12

The Vacillating Mathematician

Where Does She End Up?

K B Athreya

The problem of a mathematician who walks from her home to her office and changes her mind repeatedly during this walk is discussed. Stochastic generalizations of this problem can be used to model many real-life situations.

The Deterministic Version

A mathematician starts walking from her home to her office. Halfway through she changes her mind and starts returning home. Again halfway through that she changes her mind and starts walking towards her office. Once again halfway through that she starts returning home and so on. The problem is to determine what happens to this vacillating mathematician.

Identifying the mathematician's home as the point zero and her office as the point one we can formulate a sequence $\{X_n\}_0^\infty$ (X_n denoting the position at the n^{th} change point) of numbers in the interval [0,1] that satisfies the following simple rule:

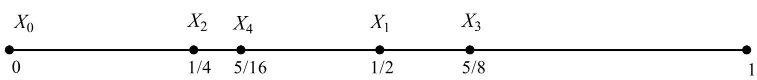

Figure 12.1

$$X_0 = 0, X_1 = \frac{1}{2}, X_2 = \frac{1}{4}, X_3 = \frac{1}{4} + \frac{1}{2}\cdot\frac{3}{4} = \frac{5}{8},$$
$$X_4 = \frac{1}{2}\cdot\frac{5}{8} = \frac{5}{16}\ldots$$

$$\left.\begin{aligned} \text{Clearly, } X_{2n+1} &= X_{2n} + \frac{(1 - X_{2n})}{2} \text{ for } n \geq 0 \\ X_{2n} &= \frac{1}{2} X_{2n-1} \text{ for } n \geq 1 \end{aligned}\right] \quad (1)$$

Letting $U_n = X_{2n-1}, V_n = X_{2n}$ we see that

$$\begin{aligned} U_{n+1} &= V_n \left(\frac{1}{2}\right) + \frac{1}{2} \\ &= \left(\frac{1}{4}\right) U_n + \frac{1}{2} \text{ for } n \geq 1. \end{aligned} \quad (2)$$

This is known as a *first order difference equation.* To solve this iterate the equation to get

$$\begin{aligned} U_{n+1} &= \left(\frac{1}{4}\right) \left(\left(\frac{1}{4}\right) U_{n-1} + \frac{1}{2}\right) + \frac{1}{2} \\ &= \left(\frac{1}{4}\right)^2 U_{n-1} + \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{2} \\ &= \left(\frac{1}{4}\right)^2 \left(\frac{1}{4} U_{n-2} + \frac{1}{2}\right) + \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{2} \\ &= \left(\frac{1}{4}\right)^3 U_{n-2} + \left(\frac{1}{4}\right)^2 \frac{1}{2} + \left(\frac{1}{4}\right) \frac{1}{2} + \frac{1}{2}. \end{aligned}$$

It is easy to guess from the above and also establish by induction that

$$U_{n+1} = \left(\frac{1}{4}\right)^n U_1 + \sum_{j=0}^{n-1} \left(\frac{1}{2}\right)\left(\frac{1}{4}\right)^j. \quad (3)$$

Since $0 \leq U_1 \leq 1$, $\left(1/4\right)^n U_1 \to 0$ as $n \to \infty$. Also the geometric series partial sum sequence

$$\sum_{0}^{n-1} \left(\frac{1}{2}\right) \left(\frac{1}{4}\right)^j \to \frac{\left(\frac{1}{2}\right)}{1 - \frac{1}{4}} = \frac{2}{3}.$$

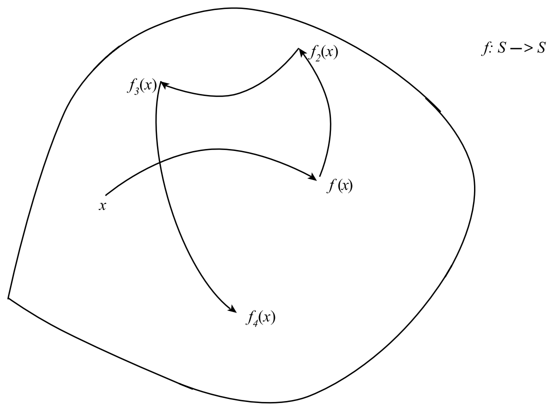

Figure 12.2

(Recall that) $\sum_0^n ar^j \to a/(1-r)$ for $|r| < 1$). Thus $U_n \to 2/3$ and since $V_n = X_{2n} = 1/2\, U_n$ it converges to $1/3$.

So our mathematician's position at odd numbered change points is non decreasing (note that $U_2 = 5/8$ is $> U_1 = 1/2$ and $U_{n+1} - U_n = (1/4)(U_n - U_{n-1}) = (1/4)^{n-1}(U_2 - U_1) > 0$) and converges to $2/3$. Similarly V_n is also nondecreasing and approaches $1/3$.

> *Eventually, our mathematician will just be hopping in the vicinity of 1/3 and 2/3.*

If $0 < X_0 < 1$ then the above arguments are still valid and the limits are the same. However, if $X_0 > 1/3$ then U_n and V_n would both be decreasing to 2/3 and 1/3 respectively. (Prove this).

Finally, note that if $X_0 = 1/3$ then $U_n = 2/3$ and $V_n = 1/3$ for all n. So 2/3 and 1/3 are *fixed points* for the $\{U_n\}$ and $\{V_n\}$ sequences or for the *dynamical systems* generated by the functions f and g respectively i.e.,

$$f(x) = \frac{1}{4}x + \frac{1}{2},\ g(x) = \frac{1}{4}x + \frac{1}{4} \qquad (4)$$

for x in [0,1]. For a function f from a set S into itself the sequence $\{ f_0(x) = x,\ f_n(x) = f(f_{n-1}(x)),\ n \geq 1\}$ is called a *dynamical system*. Currently popular topics *chaos* and *fractals* deal with dynamical systems (see Barnsley, and Ramasamy and Iyer in Suggested Reading).

The problem of the vacillating mathematician was posed by Zeev Barel (Suggested Reading). Krishnapriyan (Suggested Reading) solves this com-

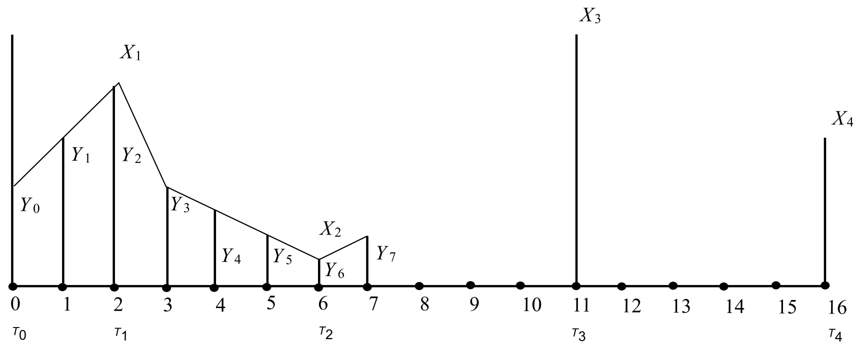

Figure 12.3

pletely and discusses an approach to this problem using difference equations, generating functions and matrix methods.

Now suppose our mathematician likes to inject some randomness in her moves. What happens to the sequence $\{X_n\}_0^\infty$ which now becomes a random sequence? The present article is devoted to answering this question. It will turn out that the set of limit points of the sequence can be (a) the entire interval [0,1] (b) a set like the *Cantor set* (defined later in the article) of length (measure) zero and (c) quite arbitrary and thus very different from the deterministic situation discussed earlier.

We conclude this section with the following two observations:

1. Suppose our mathematician when moving towards one always goes a fraction α of the remaining distance and when moving towards zero always goes a fraction β of the distance. It can be shown that in this case U_n and V_n satisfy

$$\begin{aligned} U_{n+1} &= (1-\alpha)\,(1-\beta)\,U_n + \alpha \\ V_n &= U_n\,(1-\beta) \end{aligned} \tag{5}$$

and hence

$$U_n \to \frac{\alpha}{1-r} \text{ and } V_n \to \frac{(1-\beta)\,\alpha}{1-r} \tag{6}$$

where $r = (1-\alpha)\,(1-\beta)$.

2. The random sequence $\{X_n\}$ generated by the mathematician in the stochastic case is a model that is applicable in some real life situations. For

example, let $\{Y_j\}_1^\infty$, be a random sequence denoting the level in a reservoir on period j by Y_j. Let τ_1, τ_2, τ_3 be the times at which the sequence $\{Y_j\}$ has a local minimum or maximum. Let $X_n = Y_{\tau_n}$. Suppose the reservoir has a minimum zero and a maximum normalised to be one, say. Then the sequence $\{X_n\}$ has a behaviour similar to our mathematician.

Stochastic generalizations of the vacillating mathematician provide models for some real life situations.

It is clear that we could substitute the reservoir level by stock prices, rainfall amounts, inventory level or any other randomly varying sequence in a bounded interval.

We will consider several stochastic (i.e. random) versions in the next article, *A Stochastic Version*. *Stochastic* means random. To analyze these problems we will require some concepts from probability theory which are outlined in the next section.

As a simple stochastic generalization suppose that our mathematician starts at 0, goes half way through and then flips a fair coin. If the coin comes out *heads* she continues towards one and if the coin comes out *tails* she turns back towards 0. Again half way through whatever direction she is headed she flips a fair coin and either continues in that direction or goes in the opposite direction.

To analyse this model, as before, let X_n denote the position at the n^{th} change point. Then, given X_n,

$$X_{n+1} = \begin{cases} X_n + \dfrac{(1-X_n)}{2} & \text{with probability } 1/2 \\ \dfrac{X_n}{2} & \text{with probability } 1/2 \end{cases} \qquad (7)$$

independent of $X_0, X_1, \ldots X_{n-1}$.

Thus the distribution of X_{n+1} given X_n depends only on X_n and does not depend on $X_0, X_1 \ldots X_{n-1}$ or n. In this case the sequence $\{X_n\}_0^\infty$ is called a *Markov Chain* with *stationary transition probabilities* (see Billingsley and Feller in Suggested Reading). A discussion of Markov chains follows in the next section. We will analyze this model and other stochastic generalizations in the next article.

Markov Chains

A sequence of random variables $\{X_n\}_0^\infty$ is called a Markov chain[1] if given X_n the past $(X_0, X_1, \ldots, X_{n-1})$ and the future $(X_{n+1}, X_{n+2}, \ldots)$ are stochastically independent. This property was introduced by A A Markov at the turn of the century as a simple notion of dependence in time evolution and as a departure from full independence (see Feller in Suggested Reading). Markov chains have proved to be very useful in a number of applications, especially in telephone traffic, computer traffic on the information highway, waiting lines, stock prices, etc. When the *transition probability* $P(X_{n+1} = j \mid X_n = i)$ (where $P(A \mid B)$ stands for the probability of the event A given that B has happened) depends only on i and j and not on n the Markov chain $\{X_n\}$ is said to have time homogeneous or stationary transition probabilities. Here we discuss only this case. The non time-homogeneous Markov chains also have important applications. We shall now assume that the sets of values taken by the chain $\{X_n\}_0^\infty$ known as the *state space* is a finite or countable set identified as $\{1,2,3,\ldots\}$.
Let $p_{ij}^{(n)} \equiv P(X_{n+1} = j \mid X_1 = i) = P(X_n = j \mid X_0 = i)$. Then, by the Markov property,

$$
\begin{aligned}
p_{ij}^{(2)} &\equiv P(X_2 = j \mid X_0 = i) \\
&= \sum_k P(X_2 = j, X_1 = k, \mid X_0 = i) \\
&= \sum_k P(X_2 = j \mid X_1 = k, X_0 = i)\, P(X_1 = k \mid X_0 = i) \\
&= \sum_k P(X_2 = j \mid X_1 = k)\, P(X_1 = k \mid X_0 = i) \\
&= \sum_k p_{kj}\, p_{ik}
\end{aligned}
$$

and more generally $p_{ij}^{(n)} \equiv P(X_n = j \mid X_0 = i)$ satisfy

$$p_{ij}^{(n_1+n_2)} = \sum_k p_{ik}^{(n_1)} p_{kj}^{(n_2)}$$

This is known as the *Chapman - Kolmogorov* relation. The above discussion shows that in matrix notation $P^{(2)} \equiv ((p_{ij}^{(2)}))$ is simply P^2 and $P^{(n)} = ((p_{ij}^{(n)}))$

---

[1]A sequence of random variables is called a Markov chain if the past and future of the chain are mutually independent, given the present.

is the nth power P^n of $P = ((p_{ij}))$. The main objects of interest are: (a) the probability distribution μ_n of X_n, i.e.

$$\mu_n(j) \equiv P(X_n = j) \\ = \sum_i p_{ij}^{(n)} \mu_0(i),$$

$\mu_0 = \{\mu_0(i)\}$ being the distribution of X_0, (b) the behavior of μ_n for large n and (c) the behavior of time averages of the sort $\left(\frac{1}{n}\right)\sum_1^n f(X_j)$ for bounded functions f such as the indicator function $I_A(x)$ which is one if x is in A and zero otherwise (in this case the time average is simply the proportion of visits to A by the chain during the first n steps).

A Markov chain $\{X_n\}$ with stationary transition probabilities $P = ((p_{ij}))$ is *irreducible* if for each pair i, j there is an integer n such that $p_{ij}^{(n)} \equiv P(X_n = j \mid X_0 = i)$ is strictly positive. A state i is said to be *recurrent* if $P(X_n = i$ for some $n \geq 1 \mid X_0 = i)$ is one. That is, starting from i the chain returns to i with probability one. If $T_i \equiv min\{n: n \geq 1,\ X_n = i\}$ is the *first return time* to state i, then i being recurrent is the same as $P(T_i < \infty \mid X_0 = i) = 1$. A state i is *transient* if it is not recurrent. A recurrent state i is called *positive* or *null recurrent* according as the mean value of T_i, i.e., $E(T_i \mid X_0 = i)$ is finite or infinite. It can be shown that in the irreducible case, if one state is recurrent (null or positive) all states are recurrent (null or positive). The same is true for transience. A state i has period $d_i \equiv$ g.c.d. $\{n: P(X_n = j \mid X_0 = i) > 0\}$. Then, in the irreducible case $d_i = d$ for all i.

If $d_i = 1$ then i is called *aperiodic*. The main limit theorem for Markov chains is the following.

THEOREM 1. Let $\{X_n\}_0^\infty$ be a Markov chain with stationary transition probability matrix $P = ((p_{ij}))$. Let the chain be irreducible, positive recurrent and aperiodic. Then there exists a probability distribution $\{\pi_j\}$ such that
(a) for each i, j $\lim p_{ij}^{(n)} = \pi_j$ (convergence to equilibrium)
(b) $\pi_j = \sum \pi_i p_{ij}$ (invariance or equilibrium)
(c) $\pi_j = (E(T_j \mid X_0 = j))^{-1}$ (probability of being at j = reciprocal of the mean recurrence time)
(d) $\lim_n \frac{1}{n}\sum_0^r f(X_r) = \sum_j f(j)\pi_j$
for every bounded function $f(.)$ (law of large numbers i.e. time average = ensemble average); in particular for $f = I_A$.

If the aperiodicity condition is dropped then (b), (c), (d) are still true but (a) is replaced by the *Cesaro* convergence i.e. for all *i, j*

$$\lim_{n} \frac{1}{n} \sum_{0}^{n} p_{ij}^{(r)} = \pi_j .$$

When the state space is not countable there is an appropriate extension of the above theorem.

The author would like to thank Alladi Sitaram for encouraging him to write this article and Mohan Delampady and Kavi Ramamurthy for editorial and technical assistance.

Suggested Reading

[1] W Feller. *An Introduction to Probability Theory and its Applications*. Wiley Eastern. Vol I. 1968, Vol II. 1970.

[2] Billingsley P. *Probability and Measure.* Second Edition. John Wiley and Sons. N.Y, 1990.

[3] Zeev Barel. Problem 453. *College Mathematics Journal*. Vol 22. p 255, 1991.

[4] Krishnapriyan H K. *Mathematics Spectrum*. Vol 6. p 9, 1995/6.

[5] Ramasamy B and T S K V Iyer. *Resonance*. Vol 1. No 5. May 1996.

[6] Barnsley M. *Fractals Everywhere*. Academic Press. 1988.

K B Athreya,
Department of Statistics,
Iowa State University,
Ames, IA 50011, USA