It may indeed be queried whether theoretically, with an indefinitely fine network of points, we shall be led to a unique function $h(s^2, x)$ with the common sense properties, which, from general statistical considerations, we know it should have in order to be acceptable. As with integral equations of a simpler character, the passage from a discrete network to a continuum may raise problems, but it is the author's opinion that the infinite ranges of x and s_i^2 give us the freedom which we require in the solution.

The author, however, prefers to approach the problem from the numerical behavior of the series, of which (15) gives the general terms. Here the practical issue appears to be to investigate the relation between the magnitude of the last term retained and the f_i. The author hopes in a further paper to give some results of an investigation of this character and also some tables facilitating the calculation of $h(s^2, x)$.

REFERENCE

[1] B. L. Welch, "The generalization of 'Student's' problem when several different population variances are involved". *Biometrika,* Vol. 34 (1947), pp. 28–35.

PROBABILITY SCHEMES WITH CONTAGION IN SPACE AND TIME[1]

By Félix Cernuschi[2] and Louis Castagnetto

Harvard University

1. Summary. In many natural assemblies of elements, the probability of an event for a given element depends not only on the intrinsic nature of that particular element, but also on the states of some or all of the rest of the elements belonging to the same assembly. On the basis of this general idea of "contagion" some urn schemes are developed in this paper in which one has contagious influence in space and time. The most interesting result found is that in general the points of convergence of the probability of the assembly are given by some of the roots of an equation $p = f(p)$ and that some of these roots, between zero and one, represent stable states of the assembly, or points of convergence, and others represent unstable ones, or points of divergence. The two neighboring roots, (if they are single), of a root representing a point of convergence are unstable values of the probability. Consequently, under certain conditions, the limiting probability may be made to have a finite jump by changing the initial probability by an arbitrarily small amount. The concrete cases developed in this paper can be considerably extended by similar methods by assuming more complicated and general assemblies and laws of contagion.

[1] On the suggestion of the referee, some parts of the original paper were deleted and some mathematical simplifications were introduced.

[2] Research Associate at Harvard Astronomical Observatory and Guggenheim Fellow.

2. Introduction. In the known probability schemes of contagion of Eggenberger and Polya [1], Greenwood and Yule [2], Lüders [3], Neyman [4], Feller [5] and others [6], as well as in Markoff chains different ways are considered in which the previous results in a definite series of trials may influence the probabilities of the future ones. All of these schemes consider possible influences of the results of the different trials along the time axis; and consequently might be called schemes of contagion in one dimension and one direction.

In many natural assemblies of individuals or elements, the probability of an event per individual or element depends not only on the intrinsic nature of the considered element but also on the states of the rest of the elements belonging to the same assembly.

The purpose of this paper is to develop some simple schemes with urns in which there is a contagious influence in space and time and to show some of their consequences. The method which we have used to treat certain concrete cases could be applied to more complicated assemblies and laws of influence in space and time.

3. Scheme of a closed assembly of urns in two dimensions. Let us consider a set of N urns arranged on a closed surface in such a way that each one of them is surrounded by m others. Let each urn contain a finite number of black and white balls. In this paper the probability associated with an urn will refer to the probability of obtaining a white ball if a single ball is drawn at random from the urn. We shall assume that the initial probabilities are equal for all of the urns and that the following law of influence holds: When, after a collective trial, one finds that the ball drawn from a certain arbitrary urn, taken as the central one, is white and that the corresponding results of the m surrounding urns give l white and s black balls, one multiplies the probability of obtaining a white ball out of the central urn by the factor $\alpha_{1,1}^{l}\alpha_{1,2}^{s}$; if the ball drawn from the central urn were black, without changing the given results of the surrounding urns, one multiplies the considered probability by the factor $\alpha_{2,2}^{l}\alpha_{2,1}^{s}$. Under the specified conditions, it is easily seen that the probability of obtaining a white ball from a definite urn at the $i + 1$ trial will be, by considering all the possible alternatives:

$$
\begin{aligned}
p_{i+1} &= m!\sum_{j=0}^{m}\frac{1}{j!(m-j)!}\,[p_i^2(p_i\alpha_{1,1})^{m-j}(q_i\alpha_{1,2})^{j} + p_iq_i(p_i\alpha_{2,1})^{m-j}(q_i\alpha_{2,2})^{j}] \\
&= f(p_i) = p_i^2(p_i\alpha_{1,1} + q_i\alpha_{1,2})^{m} + p_iq_i(p_i\alpha_{2,1} + q_i\alpha_{2,2})^{m},
\end{aligned}
\tag{1}
$$

where:

$$p_i + q_i = 1.$$

Consequently p_i either converges to a root of the equation $p = f(p)$ or tends to infinity. As a probability greater than one or smaller than zero has no meaning, we have to study the function $y = f(p)$ between zero and one. In (1) we have given an implicit form for $y = f(p)$, corresponding to a particular case of influence; by changing the law of influence we change the function $f(p)$. In general one can find graphically the roots of equation $p = f(p)$ by plotting $y = f(p)$ and $y = p$ and by determining the intersections of these two lines in the range $0 \leq p \leq 1$. Later we shall give the values of these roots for some concrete examples. From what we have shown it follows that if, for the considered assembly of urns and for especially chosen values of the parameters of interconnection and initial probabilities, the probability tends to some equilibrium value, this must be a root of the equation $p = f(p)$. As we shall see later, the roots in the range $0 \leq p \leq 1$ may represent stable or unstable states of the assembly.

Let us consider now a general method for finding the explicit form of the function $f(p)$ corresponding to laws of influence similar to the one used by Polya.

Assume that the trial i results in the drawing of l white balls and s black balls from the m urns surrounding the central one. Then we add lw_1 white and sb_1 black balls to the central urn if the result of the central urn was white, and lw_2 white and sb_2 black balls if it was black. It is easy to show that under these conditions the probability in the trial $i + 1$ is related to the probability in trial i by the following formula:

$$\begin{aligned} p_{i+1} = {} & p_i \int_0^1 t^{N_i - W_i - 1} \left[t_1 \frac{\partial}{\partial t_1} t_1^{W_i} (p_i t_1^{w_1} + q_i t_2^{b_1})^m \right]_{t_1 = t_2 = t} dt \\ & + (1 - p_i) \int_0^1 t^{N_i - W_i - 1} \left[t_1 \frac{\partial}{\partial t_1} t_1^{W_i} (p_i t_1^{w_2} + q_i t_2^{b_2})^m \right]_{t_1 = t_2 = t} dt \end{aligned} \tag{2}$$

where W_i and N_i are the number of white balls and the total number of balls, respectively, in the central urn before trial i. Relation (2) permits us to study several interesting schemes. It is easy to see that all the possible schemes which can be represented by relations of type (2) give only values of the probability in the interval zero and one; and consequently we do not need to make the restriction in the analysis of the equation $p = f(p)$ that was necessary in the previous scheme, represented by equation (1).

For the case $w_1 = b_1 = c_1$, $w_2 = b_2 = c_2$, we obtain from (2)

$$p_{i+1} = p_i \frac{W_i + mp_i c_1}{N_i + mc_1} + (1 - p_i) \frac{W_i + mc_2 p_i}{N_i + mc_2}. \tag{3}$$

If $c_1 = c_2$, (3) gives

$$p_{i+1} = p_i . \tag{4}$$

If one takes $c_1 = k_1 N_i$ and $c_2 = k_2 N_i$ (3) becomes

$$p_{i+1} = p_i \frac{p_i + mk_1 p_i}{1 + mk_1} + (1 - p_i) \frac{p_i + mk_2 p_i}{1 + mk_2} = f(p_i) \tag{5}$$

and the equation $p = f(p)$ has, in this case, the roots 0 and 1.

When $w_1 = b_2 = k_1N_i$ and $b_1 = w_2 = k_2N_i$, one has to replace $t_1(\partial/\partial t_1)$ by $t_2(\partial/\partial t_2)$ in the second term of (2); then if we take $m = 2$,

$$(6)\quad p_{i+1} = \frac{p_i + 2k_1}{1 + 2k_1} + p_i q_i \left(\frac{p_i}{1 + 2k_2} + 2\,\frac{p_i + k_1}{1 + k_1 + k_2} - 3\,\frac{p_i + 2k_1}{1 + 2k_1} \right) = f(p_i).$$

In particular, if $k_1 = k_2 = k$, one obtains

$$(7)\qquad p_{i+1} = \frac{1}{1 + 2k_1}\,[4kp_i^2 - (4k - 1)p_i + 2k] = f(p_i),$$

and the solutions of the equation $p = f(p)$ are $p = \frac{1}{2}$ and 1. By considering the behavior of $y = f(p)$ one finds that the stable solution is given by the root $\frac{1}{2}$; consequently if one starts with any value of $0 < p < 1$ the probability tends to the limiting value $\frac{1}{2}$. If $k_1 = 0$, $k_2 \neq 0$, by simple calculations, one obtains from (6) that the solutions of $p = f(p)$, in this case, are zero and one.

The equation $p = f(p)$, as given by (6), always has the solution 1. In order to have the other two roots real, one has to satisfy:

$$(8)\qquad \begin{aligned} k_1(1 + 2\,k_2)\,(2 + k_1 + 3\,k_2)^2 \geq 4(1 + k_1 + k_2) \\ [(k_1 + k_2)^2 + 2(k_1 - k_2) - 4\,k_2^2]. \end{aligned}$$

A simple and interesting application of relation (2) is for the case of two urns, characterized by $m = 1$. From (2) we obtain:

$$(9)\quad p_{i+1} = p_i^2\left(\frac{p_i + k_1}{1 + k_1} + \frac{1 - p_i}{1 + k_2}\right) + (1 - p_i)p_i\left(\frac{p_i + k_3}{1 + k_3} + \frac{1 - p_i}{1 + k_4}\right) = f(p_i)$$

where

$$w_1 = k_1N_i\,, \quad b_1 = k_2N_i\,, \quad w_2 = k_3N_i\,, \quad b_2 = k_4N_i\,.$$

The equation $p = f(p)$, as given by (9) has the roots 0 and 1; and one may fix the value of the third root by conveniently choosing the values of the parameters.

Applying (2) for an arbitrary value of m and integrating by parts, it is seen that in general the equation $p = f(p)$ is of degree $m + 2$ and consequently, by choosing appropriate values for the parameters k_1, k_2, k_3, k_4, each of which may be between -1 and ∞, one can expect several roots in the range $0 \leq p \leq 1$. One can easily generalize our relation (2) for cases in which w_1, w_2, b_1, b_2 are given functions of the probability p_i. Even in this most general case it is simple to see that one would have a recursion formula of the type $p_{i+1} = f(p_i)$ and, as in the elementary cases which we have considered, the points of equilibrium of the closed assembly of urns will be given by those solutions, in the range $0 \leq p \leq 1$, of the equation $p = f(p)$, where the derivative of $y = f(p)$ is negative. Consequently the two neighboring roots, if they are single, of a root representing a point of convergence are unstable values of the probability. Therefore, under certain conditions, the limiting probability may take a finite jump if the initial probability is changed by an arbitrarily small amount. This is, we think, the most important consequence of the contagion schemes that we propose. We consider that many actual cases of contagion could be better understood by schemes of the type that we are studying.

Let us consider now some simple cases of relation (1). If we take

$$\alpha_{1,1} = \alpha_{2,2} = \alpha_1 \quad \alpha_{1,2} = \alpha_{2,1} = \alpha_2 \quad \text{and} \quad m = 2,$$

representing a closed ring of urns, one obtains:

$$\begin{aligned} p_{i+1} &= p_i^2(\alpha_1 p_i + \alpha_2 q_i)^2 + p_i q_i(\alpha_2 p_i + \alpha_1 q_i)^2 \\ &= p_i^2 + (p_i^2 - p_i^3)\,[(\alpha_1 + \alpha_2)^2 - 4\,\alpha_1^2] = f(p_i). \end{aligned} \tag{10}$$

The equation $p = f(p)$, corresponding to this recursion formula, always has the solution $p = 0$. The other two solutions are given by

$$P_{1,2} = \tfrac{1}{2}\left[1 \pm \sqrt{\frac{4 - (\alpha_1 + \alpha_2)^2}{4\alpha_1 - (\alpha_1 + \alpha_2)^2}}\right]. \tag{11}$$

These roots will be between 0 and 1 when

$$\begin{matrix} 2 < \alpha_1 + \alpha_2 \\ 1 > \alpha_1 < \alpha_2 \end{matrix} \quad \text{or (12')} \quad \begin{matrix} 2 > \alpha_1 + \alpha_2 \\ 1 < \alpha_1 > \alpha_2 \end{matrix}. \tag{12}$$

We would have $P_1 > 0$ and $P_2 < 0$ if

$$\begin{matrix} 2 < \alpha_1 + \alpha_2 \\ 1 < \alpha_1 < \alpha_2 \end{matrix} \quad \text{or (13')} \quad \begin{matrix} 2 > \alpha_1 + \alpha_2 \\ 1 > \alpha_1 > \alpha_2, \end{matrix} \tag{13}$$

and $P_1 = P_2$ when

$$\alpha_1 + \alpha_2 = 2, \qquad \alpha_1 \neq 1. \tag{14}$$

Let us now study the general behavior of (10). For the conditions (12′) we have:

$$p_{i+1} - p_i = a^2 p_i(p_i - P_1)\,(p_i - P_2) \tag{15}$$

where
$$a^2 = 4\,\alpha_1^2 - (\alpha_1 + \alpha_2)^2 > 0.$$

If $0 < P_1 < P_2$, one obtains from (15) by use of elementary algebra:

$$\left|\frac{p_{i+1} - p_i}{P_1 - p_i}\right| = a^2 p_i \mid (P_2 - p_i) \mid \leq \frac{a^2 P_2^2}{4} \leq 1. \tag{16}$$

Consequently if $p_1 > P_2$ the sequence p_i increases monotonically. Otherwise p_{i+1} will lie between P_1 and p_i and will tend to P_1 without ever reaching the other side of this point. In a similar way it is possible to prove the convergence to a constant for the most general equations of the type $p = f(p)$ when they have roots between zero and one.

Let us give some numerical results. For $\alpha_1 = 0.95$ and $\alpha_2 = 1.1$, from (10) one obtains: $P_1 = 0.1$ and $P_2 = 0.9$. It is easily seen that, in this case, if

$0 < p_1 < 0.1$, the limiting value of p_i will be zero; if $p_1 > 0.1$, the limiting value will be 0.9. The interesting point is that if the initial probability is in the neighborhood of 0.1, an infinitesimal change in its value may produce a finite change in the stable limiting probabilities; and that for the initial probability equal 0.1 one would have an unstable equilibrium of the system. This consideration shows why it is important to know how the probability p_i converges towards a certain point. As we have previously shown, the points of convergence are roots of the eq. $p = f(p)$ but there roots which are not points of convergence.

Similar reasoning could be applied to more complicated systems belonging to our general scheme of contagion. Consequently, the most important result is not that the considered assembly may have a probability tending to some value in the range $0 \leq p \leq 1$, but that under certain conditions the limiting probability may jump from one value to another by changing the initial probability by an arbitrarily small amount.

REFERENCES

[1] F. Eggenberger and G. Poylá, *Zeits. fur Ang. Math. und Mech.* Vol. 3 (1923), p. 279.
[2] M. Greenwood and G. U. Yule, *Roy. Stat. Soc. Jour.*, Vol. 83 (1920), p. 255.
[3] R. Luders, *Biometrika*, Vol. 26 (1934), p. 108.
[4] J. Neyman, *Annals of Math. Stat.*, Vol. 10 (1939), p. 35.
[5] W. Feller, *Annals of Math. Stat.*, Vol. 14 (1943) p. 389.
[6] F. Cernuschi and E. Saleme, *Anales Soc. Cientifica Argentina*, Vol. 138 (1944), p. 201.

FITTING CURVES WITH ZERO OR INFINITE END POINTS

By Edmund Pinney

Oregon State College

The problem of determining a suitable equation to fit an empirically determined curve over a given interval has been of great importance in statistical work, in experimental science, and in engineering technology. Since infinitely many types of equations may be made to fit the data with required accuracy, the choice of a "suitable" type of equation depends on the qualitative nature of the empirical curve, on the use to which the equation is to be put, and upon considerations of simplicity.

As a function type, the polynomial has, because of its simplicity, been enormously useful. The function type studied here is a little more general than the polynomial type, being particularly useful in the case of empirical curves that become zero or infinity at one or both ends of the interval.

Without loss of generality the interval in which the equation is to fit the curve may be taken as $0 \leq x \leq 1$. It is assumed that, by numerical means or otherwise, a finite set of moment $\mu_m = \int_0^1 yx^m \, dx$ may be computed, y being the ordinate of the empirical curve.