Tests for Trends in Incidence Rate Ratios

M. Nurminen

Institute of Occupational Health, Helsinki, Finland

Summary

Breslow (1984) described an efficient score test for trend in incidence density rate ratios for cohort studies under a conditional Poisson or binomial model employing maximum likelihood estimation of the rate parameters. In this communication, an alternative derivation of this statistic that is based on an unconditional approach is provided, along with an examination of associated goodness-of-fit tests and methods of confidence interval estimation. The procedures are illustrated by a cohort study of ischemic heart disease mortality following industrial exposure to carbon disulfide.

Key words: Asymptotic statistics; Biometry; Cohort analysis; Constrained maximum likelihood estimation; Epidemiologic methods.

1. Introduction

Comparative analysis of several rates is a frequently occurring research task in epidemiology. When concerned with cohort studies the interest typically centers on the examination of a possible trend in incidence density rate ratios over a period of follow-up. For stratified analysis the data are usually categorized into ordered strata with numerical scores assigned to each stratum. For example, Table 1 presents coronary mortality data for carbon disulfide exposed workers and nonexposed male workers in the course of a 13-year prospective follow-up study (Nurminen and Heinonen, 1985) grouped into five disjoint time-intervals. Recently Breslow (1984) presented a score statistic for trend analysis of similar incidence density-type rates. This statistic arose as a modification of the usual test for trend in proportions (Armitage, 1955) under a binomial sampling distribution with the total frequency of occurrence of studied events kept constant. In this paper, the same problem is approached assuming a Poisson probability model for the incidences without "fixing" in advance their total number. The outlook here is mainly applied, but some of the formulations which arise appear to be new.

27 Biometrics 51 (1995) 4
2. The Statistical Model for Trend

When we are interested in a comparative rate parameter \((P)\) and its relation to a potential modifying (background) factor \((X)\) across the strata \((\text{indexed } j = 1, \ldots, J)\), a general two-parameter model to be estimated may be taken as \(P_j = P \cdot f(\beta x_j)\), where \(f\) denotes a monotonous function and \(\beta\) the parameter of trend (slope). For rate difference parameter one often specifies a simple linear model \(f(\beta x_j) = 1 + \beta x_j\), and for rate ratio (cf. Breslow, 1984) and odds ratio (cf. Zeelen, 1971) an exponential one: \(f(\beta x_j) = \exp(\beta x_j)\).

We assume that a reasonable approximation to the sampling distribution of incidence density data is to take the number of cases evolved in two populations, \(c x_j (1 = \text{exposed})\) and \(c 0 (0 = \text{nonexposed})\), in the \(j\)th period of follow-up as independent Poisson variates with expectations \(T x_j R x_j\) and \(T 0 R 0\), where \(R x_j\) and \(R 0\) are the unknown population incidence rates and \(T x_j\) and \(T 0\) are the corresponding population-times of study experience. Further we specify that the theoretical rate ratios, \(R x_j/R 0\), follow an exponential model as a function of an effect-modifier \(x_j\) (assumed positive, frequently a time-variante such as exposure-time): \(R x_j = \exp(x + \beta x_j) = (RR) \exp(\beta x_j)\), where \(RR = \exp(a)\) plays the role of baseline \((X = 0)\) rate ratio. As for scoring, one usually simply equates \(x_j\) with the mid-point of the \(j\)th stratum, for instance age-interval or time-interval following a defined zero time.

We shall in the following proceed to describe efficient tests for trends in incidence rate ratios and for choosing between alternative models.

3. Hypothesis-Testing and Rate-Estimation

In general, the main interest under the model will be in the significance testing of whether there is a uniform relation between the exposure and the disease after adjusting for their regression on the modifying factor, i.e., assessing the hypothesis that \(RR x_j = RR\). Given uniformity of \(RR\) over \(X\), testing the various null hypotheses that \(RR = RR 0\) might be called for, in particular the one that there is no association or \(RR 0 = 1\). Because it is in most cases impossible to meaningfully specify an alternative pattern for trend, the evaluation of the appropriateness of the exponential model is done in contrast to the general alternative, viz., \(RR x_j\) being fully parametrized or nonparametric.

A high-powered test of the hypothesis of a uniform rate ratio \((RR x_j = RR)\) against the alternative of an exponential trend is given by the following formula, which is derivable from the log-likelihood as the efficient score statistic (see Appendix 1):

\[
X^2 = \left(\sum U_j\right)/\left(\sum U_j\right).
\]

This statistic, in which the numerator units \(U_j = x_j T x_j / (r x_j - r 0)\) score systematically any deviation in the empirical rate ratios, \(r x_j = r x_j / r 0 = (T x_j / T 0) / (c x_j / c 0)\),
from the hypothesized model course with the passage of time, follows asymptotically a chi-square distribution with one degree of freedom (d.f.). First, we challenge the hypothesis that no trend exists under the model (i.e. $\beta = 0$). Here at the $RR$-dependent estimates of the rate parameters $R_{ij} = R_{ij}(RR)$ and $R_{ij} = c_i [T_{ij}(RR) + T_{ij}]$, $j = 1, \ldots, J$, are obtained as unconditional (conditional-free) maximum likelihood (ML) estimates under the assumption of uniform $RR$; the likelihood equation for the estimation (iterative) of $RR$ is $\sum U_j (x_j - 0) = 0$ (GART, 1973). The variance estimator of $\sum U_j$ can be found, constrained on the actual value of $RR$, to be

$$V(\sum U_j) - \text{var}(\sum U_j | \beta = 0) = \sum W_j (x_j - 2)^2,$$

where $2 = \sum W_j (x_j - 2)^2 / \sum W_j$ with $W_j^{-1} = 1 / (T_{ij} R_{ij}) + 1 / (T_{ij} R_{ij})$. Expression 2 is equal to a weighted sum of squares of the $x_j$ about their mean (cf. ASHMANI, 1982), the weights being the inverse of the constrained ML estimates of the variance of $\log (x_j)$, $W_j = V_j = (T_{ij} R_{ij})^{-1} + (T_{ij} R_{ij})^{-1}$ (NURMINEN, 1984). The resulting chi-square statistic (expression 7) is, in fact, exactly the same as that given by BRASLOW (1984), derived conditioning by the interval-specific marginal rates.

Second, provided that it is inferred that $\beta \neq 0$, to test the tenability of the model of an exponential trend against the general alternative ($RR$ unspecified) we compare the value of the goodness-of-fit statistic

$$X^2_{j-2} = \left[ \sum \left[ c_i [R_{ij}(RR) - T_{ij} R_{ij}] \right]^2 / (U_j (x_j - 2)^2) \right] \exp (J - 2) \sum \log [W_j (x_j - 2)^2]$$

with Tables of chi-square with $J - 2$ d.f. Expression 3 arises as a score test by a straightforward derivation, analogous to that given in Appendix 1. The joint likelihood equations for the estimation of $\alpha$ and $\beta$ are $\sum U_j (x_j - 0) = 0$ and $\sum U_j (x_j - 2)^2 = 0$, respectively, with $RR = \exp (x + \beta x_j)$.

For the corresponding (asymptotically equivalent) chi-square statistics based on the likelihood-ratio approach, see Appendix 2.

In case that the model of nonuniform rate ratio is not consistent with the data, the evaluation of $RR - RR_0$ may be based on the chi-square function of $RR$ (NURMINEN, 1984; MIEHTINEN and NURMINEN, 1985, p. 220) on one d.f.:

$$X^2_1 = (\sum w_j d_j)^2 / (\sum w_j d_j)^2,$$

which is a weighted average of the period-specific deviances $d_j = v_j - \{RR \} v_j$ and the weights $w_j = 1 / (T_{ij} + (RR) [T_{ij} R_{ij}])^{-1}$ being inversely proportional to the variances $V(d_j) = T_{ij} R_{ij} + (RR) R_{ij} / T_{ij}$. For the base null proposition $RR = 1$ (or $\alpha = \beta = 0$), the chi-square statistic in expression 4 reduces to the form of

$$X^2_1 = \left[ \sum T_{ij} \left( x_j - 0 \right) \right]^2 / \{ \sum \left[ T_{ij} + 1 / (T_{ij} + 1) \right] \},$$

where the marginal rate $v_j = c_i / T_{ij}$. This statistic, which has been proposed by several authors (MIEHTINEN, 1972; SHEIN et al., 1976; ROOTHM and BOISE, 1977), is the modified version of the MAZTEL-HAENZEL (1959) test for use in cohort data analyses.
Asymptotic confidence intervals for the uniform $RR$ are obtainable iteratively from
\begin{equation}
(\hat{RR}; \prod_{ij} (r_{ij} - R_0)/(\sum W_{ij})^{1/2} = \chi_{(2)}^2),
\end{equation}
where $R_0$ is the upper $1 - \alpha/2$ point of the standardized Gaussian distribution. The statistic underlying the method of setting confidence limits defined in expression 6 derives as the efficient score test of $RR = RR_0$ versus $RR + RR_0$. This latter form, however, is algebraically identical with $X^2$ given in expression 4 above. An advantage of this procedure is that $\hat{RR}$ does not have to be calculated explicitly, although often the point estimation (i.e., $x = 1$) will be of intrinsic interest. By contrast, the usual procedure is to construct the interval around $\hat{RR}$ on a logarithmic scale (see below).

4. An Illustration

NURMISCHEN and HENNINGER (1963) monitored the effects of intervention on ischemic heart disease among 343 carbon disulfide (CS) exposed male workers and 343 nonexposed paper mill workers, selected comparable ("matched") with respect to sex, age and type of work. The 15-year prospective phase of the study was carried out from 1950 to 1965. The preventive intervention program included among other actions removing all workers with coronary risk factors from exposure. The right hand column of Table 1 indicates a steady decline in the coronary death rate ratios. Thus, there is substantial interest as to whether the data are satisfactorily modeled by a log-linear trend with time: $\log (RR_j) = a + bT_j$. To improve the adequacy of the asymptotic Gaussian approximation the time intervals have been grouped so that there are deaths observed in each period. The rates are low and $T_j = T_{j+}$ so that the assumption of independence of out-

<table>
<thead>
<tr>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
</tr>
<tr>
<td>Exp.</td>
</tr>
<tr>
<td>Non Exp.</td>
</tr>
<tr>
<td>Observed deaths</td>
</tr>
<tr>
<td>Non Exp.</td>
</tr>
<tr>
<td>Rate ratio</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>j</th>
<th>$T_{ij}$</th>
<th>$T_{ij}$</th>
<th>$x_{ij}$</th>
<th>$y_{ij}$</th>
<th>$z_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0</td>
<td>1.020</td>
<td>1.029</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4.0</td>
<td>4.622</td>
<td>4.789</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>6.0</td>
<td>8.860</td>
<td>8.863</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>9.5</td>
<td>5.007</td>
<td>5.177</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>13.5</td>
<td>9.254</td>
<td>9.533</td>
<td>6</td>
<td>10</td>
</tr>
</tbody>
</table>

* The successive periods were the 1st-3rd, 4th-5th, 6th-7th, 8th-12th and 13th-15th year of follow-up.

** Mid-point of time-interval (i.e., years of follow-up)
comes from those in the previous intervals would seem to be valid. Applied to the
data of Table 1, the ML estimates for the parameters are found, employing the
Newton-Raphson method (see, e.g., McCullagh and Nelder, 1989, p. 144), to be
$a = 2.004$ and $\beta = -0.1806$. To provide for the initial values for $a$ and $\beta$, a straight
line may be fitted visually to the values $(y_i, \log(y_i))$. With the final estimates,
the rate ratio function evaluated at the mean scores for the periods yields the
$R_{k}$ estimates of 5.66, 3.61, 2.33, 1.34 and 0.65 for period 1 through 5, respectively
(see Figure 1). The test for trend ($\beta = 0$ vs $\beta \neq 0$) according to expressions 1

![Graph](image)

Fig. 1. Mortality rate of ischemic heart disease among workers exposed to carbon disulfide relative to unexposed workers during a 15-year period following intervention. Shown are the period-specific estimates given in Table 1, together with the fitted rate ratio function.

and 2 is found to be $X^2_1 = 6.68$ and, a one-sided $p$-value is $p = 0.0049$. This result
may be compared with the value of the likelihood ratio test for trend (expression 8 with $R_{k_{ML}} = 1.571$) $X^2_2 = 7.01$ (one-sided $p = 0.0041$) found by Virtanen and Hernberg (1989). As a first approximation to solving $R$, one may
take the pooled estimator $\hat{R} = 1.547$ or the Mantel-Haenszel estimator $\hat{R}_{MH} = 1.572$. The test for the adequacy of the log-linear model is computed from expression 3 to be $X^2_3 = 3.29$, indicating no more dispersion than what would be expected by random variation alone. By way of comparison, the likelihood ratio-based goodness-of-fit test (expression 9) takes on the value of $X^2_4 = 3.17$, again indicating a good agreement for these fairly small data series. We conclude that
almost all of the variability (heterogeneity) in the observed period-specific rate
ratios is due to the decremental exponential trend in the course of the 15 years.

By way of illustration, the ninetieth per cent confidence interval for the common $R$ is found (iteratively) from expression 4 to be $(0.953, 2.889)$, which coincides with that obtained from the alternative ML chi-square procedure (Miettinen and Virtanen, 1986). These limits may be compared with the ones derived, under the assumption of a log-Gaussian distribution for the ML estimator of
and supplemented with a Taylor series approximation to it (e.g., Barlow, 1984), as \( RR_{ML} \cdot \exp [\pm 1.969 (\sum V_j)]^{1/2} = (0.949, 2.602) \). Even though they give quite similar results, the latter approach can be problematic: any stratum with \( c_0 = 0 \) makes an infinite contribution to the variance \( V_j \), and thereby renders the method inapplicable.

Finally, formula 5 yields for these data \( X^2 = 3.13 \) (NS); therefore the model of uniform \( RR \) is not a sufficiently powerful (sensitive) alternative to detect a departure from the universal null hypothesis (i.e., \( RR = 1 \)).

5. Discussion

In the derivation of the score statistics for trend analysis it was assumed that the \( c_j \)'s in the subsequent intervals are unrelated to each other. For independent studies the assumption is self-evident and for dynamic populations with rapid changeover it is quite generally fulfilled as well. In case of cohort studies with low incidence densities it is reasonable to regard \( c_j \) as being independent of the events in the preceding intervals.

Under the circumstance of a uniform \( RR \) the expected number of cases \( E(c_j) = T_{1j}R_{1j} + T_{0j}R_{0j} = c_j \). (This relation holds for odds ratio but not for ratio in the context of proportion-type rates; Nurminen, 1981). It follows that conditioning by the ancillary, statistic \( c_j \), does not change the ML estimates of the \( RR_j \)'s and the \( R_0 \)'s so that we have the same efficient scores \( U_j \) for different (Poisson and binomial) sampling schemes (cf. Birch, 1963). Thus the unconditional modeling here and the conditional one (Barlow, 1984) yield completely interchangeable statistics. Yet the unconditional Poisson distribution is the natural sampling model for epidemiologic studies involving incidence density-type rates.

Extension of this technique may be adapted for use with (i) stratified prevalence (proportion-type) data, (ii) alternative comparative parameters (e.g. rate difference and odds ratio), and (iii) other model forms (such as \( RR_j = \exp (a_j S_j) \)) in a like fashion to that shown here.

Appendix 1. Efficient Score Based Trend Statistic

Suppose that \( c_{1j} \) and \( c_{0j} \) (j = 1, ..., J) are mutually independent Poisson variates having means \( T_{1j}R_{1j} = T_{0j}R_{0j}(RR) \) \( \exp (S_j) \) and \( T_{0j}R_{0j} \), respectively, all parameters being unknown and the \( T_j \)'s and \( T_j \)'s and \( a_j \)'s known constants. The log likelihood \( L_j \) for a single data set is, apart from a constant,

\[
L_j = -T_{1j}R_{1j}(RR) \exp (S_j) \\
+ c_{1j} \log T_{1j}R_{1j}(RR) \exp (S_j) \\
- T_{0j}R_{0j} + c_{0j} \log (T_{0j}R_{0j}) .
\]
If we define \( q = (RR, R_{01}, \ldots, R_{0J}) \), then by independence the efficient score statistic for testing \( \beta = 0 \) (see, e.g., Cox and Hinkley, 1974. Section 9.3) is

\[
[U_s(\beta, \hat{e}) - b_s U_s(0, \hat{e})] = \frac{1}{\sum_{j=1}^{J} U_{j,s}(\beta, \hat{e}, RR, R_{0j})},
\]

where \( b_s \) is the total vector (denoted by a dot suffix) of coefficients of regression of \( U_s(\beta, \hat{e}) \) on \( U_s(\beta, \hat{e}) \) evaluated at \((0, \hat{e})\). The part of the score vector corresponding to \( \beta \) is

\[
U_s(\beta, \hat{e}) = \sum_{j=1}^{J} U_{j,s}(\beta, RR, R_{0j}),
\]

where \( U_{j,s} = \partial E(\beta, RR, R_{0j}) / \partial \beta \) (\( j = 1, \ldots, J \)), since \( R_{0j} \) appears only for the pair \((c_{1j}, c_{0j})\) and for no other \((c_{1k}, c_{0k})\) \((k \neq j)\). In this particular instance, the components

\[
U_{j,s}(\beta, \hat{e}) = (\partial / \partial \beta RR), \quad \partial / \partial \beta R_{0j}) L_{j}(\beta, RR, R_{0j})
\]

of \( U_s(\beta, \hat{e}) = \sum_{j=1}^{J} U_{j,s}(\beta, \hat{e}) \) are zero if we substitute the ML estimates for the \( RR \) and the \( R_{0j} \)'s under the hypothesis of \( \beta = 0 \). These estimates are solvable from the joint ML equations \( \sum_{j=1}^{J} T_{1j} R_{0j}(RR) \) and \( R_{0j} = c_{0j}[(T_{1j}(RR) + T_{0j}) \) \((j = 1, \ldots, J)\). Therefore the score statistic is simply the standardized form of \( U_s(0, \hat{e}) \) evaluated at \( RR = \tilde{RR} \) and \( R_{0j} = \tilde{R}_{0j} \) \((j = 1, \ldots, J)\). We find readily from the log likelihood that

\[
U_s = \sum_{j=1}^{J} \lambda_j [c_{1j} - E(c_{1j})],
\]

where \( E(c_{1j}) = T_{1j} \tilde{R}_{1j} = T_{1j} R_{0j}(\tilde{RR}) \).

The standardizing term \( \lambda_j \) is the leading element of the inverse of the total information matrix \( I_s(0, \hat{e}) = \{i_{s,ss}(0, \hat{e})\} \), where

\[
i_{s,ss}(0, \hat{e}) = \sum_{j=1}^{J} \lambda_j E[-\partial^2 E(\beta, \hat{e}) / \partial \beta \partial \hat{e}; \beta = 0, e = \hat{e}].
\]

We partition the matrix \( I_s \) according to the partition of \((\beta, RR, R_{0j})\) and take advantage of the crucial property \( E(c_{1j}) = T_{1j} \tilde{R}_{1j} + T_{0j} \tilde{R}_{0j} = c_{1j} \) to obtain

\[
l_{s}(\beta, \tilde{RR}, R_{0j}) = \begin{bmatrix} a & b & c^T \\ b & e & f^T \\ c & f & d \end{bmatrix},
\]

where

\[
a = \sum j c_{1j} T_{1j} R_{0j}(\tilde{RR}), \quad b = \sum j c_{1j} T_{0j}, \quad c^T = \{x_1, T_{11}(\tilde{RR}), \ldots, x_1 T_{1j}(\tilde{RR})\},
\]

\[
d = \text{diag} \{T_{1j}(\tilde{RR}) + T_{0j}|R_{0j}, \ldots, [T_{1j}(\tilde{RR}) + T_{0j}|R_{0j}]\}, \quad e = \sum j T_{1j} R_{0j}(\tilde{RR}), \quad f^T = (T_{1j}, \ldots, T_{1j}).
\]
The cofactor pertaining to the element \(i, j\) in \(I_{d} = [e - (-d - 1)(d)]\) and the determinant of \(I\), is easiest to derive as

\[
|I| = \begin{bmatrix} a & b \\ b & d \end{bmatrix} = \begin{bmatrix} \frac{1}{d} \\ \frac{a}{d} \end{bmatrix} d^{-1} \begin{bmatrix} d \\ a \end{bmatrix},
\]

which, after some algebra, gives the result

\[
i_{d}^{-1} = I_{d}^{-1} |I| = \sum_{j=1}^{d} \frac{1}{d} w_{j} (\sum_{j=1}^{d} w_{j})^{-1} \sum_{j=1}^{d} w_{j} (x_{j} - \bar{x})^{2},
\]

where

\[
W_{j}^{-1} = 1/\left(\sum_{j=1}^{d} T_{ij} R_{ij} + 1/(T_{ij} R_{ij})\right) \log (\rho_{ij}) \text{ and } \bar{x} = \sum W_{j} x_{j} / \sum W_{j}.
\]

The chi-square test for trend based on the efficient score thus assumes the concise form of

\[
X_{d}^{2} = \sum_{j=1}^{d} \frac{(T_{ij} - R_{ij})}{\sum_{j=1}^{d} W_{j} (x_{j} - \bar{x})^{2}}.
\]

The asymptotic chi-square distribution of \(X_{d}^{2} = U_{d}^{2} \chi_{d}^{2}\) with one d.f. applies here when \(T_{ij}\) and \(R_{ij}\) both increase indefinitely, a necessary condition for unbounded growth if \(i_{d} = 0\), \(R_{ij}\). This is seen from the form of \(U_{d}\) by writing it as

\[
U_{d} = \sum_{j=1}^{d} c_{ij} \left[ \frac{T_{ij} R_{ij}}{T_{ij} (R_{ij})} + \frac{1}{T_{ij} (R_{ij})} \right].
\]

Appendix 2. Likelihood Ratio-Based Trend Statistics

The asymptotic likelihood ratio test for the hypothesis that no trend actually exists (i.e. \(\beta = 0\) or \(R_{ij} = R_{i}\), \(i = 1, \ldots, J\)) may be found on the chi-square statistic on one d.f. (Miettinen, 1972; cf. Cox and Hinklev, 1974, Section 9.3)

\[
X_{d}^{2} = -2 \left[ \sum_{j} c_{ij} \log \left( \frac{R_{ij} R_{ij}}{\exp (\alpha + \beta x_{j})} \right) \right] + \sum_{j} c_{ij} \log \left[ \frac{\exp (\alpha + \beta x_{j})}{R_{ij} R_{ij} + T_{ij}} \right],
\]

where \(\hat{R}_{ij}\) is solved by trial and error from the likelihood equation \(c_{ij} = U_{ij} c_{ij} / [T_{ij} (R_{ij}) + T_{ij}]\).

To test the equality of the model \(\log (R_{ij}) = \alpha + \beta x_{j}\), the large-sample likelihood ratio statistic with \(J - 2\) d.f. (\(J\) is the number of strata used for \(X\)) is

\[
X_{j-2}^{2} = -2 \left[ \sum_{j} c_{ij} \log \left( \frac{\exp (\alpha + \beta x_{j})}{\rho_{ij}} \right) \right] + \sum_{j} c_{ij} \log \left[ \frac{\rho_{ij} + T_{ij} / T_{ij}}{\exp (\alpha + \beta x_{j}) + T_{ij}} \right],
\]

where \(\rho_{ij} = (c_{ij} / T_{ij}) / (c_{ij} / T_{ij})\).
Tests for Trends in Incidence Rate Ratios

References


Manuscript received: Nov. 11, 1985

Author’s address:

Dr. Markku Nurminen
Department of Epidemiology and Biostatistics
Institute of Occupational Health
Topolanenkatu 41 a A
SF-00290 Helsinki
Finland