ON CRITERIA OF SEASONALITY

By

M. Rockenbauer

Department of Mathematics of the Faculty of Electrical Engineering,
Technical University, Budapest

Received June 20, 1978
Presented by Prof. Dr. O. Kis

Introduction

Presently there is an increased interest in investigations concerned with environmental variables and their associations with diseases. This interest is generated by an awareness of the possibility that our environment can be contaminated by nonbiologic, potentially teratogenic agents. Thus, there may be a relationship between the season of birth and the incidence of various types of disease. The prevention of these diseases depends on the detection and controllability of causes. To this aim, however, it is necessary to improve our methods of analysing seasonal data.

The monthly distribution of births of infants with different malformations has been studied by some statistical methods: for example, χ^2 test for heterogeneity between months, Edwards' method [1] and Hewitt's [2] method. This can be done either on the uncorrected material and on the material corrected for the monthly fluctuation of birth numbers.

Our investigations [3], [4] concerned the relationship of season of birth to prenatal pathology and congenital malformations. Improvements have been applied to the following three methods:

A) χ^2 *test*. If the number of births of infants increases, so does clearly the number of births of infants with malformations. But a simple χ^2 test with the transformation

$$y_i \rightarrow y_i/b_i \qquad (i = 1, 2, \ldots 12)$$

where y_i monthly number of births of patients

b_i monthly number of births

is not significant because the value of y_i/b_i is much too small. Therefore the following transformation has been introduced:

$$y_i \rightarrow y_i \cdot \bar{b}/b_i \qquad (i = 1, 2, \ldots 12)$$

where $\bar{b} = \left(\sum_{i=1}^{12} y_i\right)\Big/12$.

B) *Edwards' method* [1]:

The method consists in fitting a sinus curve to the recorded differences between found and expected numbers of malformed infants each calendar month.

$$y_i \approx \bar{y}(1 + \alpha \sin(x_i + \varphi)) \quad (i = 1, 2, \ldots 12)$$

where

$$\bar{y} = \left(\sum_{i=1}^{12} y_i\right) \Big/ 12$$

$$x_i = \frac{\pi}{12} + (i-1) \cdot \frac{\pi}{6}\,.$$

Edwards' method is likely to display a simple harmonic pattern (assuming the peak and the trough of an annual cycle to be separated by approximately six months).

According to Wehrung and Hay [5] "$0 \leq \varphi \leq 2\pi$ is the angle corresponding to the data of maximum incidence on the fitted curve" this estimation is very rough, therefore the least squares' method have been used. (Maximum likelihood is not suitable to this purpose.) The following estimates of α and φ can be made.

For $\quad m_i = y_i/\bar{y} - 1; \qquad (i = 1, 2, \ldots 12)$

$$U = \sum_{i=1}^{12} \bar{y}^2 (m_i - \alpha \sin(x_i + \varphi))^2$$

$$\frac{\partial U}{\partial \alpha} = -2\bar{y}^2 \sum_{i=1}^{12} m_i \sin(x_i + \varphi) + 2\bar{y}^2 \alpha \sum_{i=1}^{12} \sin^2(x_i + \varphi) =$$

$$= 2\bar{y}^2 \left(-\sum_{i=1}^{12} m_i \sin(x_i + \varphi) + 6\alpha\right).$$

For $\quad \dfrac{\partial U}{\partial \alpha} = 0\,;$

$$\alpha = \frac{\sum_{i=1}^{12} m_i \sin(x_i + \varphi)}{6} = \frac{\sum_{i=1}^{12} y_i \sin(x_i + \varphi)}{6\bar{y}}$$

$$\frac{\partial U}{\partial \varphi} = -2\alpha\bar{y}^2 \sum_{i=1}^{12} m_i \cos(x_i + \varphi) + \alpha^2 \bar{y}^2 \sum_{i=1}^{12} \sin(2x_i + 2\varphi) =$$

$$= -2\alpha\bar{y}^2 \sum_{i=1}^{12} (y_i/\bar{y} - 1) \cos(x_i + \varphi) =$$

$$= -2\alpha\bar{y} \sum_{i=1}^{12} y_i \cos(x_i + \varphi)\,.$$

For $\dfrac{\partial U}{\partial \varphi} = 0$;

$$\sum_{i=1}^{12} y_i \cos (x_i + \varphi) = 0,$$

consequently

$$\sum_{i=1}^{12} y_i \cos x_i \cos \varphi = \sum_{i=1}^{12} y_i \sin x_i \sin \varphi,$$

thus

$$\operatorname{tg} \varphi = \frac{\sum_{i=1}^{12} y_i \cos x_i}{\sum_{i=1}^{12} y_i \sin x_i}$$

$$\frac{\partial^2 U}{\partial \alpha^2} = 12\bar{y}^2$$

$$\frac{\partial^2 U}{\partial \varphi^2} = 2\alpha\bar{y} \sum_{i=1}^{12} y_i \sin (x_i + \varphi) = \frac{1}{3} \left(\sum_{i=1}^{12} y_i \sin (x_i + \varphi) \right)^2$$

because

$$\alpha = \frac{1}{6\bar{y}} \left(\sum_{i=1}^{12} y_i \sin (x_i + \varphi) \right)$$

$$\frac{\partial^2 U}{\partial \alpha \, \partial \varphi} = - 2\bar{y} \sum_{i=1}^{12} y_i \cos (x_i + \varphi) = 0$$

therefore

$$\begin{vmatrix} \dfrac{\partial^2 U}{\partial \alpha^2} & \dfrac{\partial^2 U}{\partial \alpha \, \partial \varphi} \\ \dfrac{\partial^2 U}{\partial \varphi \, \partial \alpha} & \dfrac{\partial^2 U}{\partial \varphi^2} \end{vmatrix} = 4 \bar{y}^2 \left(\sum_{i=1}^{12} y_i \sin (x_i + \varphi) \right)^2 > 0 \text{ if } \alpha \neq 0.$$

C) *Hewitt's method* [2]. The rank sum method also detects simple harmonic patterns but it is more critical to meandering and it can distinguish between harmonic and periodic (non-harmonic) variation [6]. It was decided to try a criterion based on the sum of the ranks of r successive months. The monthly frequencies (or preferably the monthly incidence rates) have been ranked from 12 (highest) down to 1 (lowest). On the ground that the chance probability of obtaining the largest possible rank sum is the smallest for $r = n/2$ it was chosen to use the rank sum for six successive months. But there is a mistake in [2], namely there are only 462 possible combinations of ranks, rather than 924, since if six successive months are chosen the rest of months are given. Therefore a value of the test criterion equal to or greater than 52 (instead of 50) would be regarded as significant at the conventional 5% level. It is necessary to locate the six-month segment which yields the highest value of the rank-sum. The zero distribution of this maximum rank-sum was stated [2] not to be easy to formulate or to enumerate, even on a high-speed computer, but it is a mistake.

We needed only to enumerate the value of the rank-sum for the following six-month segments: from 1 to 6, from 2 to 7, from 3 to 8, from 4 to 9, from 5 to 10 and from 6 to 11. At the same time the rest of segments are given. The correct distribution is seen in Table 1.

Table 1

HEWITT'S RANK-SUM CRITERION OF SEASONALITY

NUMBERS		EXACT DISTRIBUTION	CUMULATIVE PROBABILITY	PROBABILITY
21	57	1	+.2164502165/−02	+.2164502165/−02
22	56	1	+.4329004329/−02	+.2164502165/−02
23	55	2	+.8658008658/−02	+.4329004329/−02
24	54	3	+.1515151515/−01	+.6493506494/−02
25	53	5	+.2597402597/−01	+.1082251082/−01
26	52	7	+.4112554113/−01	+.1515151515/−01
27	51	11	+.6493506494/−01	+.2380952381/−01
28	50	13	+.9307359307/−01	+.2813852814/−01
29	49	18	+.1320346320/−00	+.3896103896/−01
30	48	22	+.1796536797/+00	+.4761904762/−01
31	47	28	+.2402597403/+00	+.6060606061/−01
32	46	32	+.3095238095/−00	+.6926406926/−01
33	45	39	+.3939393939/−00	+.8441558442/−01
34	44	42	+.4848484848/+00	+.9090909091/−01
35	43	48	+.5887445887/−00	+.1038961039/+00
36	42	51	+.6991341991/−00	+.1103896104/+00
37	41	55	+.8181818182/−00	+.1190476190/−00
38	40	55	+.9372294372/−00	+.1190476190/+00
39	39	29	+.1000000000/+01	+.6277056277/−01

Remark: It is easy to prove that the probability of the event that the minimum or the maximum value of rank-sum is k ($k = 21, 22, \ldots, 39$) and 78-k respectively, for a pre-assigned six-month segment of the year equals the probability of the event that the minimum or maximum value of the rank-sum is k and 78-k respectively, for at least one of the six-month segments of the year.

Summary

An improvement is suggested for three statistical methods applied previously in the same problem.

First, a transformation is introduced, to take the monthly fluctuation of births into consideration.

Second, a statistical estimation applying the least squares' method is given for the Edwards' model.

Third, a mistake in Hewitt's paper is corrected and a very simple enumeration is given for the maximum rank-sum.

References

1. EDWARDS, J. H.: Seasonal incidence of congenital disease in Birmingham. Ann. Hum. Genet., Lond. (1961), 25, 89.
2. HEWITT, D.—MILLNER, J.—CSIMA, A.—PAKULA, A.: On Edwards' criterion of seasonality and a non-parametic alternative. Brit. J. prev. soc. Med. (1971), 25, 147—176.
3. ROCKENBAUER, M.—CZEIZEL, E.: Investigation of the monthly incidence, the so-called seasonality of the births of persons with common congenital malformations.* Demográfia (1978), 1—2. 193—222.
4. ROCKENBAUER, M.—CZEIZEL, E.: Seasonality tests of prenatal mortality.* Gyermekgyógyászat, (1978) 29, 300—308.
5. WEHRUNG, D. A.—HAY, S.: A study of seasonal incidence of congenital malformations in the United States. Brit. J. prev. soc. Med. (1970), 24, 24. 32.
6. ROBERTS, C. J.—LOWE, C. R.—LLOYD, S.: Cyclic variations in date of last menstrual period of mothers of infants with congenital malformations in South Wales, 1964—66. Brit. J. prev. soc. Med. (1972), 26, 212—218.
7. SANDAHL, B.: Seasonal incidence of some congenital malformations in the central nervous sistem in Sweden, 1965—1972. Acta Paediatr. Scand. (1977), 66, 65—72.

Magda ROCKENBAUER H-1521 Budapest

* In Hungarian