STATISTICS IN MEDICINE, VOL. 10, 457–462 (1991)

CONFIDENCE INTERVALS FOR WEIGHTED SUMS OF POISSON PARAMETERS

ANNETTE J. DOBSON
Department of Statistics, University of Newcastle, New South Wales 2308, Australia

KARI KUULASMAA
MONICA Data Centre, National Public Health Institute, Mannerheimintie 166, SF-00300 Helsinki, Finland

ERNST EBERLE
GSF-Medis Institut, Ingolstadter Landstrasse 1, D-8042 Neuherberg, Federal Republic of Germany

AND

JOSEF SCHERER
Universitaet Muenchen, Psychiatrische Klinik, Nussbaumstrasse 7, 8000 Muenchen 2, Federal Republic of Germany

SUMMARY

Directly standardized mortality rates are examples of weighted sums of Poisson rate parameters. If the numbers of events are large then normal approximations can be used to calculate confidence intervals, but these are inadequate if the numbers are small. We present a method for obtaining approximate confidence limits for the weighted sum of Poisson parameters as linear functions of the confidence limits for a single Poisson parameter, the unweighted sum. The location and length of the proposed interval depend on the method used to obtain confidence limits for the single parameter. Therefore several methods for obtaining confidence intervals for a single Poisson parameter are compared. For single parameters and for weighted sums of parameters, simulation suggests that the coverage of the proposed intervals is close to the nominal confidence levels. The method is illustrated using data on rates of myocardial infarction obtained as part of the WHO MONICA Project in Augsburg, Federal Republic of Germany.

INTRODUCTION

Directly standardized mortality rates are weighted sums of age-specific rates. To make statistical inferences about these standardized rates it is usual to assume that the numbers of events observed in each age group are independent and have Poisson distributions (see, for example, Breslow and Day,[1] p. 59). Another example of the need for confidence limits for weighted sums of Poisson parameters occurs in meta-analysis requiring aggregation of outcome rates from several different studies, using weights related to the numbers of subjects in each study. Confidence limits can be calculated assuming that the weighted sum is approximately normally distributed, but this approach may be inadequate if the numbers of events in the age groups are small.

In this paper we begin by comparing several known methods of interval estimation for a single Poisson parameter. Then we present a simple method for obtaining approximate confidence intervals for a weighted sum of Poisson parameters.

0277–6715/91/030457–06$05.00

Received February 1990
Revised July 1990

CONFIDENCE INTERVALS FOR A SINGLE POISSON PARAMETER

Suppose that the random variable X has a Poisson distribution with parameter θ. Desirable properties for a confidence interval $[X_L, X_U]$ for θ given X are that it should be as narrow as possible and have coverage probability as near as possible to the specified confidence level $1 - 2\alpha$.

Crow and Gardner[2] described a method for obtaining confidence intervals which are as narrow as possible and among such narrowest intervals have the smallest possible upper confidence limits. The method involves three steps. First, for any given value of θ, a shortest acceptance region is obtained consisting of non-negative integers from C_1 to C_2 such that, if X is not in the interval $[C_1, C_2]$, then the hypothesis that $E(X) = \theta$ would be rejected at the significance level 2α in favour of the two-sided alternative hypothesis that $E(X) \neq \theta$. This step is performed by systematic consideration of all possible intervals. Secondly, for each acceptance region a set of possible parameter values $[\theta_1, \theta_2]$ is calculated such that $P(C_1 \leqslant X \leqslant C_2 | \theta) \geqslant 1 - 2\alpha$ for any θ in $[\theta_1, \theta_2]$. Finally the confidence interval $[X_L, X_U]$ for θ, given X, is defined as the union of all sets $[\theta_1, \theta_2]$ for which X belongs to the corresponding acceptance regions. For certain confidence levels the limits X_L and X_U were tabulated by Crow and Gardner; alternatively they can be calculated using a computer program (obtainable, for example, from the present authors).

A simpler method is to use relationship between the Poisson and chi-squared distributions (for example, see Johnson and Kotz,[3] p. 96, or Ulm[4]). Confidence limits based on approximately equal probabilities in each tail are given by

$$X_L = \tfrac{1}{2}\chi^2_{2X,\alpha} \quad \text{and} \quad X_U = \tfrac{1}{2}\chi^2_{2(X+1),1-\alpha}. \tag{1}$$

Alternatively there are numerous methods based on normal approximations and equal tail probabilities. These include

$$X \pm u\sqrt{X} \tag{2}$$

$$X + \tfrac{1}{2}u^2 \pm u(X + \tfrac{1}{4}u^2)^{1/2} \tag{3}$$

$$X + (u^2 + 2)/3 \pm u[X + (u^2 + 11)/18]^{1/2} \tag{4}$$

$$X + (2u^2 + 1)/6 \pm \{\tfrac{1}{2} + u[X + (u^2 + 2)/18 \pm \tfrac{1}{2}]^{1/2}\} \tag{5}$$

$$\left.\begin{aligned} X_L &= X[1 - 1/(9X) - u/(3x^{1/2})]^3 \\ X_U &= (X + 1)/\{1 - 1/[9(X + 1)] + u/[3(X + 1)^{1/2}]\}^3 \end{aligned}\right\}, \tag{6}$$

where u is the upper 100α per cent point of the standard normal distribution. The limits (2) and (3) are obtained from the inequalities $-u < (X - \theta)/\sqrt{\theta} < u$ by replacing $\sqrt{\theta}$ in the denominator by $\sqrt{X}$ for (2) or solving for θ for (3) (Johnson and Kotz,[3] p. 96). The limits (4) and (5) are obtained from more complicated approximations.[5] The approximations (6), suggested by Byar, are quoted by Breslow and Day,[1] p. 69. The intervals obtained by the method of Crow and Gardner are shorter than those calculated by the other methods. Among the normal approximations the simplest, (2), fails to reflect the asymmetry of the Poisson distribution, and the others, (3)–(6), all provide close approximations to the limits obtained from the chi-squared values (1).

To illustrate the coverage probabilities of these confidence intervals, simulations were performed for selected values of θ and confidence levels $100(1 - 2\alpha)$ per cent. This was done using the uniform random number generator for the S language (a combination linear-congruential and shift-register generator)[6] and the inverse distribution function for the Poisson distribution. For each simulation an observation X was obtained from a Poisson distribution with parameter θ, and the corresponding confidence limits X_L and X_U were calculated by the method of Crow and Gardner and the chi-squared method. The percentages of intervals not containing θ were

Table I. Confidence limits and empirical coverage probabilities for Poisson parameters (based on simulations of 5000 observations)

Nominal confidence level	Observed value X	Confidence limits (X_{L}, X_{U})		Poisson parameter θ	Percentage of intervals not containing θ			
		Crow and Gardner	Chi-squared		Crow and Gardner		Chi-squared	
					$X_{\mathrm{L}} > \theta$	$X_{\mathrm{U}} < \theta$	$X_{\mathrm{L}} > \theta$	$X_{\mathrm{U}} < \theta$
90	5	2·43, 9·72	1·97, 10·51	5	3·2	3·9	3·3	4·0
	10	5·98, 15·99	5·43, 16·96	10	2·6	6·4	5·3	3·0
	20	13·55, 28·09	13·25, 29·06	20	4·9	3·5	3·8	3·9
	50	38·44, 62·68	38·96, 63·29	50	3·2	6·7	4·7	4·6
95	5	1·97, 11·18	1·62, 11·67	5	3·2	0·7	1·4	0·8
	10	5·32, 17·63	4·80, 18·39	10	1·7	2·7	1·2	1·1
	20	12·82, 30·02	12·22, 30·89	20	2·2	2·1	2·2	2·1
	50	37·67, 64·95	37·11, 65·92	50	2·3	2·5	2·2	1·8

obtained from 5000 simulations. The results are shown in Table I. The coverage probabilities for both methods were larger than the nominal levels but improved as θ increased. They were nearer the nominal levels for the larger value of α. The coverage probabilities were slightly better for the Crow and Gardner intervals than for the chi-squared intervals because the former are narrower; nevertheless, the chi-square method would usually be preferred for simplicity.

These results are relevant to the calculation of confidence intervals for age-specific event rates and also for indirectly standardized ratios in which the numerator is assumed to have a Poisson distribution whose parameter is the denominator of the ratio.[4]

CONFIDENCE INTERVALS FOR WEIGHTED SUMS OF POISSON PARAMETERS

Now we consider the problem of finding a confidence interval for a weighted sum of Poisson parameters taking age-standardized death rates as an example.

Let the random variables X_i, $i = 1, \ldots, k$, represent the numbers of deaths or other outcome events in k subgroups. Suppose that the X_is are independent and that each has a Poisson distribution with expectation $n_i \lambda_i$, where λ_i is the event rate and n_i the number of person-years observed for subgroup i. The parameters λ_i are unknown whereas the n_is are assumed to be known constants.

A directly standardized event rate is defined as

$$Y = \sum_{i=1}^{k} c_i \frac{X_i}{n_i},$$

where the c_is are known constants, for example obtained from the Segi World Standard Population.[7] For simplicity we use the notation

$$Y = \sum w_i X_i,$$

where $w_i = c_i/n_i$. Also $E(X_i) = n_i \lambda_i = \theta_i$. We wish to obtain confidence limits for $E(Y) = \sum w_i \theta_i$.

It is common to use the normal approximation for the Poisson distribution for each X_i and the property that a linear combination of independent normal random variables is also normal so that, approximately, Y has the distribution $\mathrm{N}(\sum w_i \theta_i, \sum w_i^2 \theta_i)$. In practice if the X_is are small this approximation may be poor. The alternative method we propose is based on using $X = \sum X_i$, which has a Poisson distribution with parameter $\theta = \sum \theta_i$.

Table II. Empirical coverage probabilities for weighted sums of Poisson parameters, based on simulations of 5000 observations

Nominal confidence level	$\sum\theta_i$	Parameter $\sum w_i\theta_i$	Percentage of intervals not containing $\sum w_i\theta_i$					
			Crow and Gardner			Chi-squared		
			$T_L > \sum w_i\theta_i$	$T_U < \sum w_i\theta_i$	Total	$T_L > \sum w_i\theta_i$	$T_U < \sum w_i\theta_i$	Total
(a)*								
90	10	1·484	3·2	6·2	9·4	3·4	3·3	6·7
	20	2·968	3·9	5·4	9·3	3·9	3·7	7·6
	50	7·419	4·2	5·6	9·8	4·5	4·5	9·0
95	10	1·484	1·8	2·4	4·2	1·7	1·5	3·2
	20	2·968	1·8	2·9	4·7	1·4	2·0	3·4
	50	7·419	1·8	2·7	4·5	2·1	2·2	4·3
(b)*								
90	10	5·4	3·1	7·4	10·5	3·3	6·7	10·0
	50	27·0	3·6	5·8	9·4	3·7	5·5	9·2
95	10	5·4	1·8	4·6	6·4	1·2	5·1	6·3
	50	27·0	1·5	3·5	5·0	1·9	2·4	4·3

* For (a) the weights w_i were all fairly similar and the θ_is increased with i, as would be common for many chronic diseases. For (b) the weights w_i increased with i but the θ_is were all equal.

Let $T = a + bX$, where a and b are chosen so that $E(T) = E(Y)$ so that $a + b\sum\theta_i = \sum w_i\theta_i$, and $\operatorname{var}(T) = \operatorname{var}(Y)$ so that $b^2\sum\theta_i = \sum w_i^2\theta_i$. Hence

$$a = \sum w_i\theta_i - [\sum\theta_i\sum w_i^2\theta_i]^{1/2} \quad \text{and} \quad b = [\sum w_i^2\theta_i/\sum\theta_i]^{1/2}.$$

As T is a monotonically increasing function of X if (X_L, X_U) is a $100(1-2\alpha)$ per cent confidence interval for θ, then $T_L = a + bX_L$ and $T_U = a + bX_U$ would be the required confidence limits for $\sum w_i\theta_i$. However, as a and b depend on the unknown parameters θ_i they have to be replaced by estimators, for example

$$\hat{a} = \sum w_iX_i - [\sum X_i\sum w_i^2X_i]^{1/2} = Y - (XV)^{1/2}$$
$$\hat{b} = [\sum w_i^2X_i/\sum X_i]^{1/2} = (V/X)^{1/2},$$

where $V = \sum w_i^2X_i$. Thus approximate confidence limits for $\sum w_i\theta_i$ are given by

$$T_L = Y + (V/X)^{1/2}(X_L - X) \quad \text{and} \quad T_U = Y + (V/X)^{1/2}(X_U - X). \tag{7}$$

The results for confidence intervals for a single Poisson parameter provide a background against which to assess the performance of the interval (T_L, T_U), defined in (7), for a weighted sum $\sum w_i\theta_i$ of Poisson parameters.

Table II(a) shows simulation results for three cases involving small values of θ_i such as might apply for death rates per 10,000 population. Suppose there are 6 age groups each of n_i = 10,000 people, and the values c_i are {6/31, 6/31, 6/31, 5/31, 4/31, 4/31} corresponding to weights for the Segi World Standard Population[7] for the 5-year age groups from 35–39 to 60–64 years. Let the event rates λ_i be {1, 1, 2, 4, 5, 7} per 10,000. As $\theta_i = n_i\lambda_i$ and $w_i = c_i/n_i$, $\sum w_i\theta_i = \sum c_i\lambda_i$ = 2·968 and $\theta = \sum\theta_i = 20$; this is the second case. For the first case the λ_is are halved so that $\sum w_i\theta_i$ = 1·484 and $\theta = \sum\theta_i = 10$. For the third case the λ_is are multiplied by 2·5 so that $\sum w_i\theta_i$ = 7·419 and $\theta = \sum\theta_i = 50$. An observation X_i was simulated for each age group and a confidence interval (T_L, T_U) for $\sum w_i\theta_i$ was calculated using (7) with values of X_L and X_U obtained from $X = \sum X_i$ either by the method of Crow and Gardner or by the chi-squared method. This

Table III. Incidence rates for myocardial infarction in women in Augsburg in 1986 by age and reporting unit

	Reporting unit 1		Reporting unit 2	
Age (years)	Person-years, n_i	Events, X_i	Person-years, n_i	Events, X_i
35–39	7,971	0	10,276	0
40–44	7,084	0	9,365	1
45–49	9,291	1	11,623	0
50–54	7,743	2	8,684	4
55–59	7,798	4	7,926	0
60–64	8,809	10	8,375	3
Total, X		17		8
Age standardized rate per 10,000, Y		2·75		1·41
95% CI from crude normal approximation		1·43, 4·07		0·43, 2·40
95% CI from (7) and Crow and Gardner method		1·54, 4·27		0·58, 2·64
95% CI from (7) and chi-squared method		1·59, 4·42		0·61, 2·79

process was repeated 5000 times and coverage probabilities estimated. Comparison with the corresponding results for $\theta = 10$, 20 and 50 in Table I shows that the coverage probabilities are similar and slightly higher than the nominal levels (different simulations were used for Tables I and II). The Crow and Gardner intervals appear to have slightly better coverage probabilities.

Table II(b) illustrates the effects of using a wider range of w_i values $\{0{\cdot}1, 0{\cdot}1, 0{\cdot}5, 1{\cdot}0, 1{\cdot}0\}$ and differing θ_is; if $\theta_i = 2$ for $i = 1, \ldots, 5$ then $\sum w_i\theta_i = 5{\cdot}4$ and $\sum\theta_i = 10$, and if $\theta_i = 10$ for $i = 1, \ldots, 5$ then $\sum w_i\theta_i = 27$ and $\sum\theta_i = 50$. The coverage probabilities obtained were, again, close to the nominal levels and there was little difference between the two methods.

EXAMPLE

The proposed method for calculating confidence intervals for weighted sums of Poisson parameters is illustrated by an example of age-standardized rates. The data are from the WHO MONICA Project.[8] The study area in Augsburg, Federal Republic of Germany, comprises an urban area (reporting unit 1) and a rural area (reporting unit 2).[9] Since October 1984 all suspected coronary events occurring in residents of the study area have been registered. Table III gives the 1986 incidence rates for non-fatal definite myocardial infarction (according to the MONICA project criteria) in women aged 35–64 years stratified by 5-year age group and reporting unit. The weights c_i are $\{6/31, 6/31, 6/31, 5/31, 4/31, 4/31\}$ corresponding to the Segi World Standard Population.[7]

The age-standardized incidence rates per 10,000 person-years, $\sum w_i\theta_i$, estimated by $Y = \sum w_i X_i$ with $w_i = c_i/n_i$, were 2·75 and 1·41 respectively for reporting units 1 and 2. The confidence intervals using the crude normal approximation $Y \pm u\sqrt{V}$ (as suggested by Breslow and Day[1]) were (1·43, 4·07) and (0·43, 2·40). By contrast, the intervals obtained using (7) with values obtained by the method of Crow and Gardner or the chi-squared method are wider and asymmetric. These results illustrate that the conventional method may be inaccurate for small numbers of events.

REFERENCES

1. Breslow, N. E. and Day, N. E. *Statistical Methods in Cancer Research Vol. II*, International Agency for Research on Cancer, Lyon, 1987.
2. Crow, E. L. and Gardner, R. S. 'Confidence intervals for the expectation of a Poisson variable', *Biometrika*, **46**, 441–453 (1959).
3. Johnson, N. L. and Kotz, S. *Distributions in Statistics: Discrete Distributions*, Houghton Mifflin, Boston, 1969.
4. Ulm, K. 'A simple method to calculate the confidence interval of a standardized mortality ratio (SMR)', *American Journal of Epidemiology*, **131**, 373–375 (1990).
5. Molenaar, W. *Approximations to the Poisson, Binomial and Hypergeometric Distribution Functions*, Mathematical Centre Tracts 31, Mathematisch Centrum, Amsterdam, 1970.
6. Becker, R. A., Chambers J. M. and Wilks, A. R. *The New S Language: A Programming Environment for Data Analysis and Graphics*, Wadsworth & Brooks/Cole, Palm Grove, California, 1988.
7. Waterhouse, J., Muir, C., Correo, P. and Powell, J. (eds) *Cancer Incidence in Five Continents, Vol II*, IARC Scientific Publications no. 15, International Agency for Research on Cancer, Lyon, 1976.
8. WHO MONICA Project Principal Investigators. 'The World Health Organization MONICA Project (Monitoring Trends and Determinants in Cardiovascular Disease): a major international collaboration', *Journal of Clinical Epidemiology*, **2**, 105–114 (1988).
9. Keil, U. *et al.* MONICA Project. *Region Augsburg: Manual of Operations – Myocardial Infarction Register*, Bericht 21, Gesellschaft fuer Strahlen-und Umweltforschung, Muenchen, 1985.