ON A PROBLEM CONCERNING PROBABILITY AND ITS CONNECTION WITH THE THEORY OF DIFFUSION[1]

M. KAC

Introduction. The present note gives a certain probabilistic approach to the problem of diffusion. The main result is that the classical solution of the differential equation of diffusion is an asymptotic formula for the statistical problem under consideration.

The method which will be used is essentially that of Steinhaus and the present author which they applied to a similar but simpler problem of P. and T. Ehrenfest.[2]

1. **The problem.** Given an infinite sequence of boxes enumerated as follows

$$\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots$$

and N numbered balls which are distributed in a certain way in the boxes, one takes at random one of the numbers $-N, \cdots, -1$, $1, \cdots, N$ which are suppose to be equiprobable and if the number k is drawn one moves the ball number $|k|$ from its original box to the nearest to the right or to the nearest to the left according as sign k was 1 or -1. One repeats this process n times and one asks what is the "probable value" (mathematical expectation) of the "concentration" of balls in the box number s, say. By "concentration" one simply understands the ratio of the number of balls in a certain box and N. It is, of course, understood that the successive drawings are independent in the statistical sense of this word.

2. **Reduction of the problem by means of the "function of choice."** We divide the interval $(0, 1)$ into N equal parts and we define a function on $(0, 1)$ by placing $f(x) = s$ for $l - 1/N < x \leq l/N$ if the ball number l is originally in box s. This function represents the initial state of the schema. The joint length of those intervals in which $f(x) = s$ is obviously the initial "concentration" of balls in the box number s.

Let now $\omega(x)$ (the "function of choice") be 1 for $0 < x \leq 1/N$ and 0 for $1/N < x \leq 1$ and let furthermore $\omega(x+1) = \omega(x)$. Then $f(x) \pm \omega(x - (p-1)/N)$ represents obviously the state of the schema after moving the ball number p from its box to the nearest to the right $(+)$ or to the nearest to the left $(-)$.

Let r be a "random variable" which takes only the values -1, $-(N-1)/N, \cdots, -1/N, 1/N, \cdots, (N-1)/N, 1$ and every one of them with the probability $1/2N$. Then $f(x) + \text{sign } r \cdot \omega(x - |r| + 1/N)$ represents the "random state" after the first drawing. If $r_1, r_2, \cdots, r_n$ are independent "random variables" having the same distribution as r, then

$$\Sigma_n(x, r) = f(x) + \text{sign } r_1\omega(x - |r_1| + 1/N) + \cdots$$
$$+ \text{sign } r_n\omega(x - |r_n| + 1/N)$$

represents the "random state" after n independent drawings. Thus the problem is reduced to evaluating the "probable value" of the joint length of those intervals in which $\Sigma_n(x, r) = s$.

3. Evaluation by means of a "discontinuity factor." The remaining part of the solution is now of a purely technical nature. It is clear that

$$\frac{1}{2\pi} \int_0^{2\pi} \exp\left[i\xi(\Sigma_n(x, r) - s)\right]d\xi$$

is equal to 1 or to 0 according as $\Sigma = s$ or $\Sigma \neq s$, and therefore the joint length of those intervals in which $\Sigma = s$ is

$$L_n(r) = \frac{1}{2\pi} \int_0^1 \int_0^{2\pi} \exp\left[i\xi(\Sigma_n(x, r) - s)\right]d\xi dx.$$

Thus the "probable value" $P_{n,N}(s)$ of $L_n(r)$ is equal to

$$\frac{1}{2\pi} \int_0^1 \int_0^{2\pi} \text{prob. val.}\left\{\exp\left[i\xi(\Sigma_n(x, r) - s)\right]\right\}d\xi dx.$$

It is easily seen that

prob. val. $\left\{\exp\left[i\xi(\Sigma - s)\right]\right\}$

$= \exp\left[i\xi(f(x) - s)\right](\text{prob. val. } \exp\left[(i\xi \text{ sign } r)\omega(x - |r| + 1/N)\right])^n$

$= \exp\left[i\xi(f(x) - s)\right]\left(1 - \dfrac{1}{N} + \dfrac{1}{N} \cos \xi\right)^n$

and finally

$$(1)\quad P_{n,N}(s) = \frac{1}{2\pi} \int_0^1 \int_0^{2\pi} \cos \xi(f(x) - s)\left(1 - \frac{1}{N} + \frac{1}{N} \cos \xi\right)^n d\xi dx.$$

4. A particular case. The formula becomes much simpler if one assumes that $f(x) \equiv 0$, that is, all the balls were initially in the box number 0. In this case one has

$$P_{n,N}(s) = \frac{1}{2\pi} \int_0^{2\pi} \cos s\xi \left(1 - \frac{1}{N} + \frac{1}{N} \cos \xi\right)^n d\xi.$$

Suppose that $N \to \infty$ and that $n/N \to z$; then

$$P_{n,N}(s) \to \frac{e^{-z}}{2\pi} \int_0^{2\pi} \cos s\xi e^{z \cos \xi} d\xi = e^{-z} J_s(iz),$$

where $J_s(z)$ denotes the sth Bessel function.

5. **The asymptotic formula.** Suppose now that the boxes are situated on the infinite line $(-\infty, +\infty)$ and that the distance between the consecutive ones is λ. Suppose furthermore that one makes a drawing every τ seconds so that after the time t one has performed $n = [t/\tau]$ drawings. We shall be interested in evaluating the "probable value" of the "concentration" of balls on the interval (l_1, l_2) after the time t. The answer is given by the probable value of the joint length of those intervals in which $l_1 < \lambda \Sigma_n(x, r) < l_2$ and can be evaluated by means of the formula (1). However, it will be more convenient for our purpose to evaluate the "probable value" under consideration by means of the Dirichlet discontinuity factor. It is well known that

$$\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin \alpha\xi}{\xi} e^{i\gamma\xi} d\xi$$

is equal to 1, $\frac{1}{2}$ or 0 according as $|\gamma| < \alpha$, $|\gamma| = \alpha$ or $|\gamma| > \alpha$. Thus, if neither l_1 nor l_2 is a multiple of λ, the probable value one looks for is given by the following formula:

$$P_N(l_1, l_2, t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin \frac{1}{2}(l_2 - l_1)\xi}{\xi}$$

$$\cdot \exp\left(-i\xi \frac{l_1 + l_2}{2}\right) \left(\int_0^1 \exp\left(i\xi\lambda f(x)\right) dx\right)$$

$$\cdot \left(1 - \frac{1}{N} + \frac{1}{N} \cos \lambda\xi\right)^{[t/\tau]} d\xi.$$

If now $\lambda^2/N\tau \to 2k$ as $N \to \infty$, and $\lambda \to 0$, one has

$$\left(1 - \frac{1}{N} + \frac{1}{N} \cos \lambda\xi\right)^{[t/\tau]} \to \exp\left(-kt\xi^2\right)$$

uniformly in every finite interval and, assuming in addition that the distribution function of $\lambda f(x)$ tends to a distribution function $\sigma(u)$,

one readily deduces from the continuity theorem of Fourier-Stieltjes transforms[3] that

$$P_N(l_1, l_2, t) \to \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\sin \frac{1}{2}(l_2 - l_1)\xi}{\xi}$$

$$\cdot \exp\left(- i\xi \frac{l_1 + l_2}{2}\right) \left(\int_{-\infty}^{+\infty} \exp(i\xi u) d\sigma(u)\right) \exp(- kt\xi^2) d\xi.$$

The integral on the right side can easily be evaluated and its value is

$$\frac{1}{2(\pi kt)^{1/2}} \int_{l_1}^{l_2} \int_{-\infty}^{+\infty} \exp\left\{-\frac{(u - l)^2}{4kt}\right\} d\sigma(u) dt.$$

In the limiting case the concentration is given by the formula

$$\frac{1}{2(\pi kt)^{1/2}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{(u - l)^2}{4kt}\right\} d\sigma(u),$$

which is the classical solution of the differential equation of diffusion in the case of an infinite cylinder. In our case the classical formula turned out to be an asymptotic formula for a simple probabilistic process belonging to the category of the so-called "random walks."

CORNELL UNIVERSITY

---

[3] Cf. E. K. Haviland, American Journal of Mathematics, vol. 56 (1934), pp. 625–658.