A test for seasonality of events with a variable population at risk*

S. D. WALTER AND J. M. ELWOOD

Department of Epidemiology and Community Medicine, University of Ottawa, Canada

Walter, S. D. and Elwood, J. M. (1975). *British Journal of Preventive and Social Medicine*, **29**, 18-21. **A test for seasonality of events with a variable population at risk.** A statistical significance test to detect seasonality of epidemiological events is described. The method is similar to that of Edwards, but makes it possible to allow for an arbitrary pattern of variation in the population at risk, and also for the unequal lengths of time sectors of a cycle of seasons (e.g., months of a year). From the test it is possible to estimate the amplitude of seasonal variation and the time at which the maximum occurs in a postulated simple harmonic fluctuation; the adequacy of the description of the data by a curve of this kind may be evaluated using a goodness-of-fit test. A numerical example of the calculations is given using some anencephalus data, and the results are compared with those of alternative tests.

The description of incidence rates in terms of seasonal variation or cyclic trends is an important objective in many epidemiological problems, and the use of statistical tests to detect such variation is widespread; in particular, the test proposed by Edwards (1961) has occupied a central place in analyses of this kind, having been applied to a wide variety of epidemiological data, and also in other disciplines. The test has been criticized by Wehrung and Hay (1970) for being sensitive to occasional extreme values in the data, and a non-parametric alternative test has been suggested (Hewitt *et al.*, 1971) when the sample size is small. A further limitation of Edwards' test is that it does not take account of the size of the population at risk, when it is usually seasonality of rates rather than pure numbers which is of interest to epidemiologists. Examples of problems where the variable population at risk is a consideration are provided by the analysis of congenital malformations with a seasonably variable number of total births, or the incidence of postoperative deaths for an elective surgical procedure demonstrating seasonal variation because of hospital administrative arrangements. Making allowances for this factor is important in a statistical evaluation of seasonality as may be seen by two extreme examples: first, supposing that a cyclic trend of disease events was in phase with the trend for the population at risk (i.e., the maxima occur at approximately the same time), then a significant cyclic trend estimated solely from the numbers of events might have been spuriously caused by the variation in population; second, if the two trends were completely out of phase, a true cyclic variation in rates may be masked by the cancelling effect of population.

Although Edwards suggests the use of correction factors to adjust his test for variable populations at risk, it is proposed here to extend the method somewhat more rigorously. A further generalization will be in the use of time sectors (e.g., months of the year) which are not necessarily equal, as is required by Edwards' formulation. An illustrative example is given using some data on anencephalus.

METHODOLOGY

We suppose that within a certain time span (e.g., a year) there are k sectors (e.g., 12 months), and that in sector i there are n_i events (e.g., congenital malformations) from a population at risk of size m_i (e.g., total births); the total number of events is $N = \Sigma n_i$ and the total population at risk is $M = \Sigma m_i$. Following Edwards, we represent
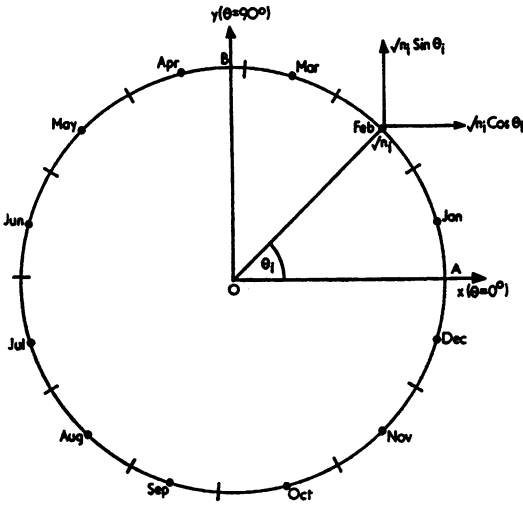
FIGURE. Typical components of moments about orthogonal axes OA and OB.

the data by weights $\sqrt{n_i}$ placed around a unit circle at points corresponding to the sector mid-points at angles θ_i to an arbitrary diameter (e.g., OA, the diameter through 1 January; see Figure). The moment about this diameter of all the weights is $\sum_i \sqrt{n_i} \sin \theta_i$, and also considering the moment about the orthogonal diameter, the sample centre of gravity of the system may be seen to be $(\bar{x}, \bar{y})$ where

$$\bar{x} = \sum_{i=1}^{k} \sqrt{n_i} \cos \theta_i / W;$$

$$\bar{y} = \sum_{i=1}^{k} \sqrt{n_i} \sin \theta_i / W \qquad (1)$$

and where $W = \sum \sqrt{n_i}$, the total weight. On the null hypothesis, the expected number of events in a sector is proportional to the population at risk in that sector, i.e., $E(n_i) = Nm_i/M$. We will suppose that the sample is large enough that to calculate the expected values of $\bar{x}$ and $\bar{y}$, we may replace n_i by its expectation everywhere in (1) and obtain

$$E(\bar{x}) = \mu_{\bar{x}} = \sum_i \sqrt{(Nm_i/M)} \cos \theta_i / \sum_i \sqrt{(Nm_i/M)}$$

$$= \sum_i \sqrt{m_i} \cos \theta_i / \sum_i \sqrt{m_i}$$

and similarly

$$E(\bar{y}) = \mu_{\bar{y}} = \sum_i \sqrt{m_i} \sin \theta_i / \sum_i \sqrt{m_i}$$

If we further suppose that the variance of $\sqrt{n_i}$ is approximately $\frac{1}{4}$, and again replace n_i by its expected value in W, we obtain from (1) that the variances of the coordinates of the sample centre of gravity are

$$\operatorname{Var}(\bar{x}) = \sigma_{\bar{x}}^2 = \sum_i \tfrac{1}{4} \cos^2 \theta_i / [\sum \sqrt{(Nm_i/M)}]^2$$

and

$$\operatorname{Var}(\bar{y}) = \sigma_{\bar{y}}^2 = \sum_i \tfrac{1}{4} \sin^2 \theta_i / [\sum \sqrt{(Nm_i/M)}]^2$$

On the assumption that $\bar{x}$ and $\bar{y}$ are normally distributed, we may take as a test statistic the quantity

$$\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \right)^2 + \left(\frac{\bar{y} - \mu_{\bar{y}}}{\sigma_{\bar{y}}} \right)^2 \qquad (2)$$

which on the null hypothesis is distributed as χ^2 with 2 d.f. The distance d of the sample centre of gravity from its null expectation is given by

$$d^2 = (\bar{x} - \mu_{\bar{x}})^2 + (\bar{y} - \mu_{\bar{y}})^2 \qquad (3)$$

If it is required to fit a simple harmonic trend to the data, we may suppose that the expected frequency in sector i is proportional to

$$c_i = m_i [1 + \alpha \cos(\theta_i - \theta^*)];$$

α is a measure of the amplitude of the cyclic variation, and θ^* gives the direction of the maximum rate, which may be estimated as

$$\theta^* = \operatorname{Tan}^{-1} \left(\frac{\bar{y} - \mu_{\bar{y}}}{\bar{x} - \mu_{\bar{x}}} \right) \qquad (4)$$

If we place equal weights of w/k (where w is a constant) at angles θ_i and radii

$$r_i = \sqrt{\{ m_i k [1 + \alpha \cos(\theta_i - \theta^*)]/M \}}$$

we may derive

$$wd = \sum_i r_i \cos(\theta_i - \theta^*) \, w/k \qquad (5)$$

If the r_i's are expanded as far as terms in α (which is presumed small), and equation (5) is rearranged, we may estimate α as

$$\alpha = \frac{2[d\sqrt{(kM)} - \sum \sqrt{m_i} \cos(\theta_i - \theta^*)]}{\sum \sqrt{m_i} \cos^2(\theta_i - \theta^*)} \qquad (6)$$

In the case $\theta_i = 2\pi i/k$, $m_i = M/k$, $i = 1, 2, \ldots k$, it may be shown that $\mu_{\bar{x}} = \mu_{\bar{y}} = 0$, and that (2), (4) and (6) reduce to the corresponding statistics of Edwards' test, which is a special case of this more general formulation. The argument of Smith (1961) concerning the non-independence of the n_i in Edwards' test may be extended to this new method. When the sample size is small, the test of Hewitt *et al.* may be a preferable alternative, using ranked rates rather than numbers of observations to take account of population size seasonality, subject to the requirement that at least six months have a non-zero number of observations. However, if the populations at risk vary substantially, the implied heteroscedasticity of the associated rates may affect the null distribution of the statistic used in Hewitt's test; further simulation to establish the null distribution would be required for specific problems of this type.

The adequacy of the description of the data afforded by a simple harmonic curve may be evaluated by a goodness-of-fit test using a further χ^2 statistic. If expected frequencies $n_i{'}$ are calculated by

$$n_i{'} = N c_i/\Sigma c_i.$$

then under the usual assumptions, the statistic

$$\Sigma (n_i - n_i{'})^2 /n_i{'} \qquad (7)$$

has a χ^2 distribution with $(k - 1)$ d.f. Wehrung and Hay (1970) describe a hypothetical example where the event frequency is constant over all months except one, where a threefold increase is observed, and Edwards' test shows a significant departure

from the null hypothesis; a χ^2 test using (7) shows that these data are quite inappropriately described by a simple harmonic curve, and we therefore do not feel that counter-examples of this kind constitute a substantial objection to the use of this test.

NUMERICAL EXAMPLE

In Table I are given monthly frequencies of cases of anencephalus and total births for Canada in the period 1954-62 (Elwood, 1975). A cursory inspection of the data reveals a distinct general excess of total births during the summer months, whereas the anencephalus cases demonstrate no consistent seasonal pattern.

The various statistical tests discussed in the previous section were applied to these data, and the results are given in Table II. Edwards' method produces a test statistic of $0\cdot80$, which is certainly not significant assuming a χ^2 distribution on 2 d.f.; the estimated maximum incidence rate occurs at $\theta^* = -18\cdot8°$, i.e., in mid-December. Results from the test proposed above are given as tests B and C of Table II, the former assuming months to be of equal length, and the latter using exact month lengths; the parameters d, θ^* and α of the simple harmonic curve were estimated from equations (3), (4), and (6) respectively. The use of exact θ_i values made very little difference to the conclusions as compared with test B with equally spaced θ_i values; the χ^2 statistic of test B was $12\cdot48$, indicating a deviation from the null hypothesis of constant rates significant at the $0\cdot5\%$ level, with an estimated maximum occurring at $\theta^* = -7\cdot4°$, i.e., in late December. For simplicity, in test C the length of February was taken as

TABLE I

ANENCEPHALUS AND TOTAL BIRTH FREQUENCIES BY MONTH, CANADA, 1954-62

Month	Total Births (m_i)	Anencephalus Cases (n_i)	Expected Frequencies		Monthly Ranking of Rates
			(i)	(ii)	
January	340,797	468	430·9	434·6	12
February	318,319	399	427·9	398·7	8
March	363,626	471	423·9	443·2	10
April	359,689	437	420·3	424·9	6
May	373,878	376	418·0	430·4	1
June	361,290	410	417·5	410·6	3
July	368,867	399	419·1	420·9	2
August	358,531	472	422·2	417·4	11
September ..	363,551	418	426·1	436·0	4
October	352,173	448	429·8	435·5	9
November ..	331,964	409	432·1	420·4	7
December	336,894	397	432·5	431·4	5
Total	4,229,579	5,104	5,103·9	5,104·0	

(i) Frequencies proportional to $1 + \alpha \cos (\theta_i - \theta^*)$, with $\alpha = 0\cdot01766$, $\theta^* = -18\cdot8°$, $\theta_i + \pi/12 = 2\pi i/12$ (Table II, test A)
    χ^2 Goodness-of-fit statistic (7) $= 27\cdot24$ (P $< 0\cdot005$)
(ii) Frequencies proportional to $m_i[1 + \alpha \cos (\theta_i - \theta^*)]$, with $\alpha = 0\cdot06014$, $\theta^* = -8\cdot1°$, θ_i's exact (Table II, test C)
    χ^2 Goodness-of-fit statistic (7) $= 23\cdot97$ (P $< 0\cdot05$)

<div align="center">TABLE II</div>

RESULTS OF VARIOUS SIGNIFICANCES TESTS FOR SIMPLE HARMONIC TREND IN ANENCEPHALUS INCIDENCE

Test	Estimated Parameters			Value of	
	d	α	θ^*	Test Statistic	Significance
A Edwards (1961), using case frequencies	0·004415	0·01766	−18·8°	0·80[1]	$P < 0·7$
B Walter and Elwood (1975), with $\theta_i + 2\pi/24 = 2\pi i/12$	0·01748	0·07048	−7·4°	12·48[2]	$P < 0·005$
C Walter and Elwood (1975), with exact θ_i	0·01746	0·06014	−8·1°	12·47[2]	$P < 0·005$
D Edwards (1961), using adjusted frequencies	0·01733	0·06933	−6·9°	12·27[1]	$P < 0·005$
E Hewitt et al.(1971) ..	—	—	—	51[3]	$P < 0·3$

[1] Statistic = $\frac{1}{2}\alpha^2 N$ ⎫ having null distribution of χ^2 with 2 d.f.
[2] Statistic given by equation (2) ⎭
[3] Largest rank sum of rates for any six-month segment of year; value 51 corresponds to period October-March.

28¼ days and a total year length of 365¼ days. To be even more correct we should have taken February to have a length of 28 + 2/9 days with a similar adjustment to the year length to take account of the exact number of leap years occurring in the time period under study; in practice, this refinement makes so little difference to the conclusions that it was decided not to incorporate this additional complexity into the general computer program used for these analyses.

Expected frequencies were calculated for the best simple harmonic curves given by tests A and C (see Table I), and these were then compared with the observed frequencies using the χ^2 goodness-of-fit test given by equation (7). Although both χ^2 statistics indicate significant departures from the fitted curves, that for Edwards' curve is somewhat larger, showing a rather poorer description of the data than that given by the test proposed here. A decomposition of the statistic corresponding to test C shows that over half the total of 23·97 comes from a large deficit of anencephalics in May and a large excess in August, but that otherwise the observed and expected frequencies are reasonably close.

Several previous authors have employed correction factors when using Edwards' test. One such possibility in the present example would be to replace the n_i by $n_i' = c\, n_i\, M/km_i$, where c is a scale factor such that $\Sigma n_i' = N$; the results of such a test are given as test D in Table II. Although the results are very similar to those of tests B and C, we feel that the latter are preferable as they have a less heuristic or empirical basis; it is quite possible that with smaller samples, the differences between the results of tests C and D would be somewhat greater than in the present example.

The non-parametric test of Hewitt *et al.* (1971) applied to these data gives a maximum rank sum for the monthly rates as 51 for the six-month period October to March; this value is not significant (test E, Table II). With a substantial amount of data, such as is available here, this test will probably have low power to detect a seasonal trend, unless it is fairly marked (i.e., has a large amplitude) and consistent. Further, this test does not allow one to estimate parameters of the simple harmonic curve even if a significant departure from the null hypothesis is indicated. On the other hand, the argument concerning the differences between the distributions of χ^2 and the Edwards' test statistic may also apply to the test proposed here; thus when the sample size is small (say $N < 50$), caution should be exercised in the interpretation of the test results, and a non-parametric approach may be a preferable alternative.

REFERENCES

EDWARDS, J. H. (1961). The recognition and estimation of cyclic trends. *Ann. hum. Genet.*, 25, 83.

ELWOOD, J. M. (1975). Seasonal variation in anencephalus in Canada. *Brit. J. prev. soc. Med.*, 29, 22.

HEWITT, D., MILNER, J., CSIMA, A., and PAKULA, A. (1971). On Edwards' criterion of seasonality and a non-parametric alternative. *Brit. J. prev. soc. Med.*, 25, 174.

SMITH, C. A. B. (1961). Note on the error variance. *Ann. hum. Genet.*, 25, 86.

WEHRUNG, D. A. and HAY, S. (1970). A study of seasonal incidence of congenital malformations in the United States. *Brit. J. prev. soc. Med.*, 24, 24.