Maximum (Max) and Mid-P Confidence Intervals and p Values for the Standardized Mortality and Incidence Ratios

Pandurang M. Kulkami, 1 Ram C. Tripathi, 2 and Joel E. Michalek 3

The authors present algorithms to compute the confidence interval for a Poisson parameter (λ) and the p value for testing the hypothesis that λ is equal to a constant, which can be used to make inferences about the standardized mortality ratio and the standardized incidence ratio. The p value and confidence interval always agree, despite the discrete nature of the Poisson distribution. The authors also give simple equations to compute numeric approximations of the confidence limits that do not require the use of any probability distributions. An example based on data arising from a study of cancer incidence is given. Am J Epidemiol 1998;147:83-6.

confidence intervals; Poisson distribution; risk; significance tests

In epidemiologic studies, interest often lies in comparing the mortality rate of a certain cohort with that of a standard population; a standard summary statistic in this regard is the standardized mortality ratio (SMR). In studies of disease incidence, the corresponding statistic is the standardized incidence ratio (SIR). Traditionally, the observed number of events is assumed to have a Poisson distribution. Thus, inference procedures for the SMR (or SIR) can be formulated based on those for the Poisson distribution. Many authors have discussed methods for constructing confidence intervals for the SMR. Some have provided tables (1–3), and others give methods based on some approximations (3–5). An exact method of constructing confidence intervals was suggested by Garwood (6), and it reappeared in an article by Ulm (7). Tests of hypothesis for the SMR have been considered by Monson (8), Armitage (9), and Rothman and Boice (10). Many tests are based on the normal approximation to the Poisson distribution. These approximations work well when the observed number of deaths is large, but they require the use of normal and chi-square distribution tables. Here, we provide simple equations to compute numeric approximations of the confidence intervals without requiring the use of tables. Further, exact tests for the SMR and their relation to confidence intervals are not well documented. Ideally, inferences based on confidence intervals and hypothesis tests should lead to the same decision. However, when using approximate procedures, this may not be the case. In this paper, we also provide procedures to calculate p values that are inverted to obtain confidence intervals. Therefore, the two procedures always lead to the same decision.

Generally, with discrete distributions it is not possible to construct confidence intervals with specified coverage (the probability that the interval contains the parameter of interest). Therefore, one uses confidence intervals with the nominal coverage as the lower bound for the actual coverage. This results in conservative p values and conservative confidence intervals, implying that the significance level of the test is less and the coverage probability of the confidence interval is greater than nominal. Procedures to give shorter intervals are available (11, 12); however, those methods are computer intensive. Here, we propose a method (known as the Mid-P method (13, 14)) to obtain p values and confidence intervals that are less conservative than the methods of Garwood (6) and Ulm (7). For this case also, we propose simple equations to generate numerically approximate confidence intervals without the use of tables. We apply the results to cancer incidence data in Air Force air crews (15).
CONFIDENCE INTERVALS AND p VALUES

Max method

Let \( X \) denote the observed number of deaths in a cohort study and assume that \( X \) follows a Poisson distribution with expectation \( \lambda \). One is usually interested in knowing if \( \lambda \) is the same as \( \lambda_0 \), the corresponding parameter from a reference population. Therefore, one is interested in testing the null hypothesis \( H_0: \lambda = \lambda_0 \). In a particular study, based on the observed \( X = x \), one wants to decide if \( H_0 \) is to be rejected. This decision can be made by constructing a confidence interval for \( \lambda/\lambda_0 \), where we reject \( H_0 \) if the interval does not include 1. In some studies, it is usual to report the \( p \) value associated with the hypothesis test. When the underlying distributions are continuous, one can obtain a confidence interval and a \( p \) value that would result in the exact nominal significance level and coverage probability, whereas with discrete distributions it is generally not possible to do so (13). For discrete distributions, an exact \( p \) value can be computed using the formula

\[
p = 2 \left[ \gamma \Pr(X = x) + \min\{\Pr(X < x), \Pr(X > x)\} \right]
\]

where \( 0 < \gamma < 1 \) and the probabilities are obtained under \( H_0 \) (13, 14). If \( X \) is continuously distributed, then \( \Pr(X = x) = 0 \) and the term associated with \( \gamma \) is zero. One can use a randomization process to choose the value of \( \gamma \) via a uniform distribution (16, 17); however, this procedure can be disturbing to a practitioner for one must add random noise to the data to produce a \( p \) value (12). Instead, it is a general practice (13) to choose \( \gamma = 1 \). This choice leads to the most conservative \( p \) values and hence to the most conservative confidence intervals obtained by inverting the \( p \) value formula. Because these procedures result from \( \gamma \) obtaining its maximum (Max) value, we refer to the resultant procedures as Max confidence intervals and \( p \) values. The confidence intervals, with \((1 - \alpha)\) 100 percent confidence, with \( \gamma = 1 \) are given by

\[
\lambda_{\text{Max, lower}} = -\frac{1}{2} \chi^2_{2x,1/2}
\]

and

\[
\lambda_{\text{Max, upper}} = \frac{1}{2} \chi^2_{2x+1,1/2}
\]

where \( \chi^2_{df,b} \) denotes the upper \( b \) cutoff point of a chi-square distribution with \( df \) degrees of freedom and for \( df = 0 \), \( \chi^2_{0,b} = 0 \). This is derived using a well-known relation between Poisson and chi-square distributions (7). Let \( p_{\text{Max}} \) denote the corresponding Max \( p \) value. Then, \( p_{\text{Max}} = 2[p_1 - p_2 + \min(1 - p_1, p_2)] \), with \( p_1 = \Pr(\chi^2_{2x} \leq 2\lambda_0) \) and \( p_2 = \Pr(\chi^2_{2(x+1)} \leq 2\lambda_0) \). Further, because the confidence intervals are obtained by simply inverting the \( p \) value formula, both the \( p \)

**TABLE 1.** Coefficients for conservative (Max) approximate confidence intervals (equations 5 and 6)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
<th>( d_1 )</th>
<th>( d_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>1.975642</td>
<td>0.742982</td>
<td>0.000600</td>
<td>-3.223824</td>
<td>4.799837</td>
<td>1.247166</td>
</tr>
<tr>
<td>0.05</td>
<td>1.060019</td>
<td>0.809256</td>
<td>0.000436</td>
<td>-2.617304</td>
<td>2.464286</td>
<td>1.187111</td>
</tr>
<tr>
<td>0.10</td>
<td>0.613212</td>
<td>0.841582</td>
<td>0.000359</td>
<td>-2.252430</td>
<td>1.537202</td>
<td>1.156689</td>
</tr>
</tbody>
</table>

**TABLE 2.** Coefficients for the less conservative (Mid-P) approximate confidence intervals (equations 7 and 8)

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( c_1 )</th>
<th>( c_2 )</th>
<th>( c_3 )</th>
<th>( c_4 )</th>
<th>( d_1 )</th>
<th>( d_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.216878</td>
<td>0.425475</td>
<td>0.485901</td>
<td>-1.243794</td>
<td>2.085711</td>
<td>0.442950</td>
</tr>
<tr>
<td>0.05</td>
<td>0.088608</td>
<td>0.436806</td>
<td>0.497343</td>
<td>-0.936323</td>
<td>1.265344</td>
<td>0.438945</td>
</tr>
<tr>
<td>0.10</td>
<td>0.034804</td>
<td>0.442301</td>
<td>0.689090</td>
<td>-0.778045</td>
<td>0.940450</td>
<td>0.437537</td>
</tr>
</tbody>
</table>
value and the confidence interval always lead to the same decision for any chosen level of significance. In figure 1 we have plotted the coverage of the conservative (Max) confidence intervals for values of \( \lambda \) between 0.1 and 50 with increments of 0.1 with \( \alpha = 0.05 \). Figure 1 also shows the corresponding plot with \( \alpha = 0.10 \). It is clear that the coverage is much more than the desired level; that is, the intervals are conservative.

**Mid-P method**

Now, we propose a method for obtaining less conservative \( p \) values and corresponding intervals. We obtain these by choosing \( \gamma = 0.5 \) instead of \( \gamma = 1.0 \). This approach has gained popularity in epidemiologic studies and has been called the Mid-P procedure (13, 18). In this case, the \( p \) value is \( p_{\text{Mid}} = 2[p_1 - p_2 + \min(1 - p_1, p_2)] \). The lower and upper Mid-P limits for \( \lambda \) are solutions of two equations; \( \lambda_{\text{mid, lower}} \) is the solution of

\[
\Pr[\chi^2_{k+1} \leq 2\lambda_{\text{mid, lower}}] + \Pr[\chi^2_{k} \leq 2\lambda_{\text{mid, lower}}] = \alpha \tag{3}
\]

and \( \lambda_{\text{mid, upper}} \) is the solution of

\[
\Pr[\chi^2_{k+1} \leq 2\lambda_{\text{mid, upper}}] + \Pr[\chi^2_{k} \leq 2\lambda_{\text{mid, upper}}] = 2 - \alpha \tag{4}
\]

In figure 1, we have also plotted the coverage of the Mid-P confidence intervals for values of \( \lambda \) between 0.1 and 50 with increments of 0.1 with \( \alpha = 0.10 \). Figure 1 additionally shows the corresponding plot with \( \alpha = 0.05 \). It is clear from figure 1 that the Mid-P procedure is less conservative than the Max procedure.

The average coverage, for the values of \( \lambda \) considered, for \( \alpha = 0.10 \) were 92.4 percent for Max and 90.5 percent for Mid, and for \( \alpha = 0.05 \) were 96.3 percent for Max and 95.3 percent for Mid.

The Mid-P method requires the solution of nonlinear equations; therefore, one must use a computer to obtain the values of \( \lambda_{\text{mid, lower}} \) and \( \lambda_{\text{mid, upper}} \). Because many epidemiologic researchers use SAS software (19), we have written a SAS code to compute confidence intervals and \( p \) values for both Max and Mid-P methods. This code is available on request. Alternatively, in the next section we give simple equations that can be used to calculate numerically approximate intervals for both procedures. These do not require any distributional tables and can be calculated with a hand calculator.

**APPROXIMATE CONFIDENCE INTERVALS**

We derived simple equations to approximate the percentile points of chi-square distributions, as in the paper by Gilbert (20). These equations can be used to obtain the confidence limits when the observed number of events is less than or equal to 50. Let \( \nu = 2x \). The Max lower and upper limits are approximated by

\[
\lambda_{\text{max, lower}} \approx 0.5[a_0 + a_1 \nu + a_2 \nu^2 + a_3 \log(\nu)] \tag{5}
\]

\[
\lambda_{\text{max, upper}} \approx 0.5[b_0 + b_1(\nu + 2) + b_2(\nu + 2)^2 + b_3 \log(\nu + 2)]. \tag{6}
\]

The Mid-P limits are approximated by

\[
\lambda_{\text{mid, lower}} \approx c_0 + c_1 \lambda_{\text{max, lower}} + c_2 x + c_3 \log(x) \tag{7}
\]

\[
\lambda_{\text{mid, upper}} \approx d_0 + d_1 \lambda_{\text{max, upper}} + d_2 x + d_3 \log(x). \tag{8}
\]

The coefficients for equations 5 through 8 are listed in tables 1 and 2 for \( \alpha = 0.01, \alpha = 0.05 \), and \( \alpha = 0.10 \).

**EXAMPLE**

We apply these methods to data arising from a study of cancer in Air Force air crews (15). Grayson and Lyons report 19 cases of cancer of the endocrine system in air crews with an SIR of 1.61 when compared with nonflying Air Force officers. They report a 99 percent confidence interval (CI), based on a normal approximation (3), of 0.81-2.83. The Max procedure gives 99 percent CI 0.82-2.83 with \( p = 0.065 \). The Mid-P procedure gives 99 percent CI 0.84-2.78 with \( p = 0.05 \). The Max approximation (equations 5 and 6) gives 99 percent CI 0.82-2.83, and the Mid-P approximation (equations 7 and 8) gives 99 percent CI 0.84-2.78. We obtained our confidence intervals by computing the 99 percent CI for \( \lambda \) and dividing the endpoints by the expected number of cases, computed as the observed number of cases divided by the SIR.

Even though we have referred to the confidence intervals arising from equations 5 through 8 as approximate, we have found that both the Max and Mid-P approximations are correct to within two decimals when \( x \) is between 2 and 50 and the expected value of \( X \) is greater than or equal to 5.

**ACKNOWLEDGMENTS**

Dr. Kulkarni’s research was partially funded by the National Research Council.
REFERENCES