Reports of the Academy of Sciences of the USSR

1968

SovietRxiv

Abstract

Full Text

Reports of the Academy of Sciences of the USSR
1968. Volume 180, No. 2

UDC 519.2:61

CYBERNETICS AND CONTROL THEORY

L. A. RVACHEV

AN EXPERIMENT IN MODELING A LARGE-SCALE EPIDEMIC ON A DIGITAL COMPUTER

(Presented by Academician V. M. Glushkov, 22 V 1967)

If one considers the equations of the mathematical theory of epidemics ([1,2]), which have a more or less constructive character, from the point of view of the scale of the processes described by these equations, then one may conclude that the greatest progress has apparently been achieved by Bartlett' s group, which constructed on a digital computer a qualitatively satisfactory model of an abstract epidemic process within a single city. The present article describes an experiment carried out on a digital computer to model an outbreak of influenza for a system of 128 cities of the USSR, with a transportation network and other parameters close to reality. Influenza was chosen as the most widespread infectious disease.

Let us denote the cities by numbers from 1 to n. We also introduce the following notation: ρ_i is the population of city i; σ_{ij} is the number of persons who have traveled from city i to city j per unit time; $\varphi_i(t, l)$ is the number of persons in city i at time t who became ill at the earlier time l (in $\varphi_i(t, l)$ are included both the sick and those who have already recovered); $x_i(t)$ is the number of nonimmune persons in city i at time t; $\lambda_i(t)$ is the probability, per unit time, of encountering someone with an intimacy sufficient for transmission of the infection; T is the maximum duration of illness; $g(\tau)$ is the probability of remaining ill after time τ from the onset of the disease. The unknown functions are $\varphi_i(t, l)$, with domain of definition $t \geq l$, and $x_i(t)$, and the initial data are $\varphi_i(0, l)$, $x_i(0)$.

To construct the model on the digital computer, the following system of $3n$ equations was used (their derivation is given in ([3])), where $i = 1, 2, \dots, n$:

$$\frac{\partial \varphi_i}{\partial t} = \sum_{j=1}^{n} \left(\frac{\sigma_{ji}}{\rho_j} \varphi_j - \frac{\sigma_{ij}}{\rho_i} \varphi_i \right),$$

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} \left(\frac{\sigma_{ji}}{\rho_j} x_j - \frac{\sigma_{ij}}{\rho_i} x_i \right) - \varphi_i(t,t), \tag{1}$$

$$\varphi_i(t,t) = \frac{\lambda_i(t)}{\rho_i} x_i(t) \int_0^T \varphi_i(t, t-\tau) g(\tau)\, d\tau,$$

where the last equation represents a boundary condition on the line $t = l$. The quantity σ_{ii} has been introduced for generality, and its numerical value is, obviously, immaterial.

Let us examine in more detail how the real values of the parameters were approximated, since this is one of the principal difficulties of the mathematical theory of epidemics. As already stated, $n = 128$, and the cities with the largest population were taken and arranged in decreasing order of population; as ρ_i, the data given in ([5]) were used. One day was chosen as the unit of time. For an approximate determination of σ_{ij}, the empirical formula $\sigma_{ij} = 2^{-32}\rho_i\rho_j$ was used. A comparison of this formula with the data of the Ministry—

data of the Ministry of Communications and the Ministry of Civil Aviation of the USSR on passenger turnover between Moscow and other cities of the selected system for October 1966 showed that the formula gives about a 15% root-mean-square error. It is clear that substituting this formula into system (1) noticeably simplifies the system.

Averaged data from various authors on the distribution of the duration of illness for a group of persons who fell ill simultaneously are given in (4); they make it possible to compute $g(\tau)$: $g(0) = 1$, $g(1) = 0.90$, $g(2) = 0.55$, $g(3) = 0.30$, $g(4) = 0.15$, $g(5) = 0.05$, $g(6) = 0$. Hence it follows that $T = 5$. Next, the so-called "proepidemicization" of the population was taken into account, i.e., the acquisition of a certain immunity in connection with the universally circulating virus (4). This fact was formalized as follows: $\lambda_i(t) = \lambda(0)x_i(t)/x_i(0)$ up to the moment $t_{\max}$ until which the quantity

$$y_i(t) = \int_0^T \varphi_i(t, t-\tau)\, g(\tau)\, d\tau,$$

increases, i.e., until the number of sick persons $y_i(t)$ begins to decrease. At subsequent moments, for $t > t_{\max}$, we shall set $\lambda_i(t) = \lambda_i(t_{\max})$.

Thus, the model is defined uniquely; but in order for it to function, its initial state must be specified. For this purpose one must consider a concrete epidemic, since for the aggregate of reasons the quantities $\varphi_i(0,l)$, $x_i(0)$, $\lambda(0)$ express the individuality of each epidemic and change when moving to another epidemic.

We settled on the large outbreak of 1965, since in this case a virus was spreading that had a definite element of novelty, which makes it possible to write, for all

Fig. 1

Figure 1: Fig. 1

cities except the initial one, $\varphi_i(0,l) = 0$. This outbreak was introduced from Western Europe into Leningrad, where the morbidity level exceeded the natural background on 4 I 1965. Therefore this date was adopted as the initial time, and, taking into account that Leningrad' s number is 2, the corresponding statistical data were recorded for $\varphi_2(0,l)$, $0 \geq l \geq -5$ (smaller values of l, in view of $T = 5$ and the structure of system (1), are obviously not needed). In view of the above-mentioned novelty of the virus, one can also write that $x_i(0) = c\rho_i$, where c does not depend on i. Since at present there are no methods for direct determination of $x_i(0)$, $\lambda(0)$, they were determined indirectly in the following way. We had at our disposal data from the sanitary-epidemiological station of Moscow on the daily morbidity $\varphi_1(t,t)$ for the period of the outbreak in the city, 18 I 1965–28 II 1965. Therefore we could regard any formula of the local epidemic (1), expressing morbidity on a certain day through morbidity on preceding days and the function $g(\tau)$, as an equation for determining $x_1(0)$, $\lambda(0)$ (specifically, 29 I 1965 was chosen). It is impossible to compose a second equation for $x_1(0)$, $\lambda(0)$ in the same way, since in all such equations the unknowns enter in the form of the product $x_1(0)\lambda(0)$. Therefore the second equation was composed on the basis of the total number S_1 of those who had been ill over the entire time of the epidemic and the threshold theorem of Kermack and McKendrick (1); moreover, since the recovery coefficient μ enters into the threshold formula, μ was first determined from $g(\tau)$ at the equilibrium level $\varphi_1(t,t)$ (background) and turned out to be equal to 1/3.

From this it was computed that $x_1(0) = 0.4\rho_1$, $\lambda(0) = 1.6$.

With the values of the constants thus obtained, the epidemic for the indicated system of 128 cities was calculated for the period January–February 1965, beginning, as already stated, in Leningrad and spreading over the entire system. In Fig. 1 the calculated course of the epidemic for Moscow (solid line) is compared with the daily data for Moscow from medical statistics over six weeks, 18 I 1965–28 II 1965 (broken dashed line). The dashed line arose as a result of connecting those points of medical statistics that are not ne-

reliable. The remaining points of medical statistics are shown in the figure as isolated points; the following can be said about them. The data for Saturdays and Sundays are certainly lower than the real values (which is clearly visible in the figure), since on these days the population less often applies for a sick leave certificate. The data for Mondays are certainly higher than the real values, since on these days some of those persons who fell ill on the preceding day also apply.

Fig. 1

It is evident from the figure that the main characteristics of the real and computed epidemics are very close, namely: the computed moment of the beginning of the outbreak (i.e., the moment when the background level is exceeded) differs from the real one by 1 day; the deviation of the computed duration of the outbreak from the real one is 6%; the computed moment of the maximum of the outbreak does not differ from the real one; the deviation of the computed value of the maximum from the real one is 11%; the deviation of the computed total number of all those who fell ill from the real total number of all those who fell ill is 10%; the computed and real epidemic curves have a similar asymmetric form. The closeness of the two curves is also apparent, and the root-mean-square deviation in this case is 25%. These facts are encouraging, since both the parameter values and the model itself were constructed on the basis of one set of data, whereas the comparison is made with another (those several numbers that were used to determine $x_1(0)$, $\lambda(0)$ were not included in our model, but in a model of the Kermack-McKendrick type). However, the significance of these facts should not be overestimated, since the method applied by us makes it possible to determine the initial state of the epidemic only after it has ended, while modern diagnostic methods, as already mentioned, do not in general make it possible to determine the initial state of an epidemic. Nevertheless, even without forecasting epidemics, considerable benefit can be obtained from "playing through" experimental epidemics on a digital computer. In particular, epidemiology differs unfavorably from other natural disciplines in the sense that the principal object it studies—the epidemic process—can neither be deliberately produced for the sake of an experiment nor observed repeatedly under identical conditions. In this respect there can be only one way out—to experiment with such mathematical and machine processes as constitute as accurate a model as possible of real epidemics.

Institute of Epidemiology and Microbiology
named after N. F. Gamaleya
Academy of Medical Sciences of the USSR

Received
12 V 1967

CITED LITERATURE

1. N. Bailey, *The Mathematical Theory of Epidemics*, London, 1957.
2. M. Bartlett, *Stochastic Population Models*, N. Y., 1960.
3. O. V. Baroyan, L. A. Rvachev, *Cybernetics*, No. 3 (1967).
4. V. M. Zhdanov, V. D. Solov' ev, F. G. Epshtein, *The Study of Influenza*, Moscow, 1958.
5. *Handbook of the Supreme Soviet of the USSR on the Administrative-Territorial Division of the USSR*. Moscow, 1965.