

Studies on the Time Series Analysis of the Seasonal Disease

by

M. Momiyama.

Meteorological Research Institute

(Received June 5, 1950)

Abstract

The writer analysed the seasonal disease, beriberi, by using the means of auto-correlation method and applying the cyclic distribution. The purpose of this study is to find out the type of the seasonal periodicity and the distribution curve of beriberi in the main cities of Japan.

1) Analysis by means of auto-correlation method :

I calculated the auto-correlation coefficient of the death rate of beriberi in 5 years from 1926 to 1930. Then I draw correlogramm by using auto-correlation coefficient.

There is the clear period of 12 months which has wide amplitude in four cities in the Kansai district and two cities in the Kanto district. While in the cities in the south Kyushu district and in the Hokkaido district, there exist two periods, one is the period of 12 months which has wide amplitude and the other is the small period of 6 months even though they are not so clear.

2) Means of applying the cyclic distribution :

I tried to express the death curve of beriberi by composition of two periods which have one year normal cyclic distribution. But the phase of the second period slips down by half a year. The first period was supposed to have the cause of outbreak in summer while the second period in winter.

The temporal distribution curve is expressed by

$$f(t) = c_1 \sum_{-\infty}^{\infty} G_1(t-nh) + c_2 \sum_{-\infty}^{\infty} G_2(t-nk)$$

In the case of Osaka City and Fukuoka City, I calculated the value of constants. Next if I calculate, by using the constants, the monthly death rate of beriberi in these two cities as the actual value, I can get the distribution curve shown in the graph, and find out that the actual measurement and calculation coincide well.

Judging from the above-mentioned fact, I can imagine that there will be some reason easy to cause beriberi death even in winter in some district.

The writer tried the time series analysis of the seasonal disease, beriberi to find out the type of the seasonal periodicity and the distribution curve of beriberi by using the means of auto-correlation method and applying the cyclic distribution.

1. Analysis by means of auto-correlation method

At first I arranged the death rates of beriberi in 5 years from 1926 to 1930 in the order of time series and then calculated the auto-correlation coefficient. The formula of calculation of auto-correlation coefficient is as follows.

$$r_k = \frac{\sum_{i=1}^{N-k} x_i x_{i+k} - (N-k) \bar{x}_1 \bar{x}_2}{(N-k) s_1 s_2}$$

$$\bar{x}_1 = \frac{1}{N-k} \sum_1^{N-k} x_i, \quad \bar{x}_2 = \frac{1}{N-k} \sum_{k+1}^N x_i$$

$$s_1^2 = \frac{1}{N-k} \sum_1^{N-k} (x_i - \bar{x}_1)^2, \quad s_2^2 = \frac{1}{N-k} \sum_{k+1}^N (x_i - \bar{x}_2)^2$$

Next if I draw correlogramm by using auto-correlation coefficient, it is as the following graph. (fig.1)

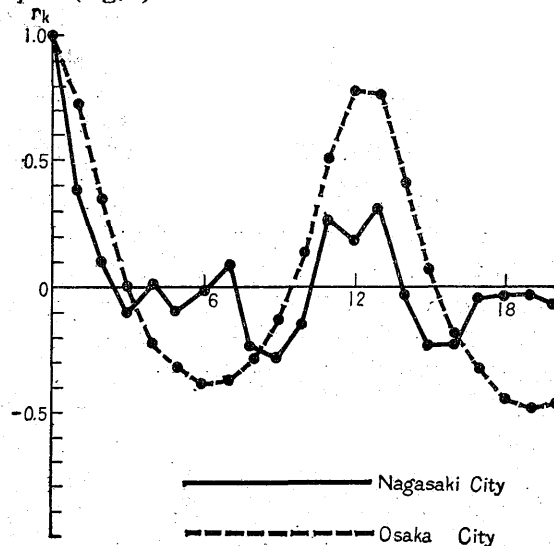


Fig. 1 Correlogramm of the death rate of beriberi.

There is the clear period of 12 months which has wide amplitude in four cities of Kyoto, Osaka, Kobe and Nagoya in the Kansai district and two cities of Tokyo and Yokohama in the Kanto district. There exists the clear period of 12 months and the other very little period of 6 months in Okayama City in the Chugoku district and two cities of Fukuoka and Yahata in the North Kyushu district. And in Nagasaki and Kagoshima in the South Kyushu district and in two cities of Sapporo and Hakodate in the Hokkaido district two periods lie one upon another, one is the period of 12 months which has wide amplitude and the other is the small period of 6 months even though they are not so clear. In Sen-

dai City in the Tohoku district, the type is irregular and no period is perceived. Thus the clear period of 12 months exists only in the cities of the Kansai and Kanto districts and in the rest of the cities except Sendai there exist two periods of 12 and 6 months.

2. Means of applying the cyclic distribution

I analysed by what kind of distribution function the distribution curve of the death rate of beriberi was made. Dr. Masuyama pointed out before that the distribution curve of the death rate of typhoid fever in Osaka City could be expressed by normal cyclic distribution of the one year period. The writer counted the distribution curve of the death rate of beriberi in Osaka City by the same method with the one used by Dr. Masuyama, but the difference between the actual value and calculated value was so great and moreover I found out that the nature of the curve was different from that of typhoid fever in its system.

This was, I thought, owing to the fact that the period of the cause of beriberi outbreak was different from that of typhoid fever and it was not appropriate to apply the normal cyclic distribution of the one year period to it. The writer, therefore, added to it another normal cyclic distribution of the one year period of which phase slipped down by half a year, that is, I tried to express the death curve of beriberi by composition of those two periods.

The first period was supposed to have the cause of outbreak in summer while the second period, in winter. Although the hypothesis that beriberi has the cause of outbreak in winter besides summer has not been proved actually yet by an experiment, it has been led to this by the stochastic analysis of the writer and others upto this time. By the following facts the writer has thought of the hypothesis that beriberi has the cause of outbreak in winter, too.

1. In most of the cities of various districts in Japan the type of the distribution of seasonal variation of beriberi death has a small peak in the time from winter to spring besides the big peak in summer, or even in such a type as not to have a small peak, the curve goes on the straight in the time from winter to spring without going down like that of typhoid fever and others.

2. As I mentioned before, if the correlogramm is drawn by the periodical analysis using auto-correlation method, there are many cities which have two periods, one is the big one year period and the other is the little half-a-year period.

I will show the formula and the method of calculation to express the distribution curve of beriberi $G_1(t)$ and $G_2(t)$ are to express distribution curves of beriberi by the two causes of outbreak respectively while $h=1$ and $k=1$ are to express the periods by the two causes of outbreak respectively, but k is slipped down by half a year from h in the phase. The temporal distribution curve is expressed by

$$f(t) = C_1 \sum_{-\infty}^{\infty} G_1(t-nh) + C_2 \sum_{-\infty}^{\infty} G_2(t-nk)$$

The distribution by the two causes of outbreak will be as follows if the normal distribution is estimated.

$$G_1(t) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2\sigma_1^2}(t-m_1)^2}$$

$$G_2(t) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2\sigma_2^2}(t-m_2)^2}$$

therefore, $\varphi(t)$ the cumulative sum of $f(t)$ is expressed by the following functions.

$$\begin{aligned} \varphi(t) = \int_0^t f(t)dt = & t + (c_1q_1 \cos 2\pi m_1 + c_2q_2 \cos 2\pi m_2) \sin 2\pi t/\pi \\ & + (c_1q_1^4 \cos 4\pi m_1 + c_2q_2^4 \cos 4\pi m_2) \sin 4\pi t/2\pi \\ & + (c_1q_1^9 \cos 6\pi m_1 + c_2q_2^9 \cos 6\pi m_2) \sin 6\pi t/3\pi \\ & + \dots \end{aligned}$$

c_1 and c_2 show the weight when G_1 is put upon G_2 and their relation is $c_1 + c_2 = 1$. m_1 and m_2 show the maximum position of the value of G_1 and G_2 and they are

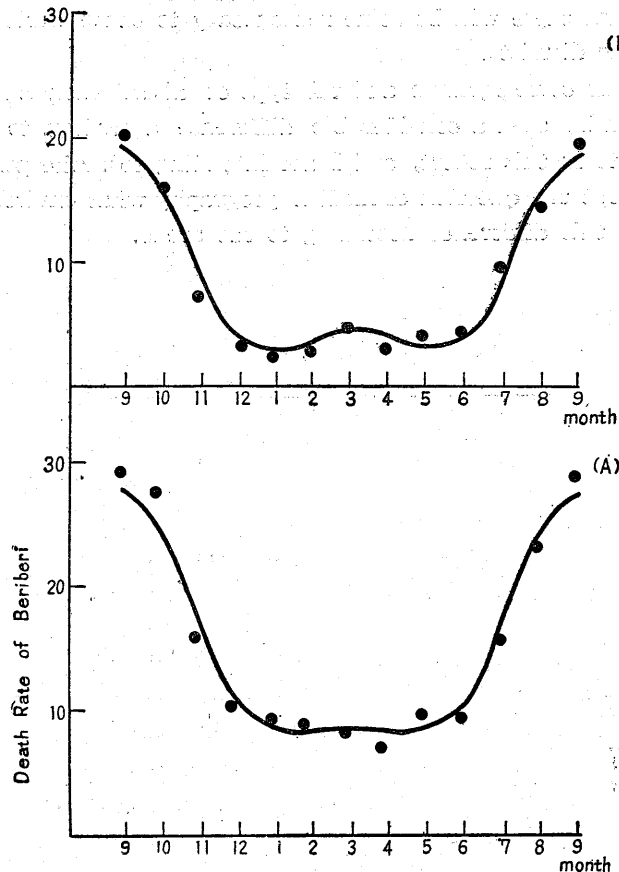


Fig. 2 Distribution curve of the death rate of beriberi.
 (A) : Osaka City (B) : Fukuoka City • Actual value — Theoretical curve

also defined by $q_1 = e^{-2\pi^2\sigma_1^2}$, $q_2 = e^{-2\pi^2\sigma_2^2}$.

In case of Osaka, if I get the general value of c_1 , c_2 , q_1 and q_2 in the graph and then I revise the value of these constants by the successive approximation, they will become as follows :

$$c_1=0.775, c_2=0.255, q_1=0.641, q_2=0.660, m_1=0.063, m_2=0.56.$$

If I calculate, by using the above-mentioned constants, the average monthly death rate of beriberi in Osaka City during 9 years from 1926 to 1934 as the actual value, I can get the distribution curve shown in the graph (Fig.2) and find out that actual measurement and calculation coincide well.

And if I draw the distribution curve in the same way for Fukuoka City where a peak exists in winter, I also find that actual measurement and calculation coincide well. The value of constant in Fukuoka City is as follows :

$$c_1=0.801, c_2=0.199, q_1=0.724, q_2=0.661, m_1=0.0, m_2=0.5.$$

Judging from the above-mentioned fact that the actual value and theoretical value of the distribution curve have become same by the composition of two yearly periods, one of them is slipped down by half-a-year period in the phase, I can imagine that there will be some reason easy to cause beriberi death even in winter in some districts.

Concerning the correlogram and the type of distribution curve of beriberi death, there can be seen a considerable difference according to the areas, as I have referred to in this report, and I am intending to make public in full at another time about the question of medical geography what environmental factors are the cause of this difference according to the areas.