A PROBLEM ON RANDOM WALK

R. SHERMAN LEHMAN
STANFORD UNIVERSITY

1. From an urn containing an equal number of each of l kinds of balls, random drawings are made. After each drawing the ball is returned to the urn, so that each drawing is independent of every other. If the drawings are repeated indefinitely, what is the probability that after some drawing in the sequence an equal number of each of the l kinds of balls will have been drawn?

When an equal number of each of the l kinds of balls have been drawn, we shall say that an equalization occurs. In this paper we shall show that for $l = 2$ or $l = 3$ the probability that an equalization will occur is 1, while for $l \geqq 4$ the probability is less than 1.[1]

2. This problem can be interpreted as a random walk in a network of "streets" in $(l - 1)$-dimensional space. The drawing of a ball of a given kind is represented by the walker's moving a fixed distance in a given direction. Equalization is represented by a return to the origin.

A similar problem has been investigated by G. Pólya [2]: the random walk in a network of "streets" in d-dimensional space where the "streets" are parallel to the coordinate axes. The probability of passing through a given point at least once tends to 1 for $d = 1$ or $d = 2$ but does not tend to 1 for $d \geqq 3$. Since in the present problem the random walk takes place in $(l - 1)$ dimensions, the results are entirely analogous.

For $l = 2$ the random walk occurs on a straight line. The walker moves one unit to the right when a ball of the first kind is drawn and one unit to the left when a ball of the second kind is drawn. It is clear that equalization is represented by a return to the origin.

For $l = 3$ the random walk takes place in a network of streets in a plane. We shall use the complex plane. Let $\omega = \frac{1}{2}(-1 + \sqrt{3}i)$, a cube root of unity.

Represent the drawing of a ball by the addition of 1, ω, or ω^2 according as a ball of the first, second or third kind is drawn. If r balls of the first kind, s of the second kind and t of the third kind have been drawn, the random walker will be at the point

$$r + s\omega + t\omega^2 .$$

If there is a return to the origin,

$$r + s\omega + t\omega^2 = 0 .$$

We know $1 + \omega + \omega^2 = 0$. Therefore, $(r - t) + (s - t)\omega = 0$. It follows that $r = s = t$. Equalization occurs. On the other hand, if equalization occurs,

$$r = s = t , \qquad r + s\omega + t\omega^2 = r(1 + \omega + \omega^2) = 0 .$$

[1] It was pointed out to me that certain parts of my result are given in chapter 12 of Professor W. Feller's recent book [1] and that the whole result can be derived by the method given there.

A return to the origin occurs when and only when there is an equalization.

The random walk takes place in a network of equilateral triangles. The junction points, where the streets meet, are all points, $a + b\omega$, where a and b are arbitrary integers. The streets are one way streets.

For $l = 4$ the random walk takes place in three dimensional space. Using vector notation,[2] let

$$\begin{aligned} a &= i + j + k, \\ b &= i - j - k, \\ c &= -i + j - k, \\ d &= -i - j + k. \end{aligned} \tag{1}$$

Represent the drawing of a ball by the addition of $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ or $\mathbf{d}$ according as a ball of the first, second, third or fourth kind is drawn. If r balls of the first kind, s of the second, t of the third and u of the fourth have been drawn, the random walker will be at the point

$$\mathbf{a}r + \mathbf{b}s + \mathbf{c}t + \mathbf{d}u.$$

If there is a return to the origin,

$$\mathbf{a}r + \mathbf{b}s + \mathbf{c}t + \mathbf{d}u = 0.$$

From (1), it follows that

$$\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = 0.$$

Therefore,

$$\mathbf{a}(r - u) + \mathbf{b}(s - u) + \mathbf{c}(t - u) = 0.$$

But $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ are linearly independent. Consequently, $r = s = t = u$ and equalization occurs. On the other hand, if $r = s = t = u$,

$$\mathbf{a}r + \mathbf{b}s + \mathbf{c}t + \mathbf{d}u = r(\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}) = 0.$$

Thus, a return to the origin occurs when and only when there is equalization.

Using (1), we have

$$\begin{aligned} \mathbf{a}r + \mathbf{b}s + \mathbf{c}t + \mathbf{d}u &= \mathbf{i}(r + s - t - u) + \mathbf{j}(r - s + t - u) + \mathbf{k}(r - s - t + u) \\ &= \mathbf{i}x + \mathbf{j}y + \mathbf{k}z. \end{aligned}$$

Then, $x \equiv y \equiv z \pmod 2$. Also,

$$\begin{aligned} 0 &= 0\cdot\mathbf{i} + 0\cdot\mathbf{j} + 0\cdot\mathbf{k}, \\ \mathbf{a} &= 1\cdot\mathbf{i} + 1\cdot\mathbf{j} + 1\cdot\mathbf{k}, \end{aligned}$$

$$\begin{array}{lll} \mathbf{a} + \mathbf{b} = 2\mathbf{i}, & \mathbf{a} + \mathbf{c} = 2\mathbf{j}, & \mathbf{a} + \mathbf{d} = 2\mathbf{k}, \\ \mathbf{c} + \mathbf{d} = -2\mathbf{i}, & \mathbf{b} + \mathbf{d} = -2\mathbf{j}, & \mathbf{b} + \mathbf{c} = -2\mathbf{k}. \end{array}$$

From this, it follows that $\mathbf{i}x + \mathbf{j}y + \mathbf{k}z$ is a junction point of the "street" network if and only if $x \equiv y \equiv z \pmod 2$.

Each point with integer coordinates is at the center of a cube with sides of length one and faces parallel to the coordinate planes. The network of "streets" joins the center of every fourth cube. The "streets" which are the diagonals of the cubes are one way streets.

[2] A vector is indicated by boldface type.

3. The probability of drawing j_1 balls of the first kind, j_2 balls of the second kind and, in general, j_k balls of the k-th kind in the first n times is the coefficient of $x_1^{j_1}x_2^{j_2}\ldots x_l^{j_l}$ in the polynomial

$$\left(\frac{x_1+x_2+\ldots+x_l}{l}\right)^n=\sum\left(\frac{n!}{j_1!\,j_2!\ldots j_l!\,l^n}x_1^{j_1}x_2^{j_2}\ldots x_l^{j_l}\right).$$

The expected number of equalizations in the first n times, where $ml \leqq n < (m+1)l$, is

$$\frac{l!}{1!^l l^l}+\frac{(2l)!}{2!^l l^{2l}}+\ldots+\frac{(ml)!}{m!^l l^{ml}}.$$

The k-th term represents the expected number of equalizations on the kl-th time. The expected number of equalizations in n times is also

$$\rho_1^{(n)}+\rho_2^{(n)}+\ldots+\rho_m^{(n)}$$

where $\rho_j^{(n)}$ is the expected number of j-th equalizations. But $\rho_j^{(n)}$ is also the probability that a j-th equalization will occur. Obviously, $\rho_j^{(n+1)} \geqq \rho_j^{(n)}$ for all n. Because it is a probability $\rho_j^{(n)} \leqq 1$ for all n. The sequence $\{\rho_j^{(n)}\}$ must, therefore, tend to a limit. Let $\lim_{n\to\infty} \rho_j^{(n)} = \rho_j$. Let π_n be the probability that an equalization will occur in the first n times. Then, $\pi_n = \rho_1^{(n)}$ and $\lim_{n\to\infty} \pi_n = \pi = \rho_1$.

The probability that a second equalization will occur in the first $2n$ times is $\rho_2^{(2n)}$. This probability is greater than or equal to the probability that a first equalization will occur not later than the n-th time followed by a second equalization not later than the $2n$-th time. This probability is greater than or equal to π_n^2. We have $\rho_2^{(2n)} \geqq \pi_n^2$.

Similarly, $\rho_3^{(3n)}$ is greater than or equal to the probability that a first equalization will occur not later than the n-th time, followed by a second equalization not later than the $2n$-th time, followed by a third equalization not later than the $3n$-th time. This probability is greater than or equal to π_n^3. In general, we have

$$\rho_k^{(kn)} \geqq \pi_n^k\,. \tag{2}$$

Letting n tend to infinity, we have

$$\rho_k \geqq \pi^k\,. \tag{3}$$

On the other hand,

$$\rho_1^{(n)} \leqq \pi\,,\qquad \rho_2^{(n)} \leqq \pi^2\,,\ \ldots\,,\ \rho_k^{(n)} \leqq \pi^k\,.$$

Letting n tend to infinity, we have

$$\rho_k \leqq \pi^k\,. \tag{4}$$

From (3) and (4), it follows that

$$\rho_k = \pi^k\,. \tag{5}$$

4. LEMMA. *If the series*

$$\sum_{m=1}^{\infty}\frac{(ml)!}{m!^l l^{ml}} \tag{6}$$

diverges, $\pi = 1$. *If the series* (6) *has the sum* σ, $\pi = \sigma/(1 + \sigma)$ *and is therefore less than* 1.

PROOF. First, using (5) we have

$$\sum_{m=1}^{M} \frac{(ml)!}{m!^l l^{ml}} = \rho_1^{(Ml)} + \rho_2^{(Ml)} + \ldots + \rho_n^{(Ml)} \tag{7}$$

$$\leq \rho_1 + \rho_2 + \ldots + \rho_n + \ldots$$

$$= \pi + \pi^2 + \pi^3 + \ldots$$

$$= \frac{\pi}{1 - \pi}.$$

Therefore,

$$\sigma = \sum_{m=1}^{\infty} \frac{(ml)!}{m!^l l^{ml}} \leq \frac{\pi}{1 - \pi}.$$

Second,

$$\sum_{m=1}^{M} \frac{(ml)!}{m!^l l^{ml}} \geq \rho_1^{(Ml)} + \rho_2^{(Ml)} + \ldots + \rho_k^{(Ml)}, \qquad k \leq M.$$

Letting M tend to infinity, we have

$$\sigma = \sum_{m=1}^{\infty} \frac{(ml)!}{m!^l l^{ml}} \geq \rho_1 + \rho_2 + \ldots + \rho_k, \qquad \text{for any } k.$$

Then,

$$\sigma = \sum_{m=1}^{\infty} \frac{(ml)!}{m!^l l^{ml}} \geq \sum_{k=1}^{\infty} \rho_k = \frac{\pi}{1 - \pi}.$$

Hence,

$$\sigma = \frac{\pi}{1 - \pi} \qquad \text{or} \qquad \pi = \frac{\sigma}{1 + \sigma}.$$

When the series (6) diverges, it follows from (7) that the series, $\pi + \pi^2 + \pi^3 + \ldots$, must diverge, and consequently the probability that an equalization will occur, π, is 1. Using Stirling's formula to approximate the general term of the series (6), we have

$$\frac{(ml)!}{m!^l l^{ml}} \sim \frac{(2\pi)^{1/2} (ml)^{ml+1/2} e^{-ml}}{(2\pi)^{l/2} (m^{m+1/2})^l e^{-ml} l^{ml}} = \frac{1}{(2\pi)^{(l-1)/2}} l^{1/2} m^{-(l-1)/2}.$$

The series $\sum_{m=1}^{\infty} m^{-(l-1)/2}$ diverges for $l = 2$ or $l = 3$ and converges for $l \geq 4$. Hence, $\pi = 1$ for $l = 2$ or $l = 3$, and $\pi < 1$ for $l \geq 4$. Using the Euler-MacLaurin summation formula, we find that for $l = 4$, $\pi = .231$.

5. In the random walk representation π is the probability of returning to the origin. We shall now determine the probability p, of reaching any given junction point. This would mean the drawing of $(m + t_1)$ balls of the first kind, $(m + t_2)$ balls of the second kind and, in general, $(m + t_k)$ balls of the k-th kind, where m is an arbitrary nonnegative integer. At least one of the fixed nonnegative integers, $t_1, t_2, \ldots, t_k, \ldots, t_l$, is zero.

Let $t_1 + t_2 + \ldots + t_l = T$. The expected number of times that the point determined by the integers, $t_1, t_2, \ldots, t_l$ is reached in the first n times is

$$\sum_{m=0}^{M} \frac{(ml+T)!}{(m+t_1)!(m+t_2)!\ldots(m+t_l)!l^{ml+T}}$$

where $Ml + T \leqq n < (M+1)l + T$.

On the other hand, it is also

$$r_1^{(n)} + r_2^{(n)} + r_3^{(n)} + \ldots + r_{n+1}^{(n)}$$

where $r_j^{(n)}$ is the expected number of j-th arrivals at the given point in the first n times. Then $r_j^{(n)}$ also represents the probability that a j-th arrival will occur.

Since $\{r_j^{(n)}\}$ is an increasing sequence bounded by 1, it tends to a limit. Let $\lim_{n\to\infty} r_j^{(n)} = r_j$. Let $r_1^{(n)} = p_n$ and $\lim_{n\to\infty} p_n = p$. The probability of reaching the given point j times in the first jn times is $r_j^{(jn)}$. It is greater than or equal to the probability that the point will be reached in the first n times, followed by $(j-1)$ returns to this point during the remaining times. Clearly, the probability of returning to this point $(j-1)$ times is not less than $\rho_{j-1}^{(jn-n)}$.

It follows from this and (2) that

$$r_j^{(jn)} \geqq p_n \rho_{j-1}^{(jn-n)} \geqq p_n \pi_n^{j-1}.$$

On the other hand, it is clear that

$$r_j^{(n)} \leqq p\pi^{j-1}.$$

Passing to the limit in both of these inequalities, we have

$$p\pi^{j-1} \geqq r_j \geqq p\pi^{j-1}.$$

Therefore, $r_j = p\pi^{j-1}$. First,

$$\begin{aligned}\sum_{m=0}^{M} \frac{(ml+T)!}{(m+t_1)!(m+t_2)!\ldots(m+t_l)!l^{ml+T}} &= r_1^{(Ml+T)} + r_2^{(Ml+T)} + \ldots + r_{M+1}^{(Ml+T)} \\ &\leqq r_1 + r_2 + \ldots + r_{M+1} + \ldots \\ &= p + p\pi + p\pi^2 + \ldots \\ &= \frac{p}{1-\pi}.\end{aligned}$$

Hence

$$S = \sum_{m=0}^{\infty} \frac{(ml+T)!}{(m+t_1)!(m+t_2)!\ldots(m+t_l)!l^{ml+T}} \leqq \frac{p}{1-\pi}.$$

Second,

$$\sum_{m=0}^{M} \frac{(ml+T)!}{(m+t_1)!(m+t_2)!\ldots(m+t_l)!l^{ml+T}} \geqq r_1^{(Ml+T)} + r_2^{(Ml+T)} + \ldots + r_k^{(Ml+T)}$$

when $k \leqq M + 1$. Then,

$$S = \sum_{m=0}^{\infty} \frac{(ml+T)!}{(m+t_1)!(m+t_2)!\dots(m+t_l)!l^{ml+T}} \geqq r_1 + r_2 + \dots + r_k$$

for any k. Also

$$S \geqq \sum_{k=1}^{\infty} r_k = \frac{p}{1-\pi} .$$

Therefore

$$S = \frac{p}{1-\pi} \quad \text{or} \quad p = (1-\pi) S = \frac{S}{1+\sigma} .$$

Since

$$\frac{(ml+T)!}{(m+t_1)!(m+t_2)!\dots(m+t_l)!l^{ml+T}} \sim \frac{(ml)!}{m!^l l^{ml}} ,$$

the series

$$\sum_{m=0}^{\infty} \frac{(ml+T)!}{(m+t_1)!(m+t_2)!\dots(m+t_l)!l^{ml+T}}$$

diverges for $l = 2$ or $l = 3$ and converges for $l \geqq 4$. Consequently for $l \geqq 4$, $p < 1$.

There remains the case for $l = 2$ or $l = 3$ in which the above series diverges. We shall show that when $\pi = 1$, $p = 1$ thus taking care of this case.

Let max $(t_k) = t$ and $t_1 + t_2 + \dots + t_l = T$. The probability of reaching the point determined by the nonnegative integers, $(t - t_1)$, $(t - t_2)$, ..., $(t - t_l)$, on the $(lt - T)$-th time is greater than zero. The probability of returning to the origin π is assumed to be 1. Then, the probability of returning to the origin from the point determined by $(t - t_1)$, $(t - t_2)$, ..., $(t - t_l)$ later than the $(lt - T)$-th time must be 1, for in this case there is no return to the origin in the first $(lt - T)$ times. But this probability is the same as the probability p of drawing $(m + t_k)$ balls of the k-th kind, for $k = 1, 2, \dots, l$ and m an arbitrary nonnegative integer. Therefore $p = 1$.

REFERENCES

[1] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Wiley, New York, 1950.

[2] G. PÓLYA, "Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Strassennetz," *Math. Annalen*, Vol. 84 (1921), pp. 149–160.