Canad. Math. Bull. Vol. **22** (3), 1979

HITTING TIME DISTRIBUTIONS WHEN $\sum X_k/d^k$ HAS A SMOOTH DENSITY

BY

DUDLEY PAUL JOHNSON

ABSTRACT. In this paper we construct the hitting time distributions for stochastic processes X_k, taking on values amongst the integers $0, 1, \ldots, d-1$ for which $\sum_{k=1}^{\infty} X_k/d^k$ has a smooth polynomial density with respect to the Lebesgue measure on $[0, 1]$.

Suppose that X_k, $k = 0, 1, 2, \ldots$ is a stochastic process on the integers $0, 1, \ldots, d-1$. Clearly the distribution of the stochastic process is uniquely determined by the distribution of the single random variable $Y = \sum_{k=0}^{\infty} X_k/d^k$ so long as the probability that $X_{n+k} = d-1$ for all k is zero for each $n \geq 0$. That is to say, one can generate the stochastic process by picking a point Y from the interval $[0, 1]$ according to a fixed distribution and then letting X_k be the kth decimal in the d-adic expansion of Y. In this paper we prove the following Theorem.

THEOREM. *Suppose that X_k, $k = 0, 1, 2, \ldots,$ is a stochastic process taking on values amongst the integers $0, \ldots, d-1$. Suppose that U is a subset of $0, \ldots, d-1$ with d_0 elements, that $a(\cdot)$ is a function mapping U^c into the reals and τ_U is the time that X_k first leaves U.*

I. If $\sum_{k=0}^{\infty} X_k/d^k$ has a density $f(x) = \sum_{n=0}^{\infty} c_n x^n$, absolutely convergent in $[0, 1]$, then

$$Ea(X_{\tau_U}) = \sum_{n=0}^{\infty} a_n c_n$$

where

$$a_0 = \frac{1}{d-d_0} \sum_{i \notin U} a(i)$$

and in general a_j can be found from $a_0, \ldots, a_{j-1}$ via

$$a_j = [1 - d_0 d^{-j-1}]^{-1} \left\{ \sum_{i \notin U} \frac{1}{j+1} \left[\left(\frac{i+1}{d} \right)^{j+1} - \left(\frac{i}{d} \right)^{j+1} \right] a(i) + d^{-j-1} \sum_{k=0}^{j-1} \binom{j}{k} \sum_{i \in U} i^{j-k} a_k \right\}.$$

II. If $\sum_{k=0}^{\infty} X_k/d^k$ has a density $f(x) = \sum_{n=0}^{\infty} \alpha_n \cos 2\pi d^n x + \beta_n \sin 2\pi d^n x$

Received by the editors February 17, 1978 and, in revised form, September 14, 1978.

https://doi.org/10.4153/CMB-1979-035-6 Published online by Cambridge University Press

then

$$Ea(X_\tau)=\sum_{n=0}^{\infty}\alpha_n a_n+\beta_n b_n$$

where

$$a_0=E\cos 2\pi X_\tau, \qquad b_0=E\sin 2\pi X_\tau$$

and in general a_n, b_n can be found from a_{n-j}, b_{n-j} via

$$a_n=\sum_{i\notin U}a(i)\int_{i/d}^{(i+1)/d}\cos(2\pi x\,d^j)\,dx+\left(d^{-1}\sum_{i\in U}\cos 2\pi i\,d^{n-1}\right)a_{n-1}$$
$$-\left(d^{-1}\sum_{i\in U}\sin 2\pi i\,d^{n-1}\right)b_{n-1}$$

$$b_n=\sum_{i\notin U}a(i)\int_{i/d}^{(i+1)/d}\sin(2\pi x\,d^j)\,dx+\left(d^{-1}\sum_{i\in U}\sin 2\pi i\,d^{n-1}\right)a_{n-1}$$
$$+\left(d^{-1}\sum_{i\in U}\cos 2\pi i\,d^{n-1}\right)b_{n-1}.$$

Proof. The stochastic process X_k, $k=0,1,2,\ldots$ induces a measure φ in the Banach space $\mathcal{M}(\Omega,\mathcal{F})$ of all bounded measures on the measurable space $(\Omega,\mathcal{F})$ of all functions ω mapping the nonnegative integers into the set $\{0,\ldots,d-1\}$ with φ defined on the cylinder sets $[X_0=i_0,\ldots,X_N=i_N]$, $X_k(\omega)=\omega(k)$, generating the σ-field $\mathcal{F}$, via

$$\varphi[X_0=i_0,\ldots,X_N=i_N]=P[X_0=i_0,\ldots,X_N=i_N].$$

We define the linear operators T and $E(i)$, $i=0,\ldots,d-1$ and the linear functional p^* on $\mathcal{M}(\Omega,\mathcal{F})$ via

$$T\psi[X_o=i_o,\ldots,X_N=i_N]=\psi[X_1=i_o,\ldots,X_{N+1}=i_N]$$
$$E(i)\psi[\Lambda]=\psi[X_o=i,\Lambda]$$
$$p^*\psi=\psi(\Omega);$$

and then we let Φ be the smallest linear subspace of $\mathcal{M}(\Omega,\mathcal{F})$ which contains φ and is invariant under the operators T and $E(i)$, $i=0,\ldots,d-1$. The collection (Φ,T,E,p^*) will be called the algebraic representation of the stochastic process X_k, $k=0,1,\ldots$. We will use the algebraic representation to find $Ea(X_{\tau_U})$ by showing that the linear functional p^* on Φ, defined by

$$a^*\varphi=\int a(X_{\tau_U})\,d\varphi,$$

is a solution of the linear equations

$$(I^*-T^*)a^*=0 \quad \text{on} \quad E(i)\Phi, \qquad i\in U$$
$$a^*=a(i)p^* \quad \text{on} \quad E(i)\Phi, \qquad i\notin U,$$

where the second equation is obvious and the first holds since for $\varphi \in E(i)\Phi$, $i \in U$ we have

$$\begin{aligned}(I^*-T^*)a^*\varphi &= a^*(I-T)\varphi \\ &= \int a(X_{\tau_U})\, d\varphi - \int a(X_{\tau_U})\, dT\varphi \\ &= \int a(X_{\tau_U})\, d\varphi - \int a(X_{\tau_U(\omega_1^+)}(\omega_1^+))\varphi(d\omega) \\ &= \int [a(X_{\tau_U(\omega)}(\omega)) - a(X_{\tau_U(\omega_1^+)}(\omega_1^+)]\varphi(d\omega) \\ &= 0\end{aligned}$$

since $\omega(0) \in U$ and so $X_{\tau_U(\omega)}(\omega) = X_{\tau_U(\omega_1^+)}(\omega_1^+)$.

Now in our case, each $\omega \in \Omega$ can be identified with a real number $x = x(\omega) = \sum_{k=0}^{\infty} \omega(k)/d^k \in [0, 1]$ so that each $\psi \in \Phi$ can be identified with a density f on $[0, 1]$. This generates a linear space $\mathscr{X}$ of densities on $[0, 1]$. Since $T\psi$ will be identified with the density $d^{-1}\sum_{k=0}^{d-1} f((k+x)/d)$ and $E(i)\psi$ with the density $I_{[i/d,(i+1)/d]}(x)f(x)$, where I_U is the indicator function of U; it follows that the algebraic representation (Φ, T, E, p^*) is isomorphic to the collection (I, T, E, p^*) where

(i) $\mathscr{X}$ is a linear space of densities on $[0, 1]$.
(ii) $Tf = d^{-1}\sum_{k=0}^{d-1} f((k+x)/d)$
(iii) $E(i)f = I_{[i/d,(i+1)/d]}(x)f(x)$
(iv) $p^*f = \int_0^1 f(x)\, dx$.

Thus we can find the linear functional a^* on Φ by looking for a linear functional a^* on $\mathscr{X}$ which satisfies

$$\begin{aligned}(I^*-T^*)a^* &= 0 \quad \text{on} \quad E(i)\mathscr{X}, \qquad i \in U \\ a^* &= a(i)p^* \quad \text{on} \quad E(i)\mathscr{X}, \qquad i \notin U.\end{aligned}$$

In case I of our theorem we take as a basis for $\mathscr{X}$ the functions f_{ij}, where $i = 0, \ldots, d-1$ and $j = 0, 1, 2 \ldots$, defined by

$$f_{ij}(x) = I_{[i/d,(i+1)/d]}(x)x^i.$$

Then, letting

$$f_k(x) = x^k = \sum_{i=0}^{d-1} f_{ik}(x),$$

we have

$$E(k)f_{ij} = \begin{cases} f_{ij}, & i = k \\ 0, & i \neq k \end{cases}$$

https://doi.org/10.4153/CMB-1979-035-6 Published online by Cambridge University Press

and

$$\begin{aligned} Tf_{ij}(x) &= d^{-1}\sum_{k=0}^{d-1} f_{ij}\left(\frac{k+x}{d}\right) \\ &= d^{-1}\sum_{k=0}^{d-1} I_{[i/d,(i+1)/d]}\left(\frac{k+x}{d}\right)\left(\frac{k+x}{d}\right)^{j} \\ &= d^{-1}\left(\frac{i+x}{d}\right)^{j} \\ &= d^{-j-1}\sum_{k=0}^{j}\binom{j}{k} i^{j-k}x^{k} \end{aligned}$$

and

$$p^*f_{ij} = \int_0^1 f_{ij}(x)\,dx = \frac{1}{j+1}\left[\left(\frac{i+1}{d}\right)^{j+1} - \left(\frac{i}{d}\right)^{j+1}\right].$$

Thus the linear functional a^* on $\mathscr{X}$ satisfies

$$\begin{aligned} a^*f_{ij} &= a^*Tf_{ij} = d^{-j-1}\sum_{k=0}^{j}\binom{j}{k} i^{j-k}a^*f_k, \qquad i \in U \\ a^*f_{ij} &= a(i)p^*f_{ij}, \qquad i \notin U. \end{aligned}$$

Summing over i now gives us

$$a^*f_j = \sum_{i\notin U}\frac{1}{j+1}\left[\left(\frac{i+1}{d}\right)^{j+1} - \left(\frac{i}{d}\right)^{j+1}\right]a(i) + a^{-j-1}\sum_{k=0}^{j}\binom{j}{k}\sum_{i\in U} i^{j-k}a^*f_k.$$

For $j=0$ this last equation becomes

$$a^*f_o = \frac{1}{d-d_o}\sum_{i\notin U} a(i).$$

For $j>0$ we have

$$a^*f_j = [1-d_o d^{-j-1}]^{-1}\left\{\sum_{i\notin U}\frac{1}{j+1}\left[\left(\frac{i+1}{d}\right)^{j+1} - \left(\frac{i}{d}\right)^{j+1}\right]a(i) + d^{-j-1}\sum_{k=0}^{j-1}\binom{j}{k}\sum_{i\in U} i^{j-k}a^*f_k\right\}.$$

Thus if

$$f(x) = \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} c_n f_n(x)$$

https://doi.org/10.4153/CMB-1979-035-6 Published online by Cambridge University Press

then by the dominated convergence theorem we have

$$\begin{aligned} a^*f &= \int f(X_\tau)\, d\varphi_f \\ &= \int \sum_{n=0}^{\infty} c_n f_n(X_\tau)\, d\varphi_f \\ &= \sum_{n=0}^{\infty} c_n \int f_n(X_\tau)\, d\varphi_f \\ &= \sum_{n=0}^{\infty} c_n a^* f_n \\ &= \sum_{n=0}^{\infty} c_n a_n, \end{aligned}$$

where $a_n = a^* f_n$ and φ_f is the measure on $(\Omega, \mathcal{F})$ induced by f. This concludes the proof of I.

To prove II, we take as our basis for I the functions

$$f_{ij}(x) = I_{[i/d,(i+1)/d]}(x) \cos 2\pi x\, d^j$$
$$g_{ij}(x) = I_{[i/d,(i+1)/d]}(x) \sin 2\pi x\, d^j$$

where $i = 0$, $d-1$ and $j = 0, 1, 2, \ldots$. Letting $f_j = \sum_{i=0}^{d-1} f_{ij} = \cos 2\pi x d^j$ and $g_j = \sum_{i=0}^{d-1} g_{ij} = \sin 2\pi x d^j$, we have

$$\begin{aligned} Tf_{ij}(x) &= d^{-1} \sum_{k=0}^{d-1} f_{ij}\left(\frac{k+x}{d}\right) \\ &= d^{-1} \sum_{k=0}^{d-1} I_{[i/d,(i+1)/d]}\left(\frac{k+x}{d}\right) \cos 2\pi\left(\frac{k+x}{d}\right) d^j \\ &= d^{-1} \cos 2\pi(id^{j-1} + xd^{j-1}) \\ &= d^{-1} \cos 2\pi id^{j-1} \cos 2\pi xd^{j-1} - d^{-1} \sin 2\pi id^{j-1} \sin 2\pi xd^{j-1} \\ &= d^{-1} \cos 2\pi\, id^{j-1} f_{j-1}(x) - (d^{-1} \sin 2\pi id^{j-1}) g_{j-1}(x). \end{aligned}$$

Similarly,

$$Tg_{ij}(x) = (d^{-1} \sin 2\pi id^{j-1}) f_j(x) + (d^{-1} \cos 2\pi id^{j-1}) g_j(x).$$

Thus a^* is a solution of

$$\begin{aligned} (I^* - T^*) a^* f_{ij} &= 0, & i &\in U \\ a^* f_{ij} &= a(i) p^* f_{ij}, & i &\notin U \end{aligned}$$

and

$$\begin{aligned} (I^* - T^*) a^* g_{ij} &= 0, & i &\in U \\ a^* g_{ij} &= a(i) p^* g_{ij}, & i &\notin U. \end{aligned}$$

https://doi.org/10.4153/CMB-1979-035-6 Published online by Cambridge University Press

which becomes

$$a^*f_{ij}=(d^{-1}\cos 2\pi i d^{j-1})f_{j-1}-(d^{-1}\sin 2\pi i d^{j-1})g_{j-1}, \qquad i\in U$$

$$a^*f_{ij}=a(i)\int_0^1 I_{[i/d,(i+1)/d]}(x)\cos(2\pi x d^j)\,dx, \qquad i\notin U$$

and

$$a^*g_{ij}=(d^{-1}\sin 2\pi i d^{j-1})f_{j-1}+(d^{-1}\cos 2\pi i d^{j-1})g_{j-1}, \qquad i\in U$$

$$a^*g_{ij}=a(i)\int_0^1 I_{[i/d,(i+1)/d]}(x)\sin(2\pi x d^j)\,dx, \qquad i\notin U.$$

Summing over i now yields

$$a^*f_j=\sum_{i\notin U}a(i)\int_{i/d}^{(i+1)/d}\cos(2\pi x d^j)\,dx+d^{-1}\Big(\sum_{i\in U}\cos 2\pi i d^{j-1}\Big)f_{j-1}-d^{-1}\Big(\sum_{i\in U}\sin 2\pi i d^{j-1}\Big)g_{j-1}$$

$$a^*g_j=\sum_{i\notin U}a(i)\int_{i/d}^{(i+1)/d}\sin(2\pi x d^j)\,dx+d^{-1}\Big(\sum_{i\in U}\sin 2\pi i d^{j-1}\Big)f_{j-1}+d^{-1}\Big(\sum_{i\in U}\cos 2\pi i d^{j-1}\Big)g_{j-1}.$$

Thus if

$$f(x)=\sum_{n=0}^{\infty}\alpha_n\cos 2\pi d^n x+\beta_n\sin 2\pi d^n x,$$

then, as in *I*, we have

$$a^*f=\sum_{n=0}^{\infty}\alpha_n a_n+\beta_n b_n$$

when $a_n=a^*f_n$ and $b_n=a^*g_n$ thus concluding the proof of II and the theorem.

References

1. N. Dunford and S. Schwartz (1958). *Linear Operators I: General theory*. Pure and Appl. Math., Vol. **7,** Interscience, New York.
2. Dudley Paul Johnson (1970). *Markov Process representations of general stochastic processes*. Proc. Amer. Math. Soc. **24,** 735–738.
3. Dudley Paul Johnson (1972). *Hitting time distributions for general stochastic processes*, Ann. Math. Statist. **43** (1972), 1716–1718.
4. Dudley Paul Johnson (1974). *Representations and classifications of stochastic processes*. Trans. Amer. Math. Soc. **188** (1974), 179–197.
5. P. A. Meyer (1966). *Probability and potentials*. Blaisdell, Waltham, Mass.

Department of Mathematics
University of Calgary
Calgary, Alberta T2N 1N4