latter advantage is becoming less important, since the cost-performance of computation is dropping rapidly while the expense of obtaining data is still rising. The computation to derive the UML estimator and the likelihood ratio test statistic in the multiple 2×2 tables contain no special difficulties. The desk-top calculator performs the computation and no special device in programming is required, since the functions appearing in the likelihood equation have favorable properties for maximization. In the case of the conditional likelihood, the computation is possibly troublesome, especially when all the numbers, n_k, m_k, s_k and t_k-s_k, are moderately large. But in such a case we can expect that the approximation of AL(β) to CL(β) is satisfactory.

Secondly, the model assumed in the analysis is not clear in the Cochran and Mantel-Haenszel procedures. The Yates estimator $\hat{\beta}_Y$ and test statistic χ^2_{CO} are appropriate when the commen standardized difference is assumed. The Mantel-Haenszel estimator $\hat{\gamma}_{MH}$ and the test statistic χ^2_{MH} are more confusing. We can not find any criterion to derive both the statistics simultaneously. As shown in Section 5, $\hat{\gamma}_{MH}$ and χ^2_{MH} are derived as approximations of the UML estimator and the likelihood ratio test under the different models. Using the recent works by Hauck [17] and Breslow and Liang [5] on the variance of the Mantel-Haenszel estimator, we may provide an alternative test statistic. But it seems that further work on this topic is needed.

The third defect of the conventional methods, which is most serious in practice, pertains to difficulties in extending the models to be fitted. Though our attention focuses on three models, other candidates for reasonable models to be fitted exist. As Nelder and Wedderburn [25], for example, emphasized, it is worthwhile to consult the goodness of fit of various candidate models. Consider a model between M_0 and M_S such that $\beta_k=\beta_1+\beta_2 z_k$ where z_k is a covariate to the stratum k. Then we can proceed straightforwardly with the analysis based on the UML and CML methods. But the use of the conventional procedure requires additional sophisticated devices to estimate parameters in this model.

We wish to express grateful thanks to a referee for a number of valuable comments.

THE INSTITUTE OF STATISTICAL MATHEMATICS
OKAYAMA UNIVERSITY OF SCIENCE

REFERENCES

[1] Andersen, E. B. (1970). Asymptotic properties of conditional maximum-likelihood estimators, *J. R. Statist. Soc.*, B, **32**, 283-301.

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

[3] Birch, M. W. (1964). The detection of partial association I: The 2×2 case, *J. R. Statist. Soc.*, B, **26**, 313-324.

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

[5] Breslow, N. E. and Liang, K. Y. (1982). The variance of the Mantel-Haenszel estimator, *Biometrics*, **38**, 943-952.

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

[7] Cox, D. R. (1970). *The Analysis of Binary Data*, Methuen and Co., London.

[8] Day, N. E. and Byar, D. P. (1979). Testing hypotheses in case-control studies—equivalence of Mantel-Haenszel statistics and logistic score tests, *Biometrics*, **35**, 623-630.

[9] Farewell, V. T. and Prentice, R. L. (1980). The approximation of partial likelihood with emphasis on case-control studies, *Biometrika*, **67**, 273-278.

[10] Fleiss, J. L. (1970). On the asserted invariance of the odds ratio, *Brit. J. Prev. Soc. Med.*, **24**, 45-46.

[11] Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*, 2nd ed., Wiley, New York.

[12] Gart, J. J. (1971). The comparison of proportions: A review of significant tests, confidence intervals and adjustments for stratification, *Rev. Int. Statist. Inst.*, **39**, 148-169.

[13] Goodman, L. A. (1969). On partitioning χ^2 and detecting partial association in three-way contingency tables, *J. R. Statist. Soc.*, B, **31**, 486-498.

[14] Haldane, J. B. S. (1955). The estimation and significance of the logarithm of a ratio of frequencies, *Ann. Hum. Genet.*, **20**, 309-311.

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

[16] Harkness, W. L. (1965). Properties of the extended hypergeometric distribution, *Ann. Math. Statist.*, **36**, 938-945.

[17] Hauck, W. W. (1979). The large sample variance of the Mantel-Haenszel estimator

of a common odds ratio, *Biometrics*, **35**, 817-819.

[18] Hauck, W. W., Anderson, S. and Leahy, F. L. (1982). Finite-sample properties of some old and some new estimators of a common odds ratio from multiple 2×2 tables, *J. Amer. Statist. Ass.*, **77**, 145-152.

[19] Hitchcock, S. E. (1962). A note on the estimation of the parameters of the logistic function, using the minimum logit χ^2 method, *Biometrika*, **49**, 250-255.

[20] Lubin, J. H. (1981). An empirical evaluation of the use of conditional and unconditional likelihoods for case-control data, *Biometrika*, **68**, 567-571.

[21] Mantel, N. and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease, *J. Nat. Cancer Inst.*, **22**, 719-748.

[22] Mantel, N., Brown, C. and Byar, D. D. (1977). Tests for homogeneity of effect in an epidemiologic investigation, *Amer. J. Epidemiol.*, **106**, 125-129.

[23] Mckinlay, S. M. (1975). The effect of bias on estimators of relative risk for pair-matched and stratified samples, *J. Amer. Statist. Ass.*, **70**, 859-864.

[24] Mckinlay, S. M. (1978). The effect of nonzero second-order interaction on combined estimators of the odds ratio, *Biometrika*, **65**, 191-202.

[25] Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models, *J. R. Statist. Soc.*, A, **135**, 370-384.

[26] Radhakrishna, S. (1965). Combination of results from several 2×2 contingency tables, *Biometrics*, **21**, 86-98.

[27] Woolf, B. (1955). On estimating the relationship between blood group and disease, *Ann. Hum. Genet.*, **19**, 251-253.

[28] Yanagimoto, T. and Kamakura, T. (1984). The maximum full and partial likelihood estimators in the proportional hazard models, *Ann. Inst. Statist. Math.*, A, **36**, 363-373.

[29] Yates, F. (1955). The use of transformations and maximum likelihood in the analysis of quantal experiments involving two treatments, *Biometrika*, **42**, 382-411.