A NOTE ON RANDOM WALK

By Herbert T. David

The Johns Hopkins University Institute for Cooperative Research

A random walk is defined as a series of discrete steps along the real line, here denoted by I. Each step is represented by the chance variable X, with sectionally continuous density function $f(x)$. The walk begins at any point a of I, and continues until a step carries us outside some subregion Ω of I. In this note, Ω is taken as a finite interval with upper bound D and lower bound $D - y$. The chance variables N and Z are, respectively, the number of steps required to end the walk, and the endpoint of the walk. The range of Z always excludes Ω.

Below, we define $x = D - a$, and consider $E(N)$ as a function $G(x, y)$ of x and y. Under specified conditions, a differential equation (32) is derived, relating $G(0, y)$ and $G(x, y)$.

Let

$$\psi_1(t) = f(t - a) \tag{1}$$

$$\psi_n(t) = \int \cdots (n-1) \cdots \int \prod_{i=1}^{n-1} f(g_i) \, f\left(t - a - \sum_{j=1}^{n-1} g_j\right) dg_1 \cdots dg_{n-1}; \qquad n > 1 \tag{2}$$

where

$$\left[a + \sum_{j=1}^{i} g_j\right] \epsilon \Omega, \quad \text{for } i: 1, 2, \cdots, n-1.$$

Then

$$P\{Z \epsilon w_1, N = n\} = \int_{w_1} \psi_n(t)\, dt \qquad \text{for} \quad w_1 \epsilon \bar{\Omega}$$

$$P\{Z \epsilon w_2, N = n\} = 0 \qquad \text{for} \quad w_2 \epsilon \Omega.$$

Hence

$$P\{N = n\} = \int_{\bar{\Omega}} \psi_n(t)\, dt$$

$$E(N=)\, i \sum_{i=1}^{\infty} \int_{\bar{\Omega}} \psi_i(t)\, dt. \tag{3}$$

The transformation $[h_i = a + \sum_{j=1}^{i} g_j \,;\, i: 1, \cdots, n-1]$ gives for $\psi_n(t)$ the more convenient expression

$$\psi_n(t) = \int_{\Omega} \cdots (n-1) \cdots \int_{\Omega} f(h_1 - a) \cdot \prod_{i=2}^{n-1} f(h_i - h_{i-1}) f(t - h_{n-1})\, dh_1 \cdots dh_{n-1}. \tag{4}$$

The n-fold integral $\int_I \psi_n(t)\, dt$ is absolutely convergent, hence may be integrated first with respect to t. This gives, keeping the notation of (4)

$$\int_I \psi_n(t)\, dt = \int_\Omega \psi_{n-1}(h_{n-1})\, dh_{n-1}. \tag{5}$$

Assuming that $E(N)$ remains finite for all considered a and Ω, series (3) may be rearranged, giving : $E(N) = \sum_{i=1}^{\infty} B_i$ where

$$B_i = \sum_{j=i}^{\infty} \int_{\bar{\Omega}} \psi_j(t)\, dt.$$

Now, $B_1 = \sum_{i=1}^{\infty} P\{N = i\} = 1$. Also, using (5) and induction on n, it is readily shown that $B_n = \int_\Omega \psi_{n-1}(t)\, dt$, so that

$$E(N) = 1 + \sum_{i=1}^{\infty} \int_\Omega \psi_i(t)\, dt \tag{6}$$

Define transformations $T_n : [g_i = D - h_i, i : 1, \cdots, n - 1; g_n = D - t]$. Substituting expressions (1) and (4) in (6), transform the jth term of the summation by T_j. This gives

$$E(N) = 1 + \sum_{n=1}^{\infty} \int_0^y \cdots (n) \cdots \int_0^y f(x - g_1) \prod_{i=1}^{n-1} f(g_i - g_{i+1})\, dg_1 \cdots dg_n \tag{7}$$

where $x = D - a$.

By (7), $E(N)$ is a function of x and y; hence we write $E(N) \equiv G(x, y)$.

Define:

$M(k)$: Max $f(t)$ for $|t| \leq k$.

K : Any number satisfying $K \leq [1 - \epsilon]/M(K)$.

R : Any region $[-\infty < x < \infty; 0 \leq y \leq K]$.

M : Max $f(t)$.

L : Any number satisfying $L \leq [1 - \epsilon]/M$.

R' : Any region $[-\infty < x < \infty; 0 \leq y \leq L]$.

In the ensuing argument, we shall assume that

$$(x, y) \in R. \tag{8}$$

This condition restricts certain one-dimensional and two-dimensional variables to regions over which some infinite series are uniformly convergent with respect to these variables. Uniform convergence is required to validate term-by-term differentiations and integrations, and to establish the continuity in one or two variables of certain functions represented by series.

Arguments dealing with the solution of integral equations (17), (20) and (25) are valid only under the more restrictive condition

(9) $$(x, y) \in R'$$

this being the general sufficiency condition for the existence of solutions. However, (17) and (20) enter the argument with respect only to the derivation of equation (21) which could have been derived, though in a more cumbersome manner, by a term by term comparison of the series expressions for $[\lambda_{01}(x, y)]$ $[G(y, y)]$ and for $[G_{01}(x, y)]$ $[\lambda(y, y)]$, this latter approach being valid under (8). Similarly, (25) is used only in obtaining (27), which could have been obtained by a direct manipulation of the series expression for $G(x, y)$, this approach also being valid under (8). Hence, all subsequent derivations hold, as long as $(x, y) \in R$

By (8), we may interchange summation and integration with respect to g_1 in (7). This gives

(10) $$G(x, y) = 1 + \int_0^y f(x - g)G(g, y)\, dg.$$

(11) Assume that $f(t)$ has a continuous derivative everywhere

Then $f(t)$ is continuous and $G(x, y)$ is continuous by (7) and (8). Hence

(12) $f(x - g)G(g, y)$ and $d/dx\, f(x - g)G(g, y)$ are continuous in (x, g)

(13) $f(x - g)G(g, y)$ is continuous in (g, y).

Let $G_{ij}(x, y)$ denote

$$\frac{d^i}{dx^i} \frac{d^j}{dy^j} G(x, y).$$

Then, by (12), we may differentiate (10) with respect to x, and, since $f_{10}(x - g) = -f_{01}(x - g)$, an integration by parts yields

(14) $$G_{10}(x, y) = f(x)G(0, y) - f(x - y)G(y, y) + \int_0^y f(x - g)G_{10}(g, y)\, dg.$$

Further, under (8), $G_{01}(x, y)$ may be obtained by differentiating (7) term by term, and is continuous in (x, y). Hence, $f(x - g)G_{01}(g, y)$ is continuous in (g, y), and we may differentiate (10) with respect to y, giving

(15) $$G_{01}(x, y) = f(x - y)G(y, y) + \int_0^y f(x - g)G_{01}(g, y)\, dg.$$

Adding (14) to (15), dividing by $G(0, y)$ which is always greater or equal to 1, and letting

(16) $$\lambda(x, y) = [G_{10}(x, y) + G_{01}(x, y)]/G(0, y)$$

we obtain

(17) $$\lambda(x, y) = f(x) + \int_0^y f(x - g)\lambda(g, y)\, dg.$$

Under (9), (17) defines a function

$$\lambda(x, y) = f(x) + \sum_{n=1}^{\infty} \int_0^y \cdots (n) \cdots \int_0^y f(x - g_1) \cdot \prod_{i=1}^{n-1} f(g_i - g_{i+1}) f(g_n)\, dg_1 \cdots dg_n. \tag{18}$$

By (8), this function is continuous in (x, y) and may be differentiated term by term with respect to y. Further, $\lambda_{01}(x, y)$ thus gotten is continuous in (x, y), so that $f(x - g)\lambda_{01}(g, y)$ is continuous in (g, y). Hence, (17) may be differentiated with respect to y, giving

$$\lambda_{01}(x, y) = f(x - y)\lambda(y, y) + \int_0^y f(x - g)\lambda_{01}(g, y)\, dg. \tag{19}$$

Since, under (9), the integral equation

$$\alpha(x, y) = f(x - y) + \int_0^y f(x - g)\alpha(g, y)\, dg \tag{20}$$

has a unique continuous solution for every fixed y, (15) and (19) give

$$\frac{\lambda_{01}(x, y)}{\lambda(y, y)} = \frac{G_{01}(x, y)}{G(y, y)}. \tag{21}$$

Hence

$$\frac{\int_0^y \lambda_{01}(x, y)\, dx}{\lambda(y, y)} = \frac{\int_0^y G_{01}(x, y)\, dx}{G(y, y)}$$

and

$$\frac{\frac{d}{dy} \int_0^y \lambda(x, y)\, dx}{\lambda(y, y)} = \frac{\frac{d}{dy} \int_0^y G(x, y)\, dx}{G(y, y)}. \tag{22}$$

$$\text{Let } f(t) = f(-t). \tag{23}$$

Then it is obvious from the definition that

$$G(0, y) = G(y, y). \tag{24}$$

Further, by (15),

$$\frac{G_{01}(x, y)}{G(y, y)} = f(x - y) + \int_0^y f(x - g)\, \frac{G_{01}(g, y)}{G(y, y)}\, dg \tag{25}$$

so that, under (9), (25) gives for $G_{01}(x, y)/G(y, y)$ the unique expression

$$f(x - y) + \sum_{n=1}^{\infty} \int_0^y \cdots (n) \cdots \int_0^y f(x - g_1) \prod_{i=1}^{n-1} f(g_i - g_{i+1}) f(g_n - y)\, dg_1 \cdots dg_n$$

which, by (23), is equal to

$$f(y - x) + \sum_{n=1}^{\infty} \int_0^y \cdots (n) \cdots \int_0^y f(y - g_n) \prod_{i=1}^{n-1} f(g_{i+1} - g_i) f(g_1 - x)\, dg_1 \cdots dg_n.$$

Since, under (8), we may interchange summation and integration with respect to x, it follows that

$$\begin{aligned}\int_0^y \frac{G_{01}(x, y)}{G(y, y)}\, dx = \int_0^y f(y - x)\, dx + \sum_{n=1}^{\infty} \int_0^y \cdots (n + 1) \cdots \\ \cdot \int_0^y f(y - g_n) \prod_{i=1}^{n-1} f(g_{i+1} - g_i) f(g_i - x)\, dg_1 \cdots dg_n\, dx\end{aligned} \tag{26}$$

which, by a change of integration indices and a referral to (7), is seen to equal $[G(y, y) - 1]$. (26) thus gives

$$\int_0^y G_{01}(x, y)\, dx = G(y, y)[G(y, y) - 1]. \tag{27}$$

Further, by (16), (24), and (27),

$$\int_0^y \lambda(x, y)\, dx = G(0, y) - 1 \tag{28}$$

so that

$$\frac{d}{dy} \int_0^y \lambda(x, y)\, dx = \frac{d}{dy} G(0, y) \tag{29}$$

while (24) and (27) also yield

$$\frac{d}{dy} \int_0^y G(x, y)\, dx = [G(0, y)]^2. \tag{30}$$

Hence, by (22), (29), and (30),

$$\lambda(y, y) = \frac{d}{d_y} G(0, y)/G(0, y). \tag{31}$$

Finally, substituting (31) in (21), and remembering the definition of λ given in (16), we get, using (24),

$$G(0, y)[G_{11}(x, y) + G_{02}(x, y)] = \frac{d}{dy} G(0, y)[G_{10}(x, y) + 2G_{01}(x, y)]. \tag{32}$$

The conditions under which (32) holds are, in summary, (8), (11), and (23). If $f(t)$ has an expansion

$$f(t) = \sum_{i=0}^{\infty} A_i t^i; \qquad |t| < T \tag{33}$$

it is clear from (7) that

$$G(x, y) = \sum_{i,j=0}^{\infty} B_{ij} x^i y^j \tag{34}$$

for $(x, y) \in S$, where $S : [T_0 \leq x \leq T_1; 0 \leq y \leq T_1 + T_0]$; $T_0 \leq 0$, $T_1 < T$.

Substituting (34) in (32), and equating coefficients of like powers of (x, y), we obtain the recursion formulae

$$(35)\quad \sum_{j+k=n} B_{ij} B_{0k}[j][2k-j+1] = \sum_{j+k=n-1} B_{i+1,j} B_{0k}[i+1][j-k]; \quad i: 0, 1, \cdots.$$

From (10), it is readily verified that $B_{i0} = 0$ for $i \neq 0$, so that equations (35) give solutions for the B_{ij} in terms of the B_{0k}. These solutions are of interest since they show a one-to-one correspondence between the functions $G(0, y)$ and $G(x, y)$, for $(x, y) \in [R \cap S]$.

NUMERICAL INTEGRATION FOR LINEAR SUMS OF EXPONENTIAL FUNCTIONS

By Robert E. Greenwood

The University of Texas and the Institute for Numerical Analysis[1]

1. Introduction. The methods of numerical integration going by the names trapezoidal rule, Simpson's rule, Weddle's rule, and the Newton-Cotes formulae are of the type

$$(1)\qquad \int_{-1}^{1} f(x)\,dx \simeq \sum_{i=0}^{n} \lambda_{in} f(x_{in})$$

where the abscissae $\{x_{in}\}$ are uniformly distributed on a finite interval, chosen as $(-1, 1)$ for convenience,

$$(2)\qquad x_{in} = -1 + \frac{2i}{n}, \qquad i = 0, 1, 2, \cdots, n,$$

and where the set of constants $\{\lambda_{in}\}$ depend on the name of the rule and the value of n but not on the function $f(x)$. Throughout this note all abscissae will be assumed to be uniformly distributed on $(-1, 1)$ unless the contrary is explicitly stated.

Since correspondence relation (1) involves $(n + 1)$ constants $\{\lambda_{in}\}$, it might be possible to choose $(n + 1)$ arbitrary functions $g_j(x)$, $j = 0, 1, 2, \cdots, n$, and require that the set $\{\lambda_{in}\}$ be the solution, if such exists, of the $(n + 1)$ simultaneous linear equations

$$(3)\qquad \int_{-1}^{1} g_j(x)\,dx = \sum_{i=0}^{n} \lambda_{in} g_j(x_{in}), \qquad j = 0, 1, 2, \cdots, n.$$

Indeed, the selection

$$(4)\qquad g_j(x) = x^j, \qquad j = 0, 1, 2, \cdots, n,$$

will give a set of $(n + 1)$ simultaneous equations of form (3) and the solution $\{\lambda_{in}\}$ is the set of Newton-Cotes weights for that value of n. The numerical evaluation

[1] This work was performed with the financial support of the Office of Naval Research of the Navy Department.