Autoregressive Processes and First-Hit Probabilities for Randomized Random Walks

by

Raisa Epstein Feldman

University of California, Berkeley
and University of California, Santa Barbara

Technical Report No. 212
July 1989

Research was supported by ONR Grant No. N00014/89/J/1870

Department of Statistics
University of California
Berkeley, California 94720

Autoregressive Processes and First-Hit Probabilities for Randomized Random Walks

by

Raisa Epstein Feldman[1]

University of Calif. at Berkeley
and University of Calif. at Santa Barbara

Abstract

We find the first-hit distributions of symmetric randomized random walks on $\mathcal{R}$ with exponential lifetime using prediction formulas for Gaussian stationary autoregressive processes, associated with the random walks.

July 1989

AMS 1980 subject classifications: Primary, 60J15, 62M20; secondary, 62M15, 60G15
Key words and phrases: randomized random walk, autoregressive processes, prediction of stationary processes, first-hit probabilities of random walk
[1]This research was supported by ONR Grant No. N00014/89/J/1870

1 Introduction

A randomized random walk is a continuous-time stochastic process which is obtained from an ordinary random walk by specifying that the epochs of the jumps are regulated by a Poisson process. We consider only symmetric processes with integer values, i.e. the jumps themselves are random variables assuming values $\pm 1, \ldots, \pm k$ with probabilities $p_1/2, \ldots, p_k/2$, independent of each other and of the Poisson process. The lifetime of this time-dependent process if regulated by an exponentially distributed random variable independent of the process.

The classic book of Feller (1968 and 1971) devotes substantial space to the topic of random walks (c.f. chapters III and XII of Feller (1968) and chapters II, XII, XIV, XVII of Feller (1971)). Problems connected with first passages, path maxima, hitting points, etc., are treated using combinatorial methods and Laplace and Fourier transforms. The study of randomized random walks, however, is more restricted and the treatment (Feller (1971), §II.7 and XIV.6) is less developed.

In this paper we find first-hit distributions of symmetric randomized random walks with exponential lifetimes (i.e. if $X_u, u \geq 0$ is the process described above and τ is the time of first entry of a set S by X when $X_0 \in S$, then the probabilities of interest are $\{P(X_\tau = s),\ s \notin S\}$). We use a rather unusual method based on prediction of time series and on the theory of Dynkin, which relates Markov processes X with Gaussian processes Φ. Specifically, we use the prediction formula of Dynkin (1980) which expresses the coefficients of the best least square predictor of Φ in terms of the first-hit distribution of X. (See formula (2) in Section 2 below).

Ylvisaker (1987) first proposed use of this theoretical result for practical applications, in his case, prediction and design of Gaussian random fields through simulation of underlying Markov processes. We in some sense invert his idea. As noted by Ylvisaker, if X is a randomized symmetric random walk with exponential lifetime, the Gaussian process Φ associated with X is an autoregressive sequence. This means that the coefficients of the least square estimate of Φ can be found through well-known time series algorithms; these coefficients are the first-hit probabilities of X, according to Dynkin's prediction formula.

In section 2 we review Dynkin's theory of Gaussian processes associated with Markov. In section 3 we discuss the AR(k) sequences which arise from association with randomized random walks. In section 4 we get explicit formulas for first-hit probabilities of random walks and compare our results with the ones in Feller (1968, 1971).

2 Gaussian Processes Associated with Markov Processes

Let us review the results of Dynkin (1980). Here, we follow the paper of Ylvisaker (1987). Let $\hat{X}_u, u \geq 0$ be a symmetric homogeneous right Markov process with the state space $T \subseteq R$ and transition density $p_u(t,s) = p_u(s,t), s,t \in T$. This means that for all Borel $B \subseteq R$

$$P(\hat{X}_u \in B | \hat{X}_{u_1} = t_1, \ldots, \hat{X}_{u_n} = t_n, 0 \leq u_1 < \ldots < u_n < u < \zeta)$$

$$= P(\hat{X}_u \in B | \hat{X}_{u_n} = t_n, 0 \leq u_n < u < \zeta) = \int_B p_{u-u_n}(t_n, t)\, dt.$$

Let ζ be an exponential random variable with mean $\theta, \theta > 0$, independent of $\hat{X}$, which determines the lifetime of $\hat{X}$. Denote by X the "killed version" of $\hat{X}$. This means that $X_u = \hat{X}_u$, $u < \zeta$ and $X_u = \Delta$("the cemetary"), $u \geq \zeta$. Then $X_u, u \geq 0$ is a Markov process with transition density $e^{-\theta u} p_u(t,s), t,s \in T$ and the Green's function

$$g(t,s) = \int_0^\infty e^{-\theta u} p_u(t,s) du \equiv g(s,t). \tag{1}$$

(The function g (as well as the process X) depends on the parameter θ, but we omit the subscript θ for notational simplicity). The function g is the average time spent by X at s when started at t and the "killing" of the process keeps this quantity finite for almost all s for every t. We shall assume throughout that g is finite for all s and t in T. Using the Chapman-Kolmogorov identity and the symmetry of p_u we have that g must be non-negative definite: for any real constants $c_1, \ldots, c_n$

$$\sum c_i c_j g(t_i, s_j) = \sum c_i c_j \int_0^\infty e^{-\theta u} p_u(t_i, s_j) du$$

$$= \sum c_i c_j \int_0^\infty e^{-\theta u} \int_R p_{u/2}(t_i, r) p_{u/2}(r, s_j)\, dr\, du$$

$$= \int_0^\infty e^{-\theta u} \int_R \left(\sum c_i p_{u/2}(t_i, r)\right)^2 dr\, du \geq 0$$

This means that there exists a Gaussian process $\Phi_t, t \in T$ whose parameter space T is the state space of X, with mean zero and covariance function $g(t,s)$. Following Dynkin, we will call Φ *a Gaussian process associated with a Markov process* X. (In Ylvisaker's terminology, Φ is called the G-MAP.) For these Gaussian processes Φ, here is a special case of Dynkin's theorem:

Theorem. *(Dynkin(1980)): Let p_t^S be the first-hit distribution of X on a countable subset S of T starting from the state $t \in T \backslash S$:*

$p_t^S(s) = P(X_{\tau_S} = s | X_0 = t)$, where $\tau_S = \inf\{u \geq 0 : X_u \in S\}$ and $s \in S$.

Then

$$E(\Phi_t | \Phi_s, s \in S) = \sum_{s \in S} p_t^S(s) \Phi_s \tag{2}$$

It is clear that not every Gaussian process can be created in the above described way. We notice that Φ always has a positive correlation function and the coefficients in the prediction formula (2) satisfy

$$p_t^S(s) \geq 0; \quad \sum_{s \in S} p_t^S(s) \leq 1. \tag{3}$$

For further discussion of the relationship between X and Φ, and examples, see Ylvisaker (1987), Dynkin(1980), and Adler and Epstein(1987).

3 Randomized Random Walks and AR Processes

We will concentrate on the special case of stationary Gaussian processes. A zero-mean Gaussian process is stationary if and only if its covariance function depends only on the difference between its arguments:

$$g(t, s) = g(t - s) = g(s - t) \tag{4}$$

Thus, the transition density of the Markov process X associated with Φ satisfies

$$p_u(t, s) = p_u(t - s) = p_u(s - t) \tag{5}$$

According to Blumenthal and Getoor (1968), p. 17, this condition implies that X is a Levy process, i.e. a Markov process with independent stationary increments. Since the Fourier transform of p_u is known for such processes we immediately have:

Proposition. *A Gaussian process Φ_t, $t \in T \subseteq R$, associated with a Markov process X_u, $u \geq 0$, is stationary if and only if its spectral density $f(\lambda)$ is given by the following expression:*

$$f(\lambda) = 2\pi^{-1}\left(\theta - i\lambda\beta + \sigma^2\lambda^2/2 + \int_R \left(1 - i\lambda x/(1 + x^2) - e^{-i\lambda x}\right) \nu(dx)\right)^{-1} \tag{6}$$

where ν is a Levy measure on $R\backslash\{0\}$, $\int_R x^2/(1 + x^2)\nu(dx) < \infty$, $\sigma^2 \geq 0$, $-\infty < \beta < \infty$.

Proof. For a Markov process X_u, $u \geq 0$ with independent stationary increments:

$$\hat{p}_u(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda x} p_u(x) dx = Ee^{i\lambda X_u} = e^{u\psi(\lambda)} \tag{7}$$

with

$$\psi(\lambda) = -i\lambda\beta - \sigma^2\lambda^2/2 - \int_R \left(1 + i\lambda x/(1+x^2) - e^{-i\lambda x}\right) \nu(dx) \tag{8}$$

(see Blumenthal and Getoor(1968), p.18). The relationship between the spectral density f and the covariance g is

$$f(\lambda) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-i\lambda x} g(x)\, dx,$$

and from (1) and (7) we have:

$$\begin{aligned} f(\lambda) &= (2\pi)^{-1} \int_0^{\infty} e^{-\theta u} e^{u\psi(-\lambda)} du \\ &= (2\pi)^{-1}(\theta - \psi(-\lambda))^{-1}; \end{aligned}$$

for the discrete case, replace integrals with sums. Put (8) into the last formula and we are done.

Corollary. *The Gaussian process associated with Brownian Motion is an Ornstein-Uhlenbeck process with spectral density*

$$f(\lambda) = (2\pi)^{-1}(\theta + \sigma^2\lambda^2/2)^{-1}.$$

According to a result of Dynkin (1980), a Gaussian process Φ associated with a Markov process X is itself Markovian if X is continuous. Since there exists only one continuous Markov process with stationary independent increments—Brownian motion, we can construct only one stationary Markov Gaussian process. This, of course, is well-known (see discussion in Adler and Epstein (1987) §5).

As another corollary, we obtain the following:

Theorem 1. *Let $X_u, u \geq 0$ be a randomized symmetric random walk with exponential lifetime:*
for $u < \zeta$, when $\zeta \sim \exp(\theta), \theta > 0$,

$$X_u = \sum_{i=1}^{N_u} \xi_i, \;\; N_u \sim \text{Poisson}(\mu u), \;\; \mu > 0, \tag{9}$$

where ξ_i are i.i.d. random variables and

$$P(\xi_i = \pm l) = \frac{p_l}{2}, \;\; l = 1, \ldots, k, \;\; p_1 + \ldots + p_k = 1.$$

Then the Gaussian process $\Phi_t, t = 0, \pm 1, \pm 2, \ldots$ associated with X, is a causal $AR(k)$ process with spectral density

$$f(\lambda) = (2\pi)^{-1}\left(\theta + \mu - \mu \sum_{l=1}^{k} p_l \cos \lambda l\right)^{-1} \tag{10}$$

Proof. For given X, $\beta = 0, \sigma^2 = 0$ and the measure ν is concentrated on $\pm 1, \ldots, \pm k$:

$$\nu(dx) = \mu \sum_{i=1}^{k} \frac{p_l}{2}(\delta_{-l}(x) + \delta_l(x))\, dx. \tag{11}$$

Note that $\int x/(1+x^2)\nu(dx) = 0$. Then

$$\begin{aligned} f(\lambda) &= (2\pi)^{-1}\left(\theta + \mu \sum_{l=1}^{k} \frac{p_l}{2}(1 - e^{-i\lambda l} + 1 - e^{i\lambda l})\right)^{-1} \\ &= (2\pi)^{-1}\left(\theta + \mu - \mu \sum_{l=1}^{k} p_l \cos \lambda l\right)^{-1}. \end{aligned}$$

Since $f(\lambda)$ is a positive symmetric integrable bounded function, it is the spectral density of a real-valued stationary process. It can be written in the form

$$f(\lambda) = (2\pi)^{-1}\left(\theta + \mu - \mu \sum_{l=-k,\ldots,k,\, l\neq 0} \frac{p_l}{2} e^{i\lambda l}\right)^{-1}.$$

i.e. $f(\lambda)$ is a rational function of $e^{i\lambda}$. According to Yaglom (1962), p. 121, it can be then represented in the form

$$f(\lambda) = \frac{|B(e^{i\lambda})|^2}{|A(e^{i\lambda})|^2} = \frac{|B_0 e^{iM\lambda} + B_1 e^{i(M-1)\lambda} + \ldots + B_M|^2}{|A_0 e^{iN\lambda} + A_1 e^{i(N-1)\lambda} + \ldots + A_N|^2},$$

where A_j and B_j are real, A_0, B_0, A_N, and B_M are all non-zero, all zeros of the polynomial $A(z)$, $z \in C$, lie inside of the unit circle and the zeros of the polynomial $B(z)$ have absolute values which do not exceed unity.

In our case the complex-variable function $f^*(z) := f(e^{i\lambda})$ can be written as

$$f^*(z) = (2\pi)^{-1}(-2/\mu p_k) z^k [P_{2k}(z)]^{-1}, \tag{12}$$

where

$$P_{2k}(z) = z^{2k} + \frac{p_{k-1}}{p_k} z^{2k-1} + \ldots + \frac{p_1}{p_k} z^{k+1} - \frac{2(\theta + \mu)}{\mu p_k} z^k + \frac{p_1}{p_k} z^{k-1} + \ldots + 1. \tag{13}$$

The right-hand side of (12) has no zeros for $z \neq 0, |z| \leq 1$ and $f^*(z) = |B(z)|^2/|A(z)|^2$ has M zeros inside the closure of the unit circle, all of which are non-zero. This is possible only if $M = 0$, $N = k$ and

$$f(\lambda) = \left(\frac{|A_0|^2}{|B_0|^2}\left|1 + \frac{A_1}{A_0}e^{-i\lambda} + \ldots + \frac{A_k}{A_0}e^{-ik\lambda}\right|^2\right)^{-1}. \tag{14}$$

From Brockwell and Davis(1987), §4, we learn that this is the spectral density of a causal AR(k) process

$$\Phi_t + \frac{A_1}{A_0}\Phi_{t-1} + \ldots + \frac{A_k}{A_0}\Phi_{t-k} = Z_t,\ Z_t \sim \text{White Noise}(0, 2\pi\frac{|B_0|^2}{|A_0|^2}). \tag{15}$$

We are done.

Corollary. *The AR(k) process Φ associated with random walk (9) has representation*

$$\Phi_t = A_1\Phi_{t-1} + \ldots + A_k\Phi_{t-k} + Z_t,\ Z_t \sim \text{White Noise}(0, 2\pi|B_0|^2)$$

where the constants $B_0, A_1, \ldots, A_k$ can be found from the system

$$\begin{aligned} \alpha_1 + \alpha_2 + \ldots + \alpha_k &= A_1 \\ \alpha_1\alpha_2 + \alpha_1\alpha_3 + \ldots + \alpha_{k-1}\alpha_k &= -A_2 \\ \ldots \qquad \ldots &\quad \ldots \\ \alpha_1\alpha_2\ldots\alpha_k &= (-1)^{k+1}A_k \\ \frac{1}{2\pi}\frac{2}{\mu p_k} &= \frac{|B_0|^2}{A_k} \end{aligned} \tag{16}$$

As before $\alpha_1, \ldots, \alpha_k$ are k roots of the polynomial P_{2k}, defined in (13), which lie inside of the unit circle.

This follows from Vieta's theorem and the fact that $\alpha_1, \ldots, \alpha_k$ are k roots of the polynomial $A(z)$ ($A(z)$ was defined while proving Theorem 1). Since only A_j/A_0 matter, we took $A_0 = -1$. Of course, if all roots $\alpha_1, \ldots, \alpha_k$ are distinct from each other, the coefficients $A_1, \ldots, A_k$ can be found from a linear system:

$$\begin{aligned} A_1\alpha_1^{k-1} + A_2\alpha_1^{k-2} + \ldots + A_k &= \alpha_1^k \\ A_1\alpha_2^{k-1} + A_2\alpha_2^{k-2} + \ldots + A_k &= \alpha_2^k \\ \ldots \qquad \ldots &\quad \ldots \\ A_1\alpha_k^{k-1} + A_2\alpha_k^{k-2} + \ldots + A_k &= \alpha_k^k \end{aligned}$$

Example 1. Let ξ_i be Bernoulli random variables, $P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$. Then the Gaussian process Φ associated with the randomized random walk X has spectral density

$$f(\lambda) = \frac{1}{2\pi}\left(\theta + \mu - \mu\cos\lambda\right)^{-1}.$$

This is the spectral density of an AR(1) process Φ: $\Phi_t = A_1\Phi_{t-1} + Z_t$ with

$$A_1 = (1+\theta/\mu) - \sqrt{(1+\theta/\mu)^2 - 1};\ \ Z_t \sim \text{White Noise}(0, 2A_1/\mu).$$

Example 2. Let k=2, $P(\xi_i = \pm 1) = p_1/2$ and $P(\xi_i = \pm 2) = p_2/2$; $p_1 + p_2 = 1$. Then the spectral density of the Gaussian process Φ is

$$f(\lambda) = \frac{1}{2\pi}\left(\theta + \mu - \mu p_1\cos\lambda - \mu p_2\cos 2\lambda\right)^{-1}.$$

This is the spectral density of AR(2) process. To find the coefficients A_1 and A_2 we write:

$$P_4(z) = z^4 + \frac{p_1}{p_2}z^3 - \frac{2(\theta+\mu)}{\mu p_2}z^2 + \frac{p_1}{p_2}z + 1 = 0.$$

Let $z + \frac{1}{z} = y$; then $y^2 = z^2 + \frac{1}{z^2} + 2$ and we have:

$$y^2 + \frac{p_1}{p_2}y - \left(2 + \frac{2}{p_2}\left(1 + \frac{\theta}{\mu}\right)\right) = 0;$$

$$y_1 = -\frac{p_1}{2p_2} + \sqrt{\frac{p_1^2}{4p_2^2} + 2 + \frac{2}{p_2}\left(1+\frac{\theta}{\mu}\right)}$$

$$y_2 = -\frac{p_1}{2p_2} - \sqrt{\frac{p_1^2}{4p_2^2} + 2 + \frac{2}{p_2}\left(1+\frac{\theta}{\mu}\right)}.$$

The roots α_1 and α_2 of $P_4(z)$, which are inside the unit circle, can be obtained from

$$\alpha_1 = \frac{y_1}{2} - \sqrt{\frac{y_1^2}{4} - 1},\ \ \alpha_2 = \frac{y_2}{2} + \sqrt{\frac{y_2^2}{4} - 1}, \tag{17}$$

Then

$$\begin{aligned} A_1 &= \alpha_1 + \alpha_2 \\ A_2 &= -\alpha_1\alpha_2 \\ Z_t &\sim \text{White Noise}(0, (2/\mu p_2)A_2) \end{aligned} \tag{18}$$

For example, if $\theta = 1/16, \mu = 5/4, p_1 = 3/5$, and $p_2 = 2/5$ then $y_{1,2} = (-3 \pm 5\sqrt{5})/4$; $\alpha_1 = (1+\sqrt{5})/4 \approx 0.809, \alpha_2 = (1-\sqrt{5})/4 \approx -0.309, A_1 = 0.5, A_2 = 0.25, |B_0|^2 = 1/2\pi$.
We get $\Phi : \Phi_t = (1/2)\Phi_{t-1} + (1/4)\Phi_{t-2} + Z_t,\ Z_t \sim \text{White Noise}(0,1)$.

4 First-Hit Probabilities of Randomized Random Walks.

The main result of section 3 is that Gaussian processes associated with randomized random walks are causal stationary AR(k) processes with spectral density given by (10). Applying the theory of extrapolation of stationary sequences with rational spectral density in $e^{i\lambda}$ for density (10) we obtain the following prediction procedure:

1. Let $\alpha_1, \ldots, \alpha_k$ be roots of the polynomial

$$P_{2k}(z) = z^{2k} + \frac{p_{k-1}}{p_k} z^{2k-1} + \ldots + \frac{p_1}{p_k} z^{k+1} + \left(-\frac{2(\theta+\mu)}{\mu p_k} \right) z^k + \frac{p_1}{p_k} z^{k-1} + \ldots + 1 = 0 \quad (19)$$

such that $|\alpha_j| < 1$, $j = 1, \ldots, k$. Assume that the roots α_j are all distinct.

2. For $m \geq 0$ find the coefficients $\gamma_0^{(m)}, \ldots, \gamma_{k-1}^{(m)}$ of the linear system

$$\begin{aligned} \gamma_{k-1}^{(m)} \alpha_1^{k-1} + \gamma_1^{(m)} \alpha_1 + \gamma_0^{(m)} \cdot 1 &= \alpha_1^{m+k} \\ \ldots \qquad\qquad \ldots \quad &\qquad \ldots \\ \gamma_{k-1}^{(m)} \alpha_k^{k-1} + \gamma_1^{(m)} \alpha_k + \gamma_0^{(m)} \cdot 1 &= \alpha_k^{m+k} \end{aligned} \quad (20)$$

3. Then find the $(m+1)$-step predictor from:

$$\hat{\Phi}(t+m) := E(\Phi_{t+m} | \Phi_s,\ s \leq t-1) = \gamma_{k-1}^{(m)} \Phi(t-1) + \ldots + \gamma_0^{(m)} \Phi(t-k).$$

Note that for $m = 0$, $\gamma_{k-j}^0 = A_j$.

In the general case, if $\alpha_i = \alpha_j$ for some $i \neq j$, we can use the recursive methods of prediction (Brockwell and Davis (1987)) or methods of Yaglom (1962), §24, relying on the knowledge of roots of the polynomial $P_{2k}(z)$. Combining the prediction formula for Φ with the theorem of Dynkin from section 2 we have the following results:

Result 1. *Let $X_u, u \geq 0$ be a randomized symmetric random walk with jumps of size ± 1 and exponential lifetime:*
for $u < \zeta$, when $\zeta \sim \exp(\theta), \theta > 0$,

$$X_u = \sum_{i=1}^{N_u} \xi_i, \quad N_u \sim \text{Poisson}(\mu u), \quad \mu > 0, \quad (21)$$

where ξ_i are i.i.d. and $P(\xi_i = \pm 1) = 1/2$.
Let $l \geq 1$. Then

$$P(X_u \geq l \text{ for some } u < \zeta) = A^l; \quad A = \left(1 + \frac{\theta}{\mu}\right) - \sqrt{\left(1 + \frac{\theta}{\mu}\right)^2 - 1}. \quad (22)$$

Proof: The Gaussian Process Φ associated with X is an AR(1) process:

$$\Phi_t = A_1 \Phi_{t-1} + Z_t$$

$$A_1 = \left(1 + \frac{\theta}{\mu}\right) - \sqrt{\left(1 + \frac{\theta}{\mu}\right)^2 - 1}; \; Z_t \sim \text{White Noise}(0, \frac{2}{\mu} A_1)$$

(see example 1 of the previous section). The prediction formula for Φ is:

$$\hat{\Phi}(t+m) = A^{m+1} \Phi_{t-1}, \;\; m \geq 0,$$

i.e.

$$E(\Phi_{t+m} \mid \Phi_s, \, s \leq t-1) = \sum_{s \leq t-1} p_{t+m}^{\{s:s \leq t-1\}}(s) \Phi_s = A^{m+1} \Phi_{t-1},$$

which means that

$$p_{t+m}^{\{s:s \leq t-1\}}(t-1) \equiv A^{m+1},$$

while $p_{t+m}^{\{s:s \leq t-1\}}(s) = 0$ for $s < t-1$. But, using the symmetry of X, we have:

$$\begin{aligned} p_{t+m}^{\{s:s \leq t-1\}}(t-1) &= P(X, \text{ starting at } t+m, \text{ enters } \{s : s \leq t-1\} \text{ at point } t-1) \\ &= P(X, \text{ starting at } 0, \text{ enters } \{s : s \leq -m-1\} \text{ at point } -m-1) \\ &= P(X, \text{ starting at } 0, \text{ enters } \{s : s \geq m+1\} \text{ at point } m+1) \\ &= P(\max(X_u; u < \zeta) \geq m+1). \end{aligned}$$

Replace $m+1$ by l and we are done.

Remark. For comparison, we compute (22) using combinatorial methods of Feller(1968). Consider only $l = 2$. First, let us find the probability π_N that the path of length N (i.e. the number of steps is fixed to be N) starting from the origin has maximum less than 2. This is the same as having minimum greater than -2. Let us fix an end point $(N, 2k)$, where $k \geq 0$, $N = 2n$, and $n \geq 1$. The number of paths from the origin to the point $(2n, 2k)$ is

$$N_{2n,k} = \binom{2n}{n+k};$$

each has probability 2^{-2n} (see (III.2.2) of Feller(1968)). The number of paths from the origin to $(2n, 2k)$ which visit point -2 is the same as the number of paths from $(0, -4)$ to $(2n, 2k)$ (by the reflection principle) and it is equal to the number of paths from $(0,0)$ to $(2n, 2k+4)$, i.e. $N_{2n,2k+4}$. Thus, the probability that the path from $(0,0)$ to $(2n, 2k)$ will not visit -2 is $2^{-2n}(N_{2n,2k} - N_{2n,2k+4})$, and the probability that the path with $N = 2n$ steps, starting from the origin, will be above level -2 (or below level 2) is

$$\pi_N = \sum_{k=0}^{\infty} 2^{-2n}(N_{2n,2k} - N_{2n,2k+4}) = 2^{-2n}(N_{2n,0} + N_{2n,2})$$

$$= 2^{-2n}\left(\binom{2n}{n}+\binom{2n}{n+1}\right) = 2^{-2n}\binom{2n+1}{n+1} = 2(-1)^{n+1}\binom{-1/2}{n+1} \qquad (23)$$

(we used formulas (II.8.6) and (II.12.5) of Feller (1968)). This quantity also represents the probability of interest for a path with $N = 2n+1$ steps, since a path which is below level 2 for $2n$ steps remains there for at least one additional step. Thus,

$$\begin{aligned}
& P(\max(X_u; u < \zeta) \geq 2) \\
= & \int_0^\infty \theta e^{-\theta z}\, dz \sum_{N=2}^\infty \frac{e^{-\mu z}}{N!}(\mu z)^N (1-\pi_N) \\
= & \sum_{N=2}^\infty \frac{\theta}{\theta+\mu}\left(\frac{\mu}{\theta+\mu}\right)^N (1-\pi_N) \\
= & \left(\frac{\mu}{\theta+\mu}\right)^2 - \sum_{n=1}^\infty \frac{\theta}{\theta+\mu}\left(\frac{\mu}{\theta+\mu}\right)^{2n}\left(1+\frac{\mu}{\theta+\mu}\right)2(-1)^{n+1}\binom{-1/2}{n+1} \\
= & \left(\frac{\mu}{\theta+\mu}\right)^2 - 2\frac{\theta(\theta+2\mu)}{(\theta+\mu)^2}\sum_{n=1}^\infty \left(\frac{\mu}{\theta+\mu}\right)^{2n}(-1)^{n+1}\binom{-1/2}{n+1} \\
= & \left(\frac{\mu}{\theta+\mu}\right)^2 - 2\frac{\theta(\theta+2\mu)}{(\theta+\mu)^2}\left(\frac{\theta+\mu}{\mu}\right)^2\sum_{n=2}^\infty \left(\frac{\mu}{\theta+\mu}\right)^{2n}(-1)^{n}\binom{-1/2}{n} \\
= & \left(\frac{\mu}{\theta+\mu}\right)^2 - 2\frac{\theta(\theta+2\mu)}{\mu^2}\left[\left(1-\left(\frac{\mu}{\theta+\mu}\right)^2\right)^{-\frac{1}{2}} - 1 + \left(\frac{\mu}{\theta+\mu}\right)^2\binom{-1/2}{1}\right] \\
= & \left(\frac{\mu}{\theta+\mu}\right)^2 - 2\frac{\theta(\theta+2\mu)}{\mu^2}\frac{\theta+\mu}{\sqrt{\theta(\theta+2\mu)}} + 2\frac{\theta(\theta+2\mu)}{\mu^2} + \frac{\theta(\theta+2\mu)}{\mu^2}\left(\frac{\mu}{\theta+\mu}\right)^2 \\
= & \left(\frac{\mu}{\theta+\mu}\right)^2 - 2\frac{\sqrt{\theta(\theta+2\mu)}(\theta+\mu)}{\mu^2} + 2\frac{\theta(\theta+2\mu)}{\mu^2} + \frac{\theta(\theta+2\mu)}{(\theta+\mu)^2} \\
= & 1 - 2\left(1+\frac{\theta}{\mu}\right)\sqrt{\frac{\theta}{\mu}\left(2+\frac{\theta}{\mu}\right)} + 2\frac{\theta}{\mu}\left(2+\frac{\theta}{\mu}\right) \\
= & 1 + 2\frac{\theta}{\mu} + \left(\frac{\theta}{\mu}\right)^2 - 2\left(1+\frac{\theta}{\mu}\right)\sqrt{\frac{\theta}{\mu}\left(2+\frac{\theta}{\mu}\right)} + \frac{\theta}{\mu}\left(2+\frac{\theta}{\mu}\right) \\
= & \left(\left(1+\frac{\theta}{\mu}\right) - \sqrt{\left(1+\frac{\theta}{\mu}\right)^2 - 1}\right)^2 \\
= & A^2.
\end{aligned}$$

Case $l > 2$ involves even more complicated combinatorics and summations. In fact,

§7 of Chapter II and §6 of Chapter XIV of Feller (1971), which discuss randomized random walks, deal with jumps of size ± 1 only and use Bessel functions.

Result 2. *Let $X_u, u \geq 0$ be a randomized symmetric random walk (21) with jumps of size $\pm 1, \pm 2$, i.e. in (21)*

$$P(\xi_i = \pm 1) = p_1/2, \quad P(\xi_i = \pm 2) = p_2/2; \; p_1 + p_2 = 1.$$

Let $\tau = \min\{u > 0 : X_u \in \{1, 2, \ldots\}\}$. Then

$$P(X_\tau = 1) = \frac{y_1 + y_2}{2} - \sqrt{\frac{y_1^2}{4} - 1} + \sqrt{\frac{y_2^2}{4} - 1}$$

$$P(X_\tau = 2) = \left(\frac{y_1}{2} - \sqrt{\frac{y_1^2}{4} - 1}\right)\left(\frac{y_2}{2} + \sqrt{\frac{y_2^2}{4} - 1}\right)$$

where

$$y_1 = -\frac{p_1}{2p_2} + \sqrt{\frac{p_1^2}{4p_2^2} + 2 + \frac{2}{p_2}\left(1 + \frac{\theta}{\mu}\right)};$$

$$y_2 = -\frac{p_1}{2p_2} - \sqrt{\frac{p_1^2}{4p_2^2} + 2 + \frac{2}{p_2}\left(1 + \frac{\theta}{\mu}\right)}.$$

Proof. The Gaussian process Φ associated with X is an AR(2) process $\Phi_t = A_1\Phi_{t-1} + A_2\Phi_{t-2} + Z_t$ with A_1, A_2, Z_t defined in (17) - (18). The one-step predictor for Φ is $\hat{\Phi}(t) = A_1\Phi_{t-1} + A_2\Phi_{t-2}$. The coefficients A_1, A_2 of the prediction formula are expressed through the first-hit probabilities of the random walk:

$$A_1 = p_t^{\{s:s\leq t-1\}}(t-1), \; A_2 = p_t^{\{s:s\leq t-1\}}(t-2).$$

The symmetry of X gives us the desired formulas.

Using symmetry and reflection properties of random walks, we obtain the following result:

Result 3. *Let $X_u, u \geq 0$ be as in Result 2. Let*

$$\tau_n = \min\{u > 0 : X_u \in \{n, n+1, \ldots\}\}$$

Then

$$\begin{aligned} P(X_{\tau_n} = n+1) &= \alpha_1\alpha_2\frac{\alpha_2^n - \alpha_1^n}{\alpha_2 - \alpha_1} \\ &= (\alpha_1\alpha_2)\left(\alpha_1^{n-1} + \alpha_1^{n-2}\alpha_2 + \ldots + \alpha_2^{n-1}\right) \\ P(X_{\tau_n} = n) &= \frac{\alpha_2^{n+1} - \alpha_1^{n+1}}{\alpha_2 - \alpha_1} \\ &= \alpha_1^n + \alpha_1^{n-1}\alpha_2 + \ldots + \alpha_2^n \end{aligned}$$

where α_1 and α_2 are defined in (17).
We get this by direct solution of system (20) for the case $k = 2$, and the symmetry of the random walk.

For example, if the parameters in Result 2 are $p_1 = 3/5$, $p_2 = 2/5$, $\theta = 1/16$, and $\mu = 5/4$, then from the calculations at the end of example 2 of the previous section we find that $P(X_\tau = 1) = 1/2$ and $P(X_\tau = 2) = 1/4$.

Result 4. *Let X_u, $u \geq 0$ be a randomized symmetric random walk (21) with jumps of size $\pm 1, \pm 2, \ldots, \pm k$, i.e. in (21).*

$$P(\xi_i = \pm l) = p_{l/2};\ l = 1, \ldots, k;\ p_1 + \ldots + p_k = 1.$$

Assume that all roots $\alpha_1, \ldots, \alpha_k$, $|\alpha_j| < 1$, of the polynomial P_{2k} from (19) are distinct. Let $\tau_n = \min\{u > 0 : X_u \in \{n, n+1, \ldots\}\}$. Then $P(X_{\tau_n} = n + l) = \gamma_{k-l-1}^{(n-1)}$, $l = 0, 1, \ldots, k-1$, where $\gamma_{k-1}^{(n-1)}, \ldots, \gamma_0^{(n-1)}$ are found from system (20).

Result 5. *Let X be as in Result 4, except without conditions on the roots of P_{2k}. Then $P(X_{\tau_1} = l + 1) = \gamma_{k-l-1}^0 = A_{l+1}$, $l = 0, 1, \ldots, k-1$ where $A_1, \ldots, A_k$ are found from (16).*

In general, $P(X_{\tau_n} = n + l), l = 0, \ldots, k-1$, can be found as coefficients of Φ_{t-l-1} in the prediction formula, estimating Φ_{t+n-1} from observed values $\Phi_{t-1}, \Phi_{t-2}, \ldots$

Until now we were interested in first-hit probabilities for sets of the form $\{n, n+1, ...\}$ (the first-hit probabilites for sets $\{\ldots, -(n+1), -n\}$ are the same because of the symmetry of this random walk). The following result involves first-hit probabilities for sets of the form $\{\ldots, q, q+m+1, \ldots\}$ when the process starts from points $q + l$ of the set $\{q+1, \ldots, q+m\}$. A similar situation is discussed in Feller (1968), ch. XIV, §8, where the problem is related to sequential analysis. As before, we consider a randomized symmetric random walk, for simplicity allowing only jumps of ± 1 at one step.

Result 6. *Let $X_u, u \geq 0$ be a symmetric randomized random walk (21) with ξ_i taking values ± 1 only, each with probability $1/2$, except that let now $X_0 = q + l$. Let $\tau = \min\{u > 0 : X_u \in \{\ldots, q, q+m+1, \ldots\}\}$, for some $m \geq 1$ and $1 \leq l \leq m$. Then*

$$\begin{aligned} P(X_\tau = q) &= A^l - A^{2m+2-l} \frac{1 - A^{2l}}{1 - A^{2m+2}} \\ P(X_\tau = q + m + 1) &= A^{m+1-l} \frac{1 - A^{2l}}{1 - A^{2m+2}}, \end{aligned}$$

where

$$A = (1 + \frac{\theta}{\mu}) - \sqrt{\left(1 + \frac{\theta}{\mu}\right)^2 - 1}$$

Proof. To prove the result we have to show that for an AR(1) process Φ_t satisfying $\Phi_t = A\Phi_{t-1} + Z_t$, $Z_t \sim$ White Noise$(0, 2A/\mu)$, the following prediction formula holds:

$$\begin{aligned} \Phi_{q+l|q,m} &:= E(\Phi_{q+l}|\Phi_1, \ldots, \Phi_{q-1}, \Phi_q, \Phi_{q+m+1}, \Phi_{q+m+2}, \ldots, \Phi_n) \\ &= \left[A^l - A^{2m+2-l}\frac{1-A^{2l}}{1-A^{2m+2}}\right]\Phi_q + \left[A^{m+1-l}\frac{1-A^{2l}}{1-A^{2m+2}}\right]\Phi_{q+m+1} \end{aligned}$$

where $n \geq q+m+1$ and $q \geq 1$. The latter follows easily from application of the Kalman Fixed-Point Smoother to estimation of the missing value Φ_{q+l} of our AR(1) process (see Brockwell and Davis (1987), §12.7).

The formulas of result 6 can be easily checked for the simplest case, $m = 1$:

$$P(X_\tau = q) = P(X_\tau = q+2) = 1/2P(\text{path length is at least } 1)$$

$$= 1/2\int_0^\infty \theta e^{-\theta z} dz(1 - e^{-z\mu}) = 1/2\left(1 - \frac{\theta}{\theta+\mu}\right) = \frac{\mu}{2(\theta+\mu)} = \frac{A}{1+A^2},$$

where the last equality follows from the identity $A^2 - 2(\theta+\mu)/\mu A + 1 = 0$. The case of $m > 1$, however, involves calculations similar to ones following result 1. The advantage of the method appears when we try to calculate first-hit probabilities for random walks with jumps $\pm 1, \ldots, \pm k$, $k > 1$, $m \geq 1$, through the theory of estimation of missing values of an AR(k) process. Programs for performing such calculations can be found in Wampler (1986).

References

Adler, R. J. and Epstein, R. (1987), Some central limit theorems for Markov paths and some properties of Gaussian random fields, *Stoch. Proc. Appls.*, **24**, 157–202.

Blumenthal, R. M. and Getoor R. K (1968), *Markov Processes and Potential Theory*, Academic Press, New York.

Brockwell, P. J. and Davis, R. A. (1987), *Time Series: Theory and Methods*, Springer-Verlag, New York.

Dynkin, E. B. (1980), Markov processes and random fields, *Bull. Amer. Math. Soc.*, **3**, 975–999.

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Vol. I, Wiley, New York.

Feller, W. (1971), *An Introduction to Probability Theory and Its Applications*, Vol. II, Wiley, New York.

Wampler, S. (1968), Missing values in time series analysis, Statistics Dept. of Colorado State University.

Yaglom, A. M. (1962), *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall, Englewood Cliffs.

Ylvisaker, D. (1987), Prediction and design, *Ann. Statist.*, **15**, 1-19.

TECHNICAL REPORTS

Statistics Department

University of California, Berkeley

1. BREIMAN, L. and FREEDMAN, D. (Nov. 1981, revised Feb. 1982). How many variables should be entered in a regression equation? Jour. Amer. Statist. Assoc., March 1983, 78, No. 381, 131-136.

2. BRILLINGER, D. R. (Jan. 1982). Some contrasting examples of the time and frequency domain approaches to time series analysis. Time Series Methods in Hydrosciences, (A. H. El-Shaarawi and S. R. Esterby, eds.) Elsevier Scientific Publishing Co., Amsterdam, 1982, pp. 1-15.

3. DOKSUM, K. A. (Jan. 1982). On the performance of estimates in proportional hazard and log-linear models. Survival Analysis, (John Crowley and Richard A. Johnson, eds.) IMS Lecture Notes - Monograph Series, (Shanti S. Gupta, series ed.) 1982, 74-84.

4. BICKEL, P. J. and BREIMAN, L. (Feb. 1982). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. Ann. Prob., Feb. 1982, 11. No. 1, 185-214.

5. BRILLINGER, D. R. and TUKEY, J. W. (March 1982). Spectrum estimation and system identification relying on a Fourier transform. The Collected Works of J. W. Tukey, vol. 2, Wadsworth, 1985, 1001-1141.

6. BERAN, R. (May 1982). Jackknife approximation to bootstrap estimates. Ann. Statist., March 1984, 12 No. 1, 101-118.

7. BICKEL, P. J. and FREEDMAN, D. A. (June 1982). Bootstrapping regression models with many parameters. Lehmann Festschrift, (P. J. Bickel, K. Doksum and J. L. Hodges, Jr., eds.) Wadsworth Press, Belmont, 1983, 28-48.

8. BICKEL, P. J. and COLLINS, J. (March 1982). Minimizing Fisher information over mixtures of distributions. Sankhyā, 1983, 45, Series A, Pt. 1, 1-19.

9. BREIMAN, L. and FRIEDMAN, J. (July 1982). Estimating optimal transformations for multiple regression and correlation.

10. FREEDMAN, D. A. and PETERS, S. (July 1982, revised Aug. 1983). Bootstrapping a regression equation: some empirical results. JASA, 1984, 79, 97-106.

11. EATON, M. L. and FREEDMAN, D. A. (Sept. 1982). A remark on adjusting for covariates in multiple regression.

12. BICKEL, P. J. (April 1982). Minimax estimation of the mean of a mean of a normal distribution subject to doing well at a point. Recent Advances in Statistics, Academic Press, 1983.

14. FREEDMAN, D. A., ROTHENBERG, T. and SUTCH, R. (Oct. 1982). A review of a residential energy end use model.

15. BRILLINGER, D. and PREISLER, H. (Nov. 1982). Maximum likelihood estimation in a latent variable problem. Studies in Econometrics, Time Series, and Multivariate Statistics, (eds. S. Karlin, T. Amemiya, L. A. Goodman). Academic Press, New York, 1983, pp. 31-65.

16. BICKEL, P. J. (Nov. 1982). Robust regression based on infinitesimal neighborhoods. Ann. Statist., Dec. 1984, 12, 1349-1368.

17. DRAPER, D. C. (Feb. 1983). Rank-based robust analysis of linear models. I. Exposition and review. Statistical Science, 1988, Vol.3 No. 2 239-271.

18. DRAPER, D. C. (Feb 1983). Rank-based robust inference in regression models with several observations per cell.

19. FREEDMAN, D. A. and FIENBERG, S. (Feb. 1983, revised April 1983). Statistics and the scientific method, Comments on and reactions to Freedman, A rejoinder to Fienberg's comments. Springer New York 1985 Cohort Analysis in Social Research, (W. M. Mason and S. E. Fienberg, eds.).

20. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Jan. 1984). Using the bootstrap to evaluate forecasting equations. J. of Forecasting. 1985, Vol. 4, 251-262.

21. FREEDMAN, D. A. and PETERS, S. C. (March 1983, revised Aug. 1983). Bootstrapping an econometric model: some empirical results. JBES, 1985, 2, 150-158.

22. FREEDMAN, D. A. (March 1983). Structural-equation models: a case study.

23. DAGGETT, R. S. and FREEDMAN, D. (April 1983, revised Sept. 1983). Econometrics and the law: a case study in the proof of antitrust damages. Proc. of the Berkeley Conference, in honor of Jerzy Neyman and Jack Kiefer. Vol I pp. 123-172. (L. Le Cam, R. Olshen eds.) Wadsworth, 1985.

24. DOKSUM, K. and YANDELL, B. (April 1983). Tests for exponentiality. Handbook of Statistics, (P. R. Krishnaiah and P. K. Sen, eds.) 4, 1984, 579-611.

25. FREEDMAN, D. A. (May 1983). Comments on a paper by Markus.

26. FREEDMAN, D. (Oct. 1983, revised March 1984). On bootstrapping two-stage least-squares estimates in stationary linear models. Ann. Statist., 1984, 12, 827-842.

27. DOKSUM, K. A. (Dec. 1983). An extension of partial likelihood methods for proportional hazard models to general transformation models. Ann. Statist., 1987, 15, 325-345.

28. BICKEL, P. J., GOETZE, F. and VAN ZWET, W. R. (Jan. 1984). A simple analysis of third order efficiency of estimate Proc. of the Neyman-Kiefer Conference, (L. Le Cam, ed.) Wadsworth, 1985.

29. BICKEL, P. J. and FREEDMAN, D. A. Asymptotic normality and the bootstrap in stratified sampling. Ann. Statist. 12 470-482.

30. FREEDMAN, D. A. (Jan. 1984). The mean vs. the median: a case study in 4-R Act litigation. JBES. 1985 Vol 3 pp. 1-13.

31. STONE, C. J. (Feb. 1984). An asymptotically optimal window selection rule for kernel density estimates. Ann. Statist., Dec. 1984, 12, 1285-1297.

32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.

33. STONE, C. J. (Oct. 1984). Additive regression and other nonparametric models. Ann. Statist., 1985, 13, 689-705.

34. STONE, C. J. (June 1984). An asymptotically optimal histogram selection rule. Proc. of the Berkeley Conf. in Honor of Jerzy Neyman and Jack Kiefer (L. Le Cam and R. A. Olshen, eds.), II, 513-520.

35. FREEDMAN, D. A. and NAVIDI, W. C. (Sept. 1984, revised Jan. 1985). Regression models for adjusting the 1980 Census. Statistical Science. Feb 1986, Vol. 1, No. 1, 3-39.

36. FREEDMAN, D. A. (Sept. 1984, revised Nov. 1984). De Finetti's theorem in continuous time.

37. DIACONIS, P. and FREEDMAN, D. (Oct. 1984). An elementary proof of Stirling's formula. Amer. Math Monthly. Feb 1986, Vol. 93, No. 2, 123-125.

38. LE CAM, L. (Nov. 1984). Sur l'approximation de familles de mesures par des familles Gaussiennes. Ann. Inst. Henri Poincaré, 1985, 21, 225-287.

39. DIACONIS, P. and FREEDMAN, D. A. (Nov. 1984). A note on weak star uniformities.

40. BREIMAN, L. and IHAKA, R. (Dec. 1984). Nonlinear discriminant analysis via SCALING and ACE.

41. STONE, C. J. (Jan. 1985). The dimensionality reduction principle for generalized additive models.

42. LE CAM, L. (Jan. 1985). On the normal approximation for sums of independent variables.

43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?

44. BRILLINGER, D. R. (1985). The natural variability of vital rates and associated statistics. Biometrics, to appear.

45. BRILLINGER, D. R. (1985). Fourier inference: some methods for the analysis of array and nonGaussian series data. Water Resources Bulletin, 1985, 21, 743-756.

46. BREIMAN, L. and STONE, C. J. (1985). Broad spectrum estimates and confidence intervals for tail quantiles.

47. DABROWSKA, D. M. and DOKSUM, K. A. (1985, revised March 1987). Partial likelihood in transformation models with censored data. Scandinavian J. Statist., 1988, 15, 1-23.

48. HAYCOCK, K. A. and BRILLINGER, D. R. (November 1985). LIBDRB: A subroutine library for elementary time series analysis.

49. BRILLINGER, D. R. (October 1985). Fitting cosines: some procedures and some physical examples. Joshi Festschrift, 1986. D. Reidel.

50. BRILLINGER, D. R. (November 1985). What do seismology and neurophysiology have in common? - Statistics! Comptes Rendus Math. Rep. Acad. Sci. Canada. January, 1986.

51. COX, D. D. and O'SULLIVAN, F. (October 1985). Analysis of penalized likelihood-type estimators with application to generalized smoothing in Sobolev Spaces.

52. O'SULLIVAN, F. (November 1985). A practical perspective on ill-posed inverse problems: A review with some new developments. To appear in Journal of Statistical Science.

53. LE CAM, L. and YANG, G. L. (November 1985, revised March 1987). On the preservation of local asymptotic normality under information loss.

54. BLACKWELL, D. (November 1985). Approximate normality of large products.

55. FREEDMAN, D. A. (June 1987). As others see us: A case study in path analysis. Journal of Educational Statistics. 12, 101-128.

56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.

57. LE CAM, L. (February 1986). On the Bernstein - von Mises theorem.

58. O'SULLIVAN, F. (January 1986). Estimation of Densities and Hazards by the Method of Penalized likelihood.

59. ALDOUS, D. and DIACONIS, P. (February 1986). Strong Uniform Times and Finite Random Walks.

60. ALDOUS, D. (March 1986). On the Markov Chain simulation Method for Uniform Combinatorial Distributions and Simulated Annealing.

61. CHENG, C-S. (April 1986). An Optimization Problem with Applications to Optimal Design Theory.

62. CHENG, C-S., MAJUMDAR, D., STUFKEN, J. & TURE, T. E. (May 1986, revised Jan 1987). Optimal step type design for comparing test treatments with a control.

63. CHENG, C-S. (May 1986, revised Jan. 1987). An Application of the Kiefer-Wolfowitz Equivalence Theorem.

64. O'SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.

65. ALDOUS, D. (JUNE 1986). Finite-Time Implications of Relaxation Times for Stochastically Monotone Processes.

66. PITMAN, J. (JULY 1986, revised November 1986). Stationary Excursions.

67. DABROWSKA, D. and DOKSUM, K. (July 1986, revised November 1986). Estimates and confidence intervals for median and mean life in the proportional hazard model with censored data. Biometrika, 1987, 74, 799-808.

68. LE CAM, L. and YANG, G.L. (July 1986). Distinguished Statistics, Loss of information and a theorem of Robert B. Davies (Fourth edition).

69. STONE, C.J. (July 1986). Asymptotic properties of logspline density estimation.

71. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.

72. LEHMANN, E.L. (July 1986). Statistics - an overview.

73. STONE, C.J. (August 1986). A nonparametric framework for statistical modelling.

74. BIANE, PH. and YOR, M. (August 1986). A relation between Lévy's stochastic area formula, Legendre polynomial, and some continued fractions of Gauss.

75. LEHMANN, E.L. (August 1986, revised July 1987). Comparing Location Experiments.

76. O'SULLIVAN, F. (September 1986). Relative risk estimation.

77. O'SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.

78. PITMAN, J. & YOR, M. (September 1987). Further asymptotic laws of planar Brownian motion.

79. FREEDMAN, D.A. & ZEISEL, H. (November 1986). From mouse to man: The quantitative assessment of cancer risks. To appear in Statistical Science.

80. BRILLINGER, D.R. (October 1986). Maximum likelihood analysis of spike trains of interacting nerve cells.

81. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.

82. DOKSUM, K.J. and LO, A.Y. (Nov 1986, revised Aug 1988). Consistent and robust Bayes Procedures for Location based on Partial Information.

83. DABROWSKA, D.M., DOKSUM, K.A. and MIURA, R. (November 1986). Rank estimates in a class of semiparametric two-sample models.

84. BRILLINGER, D. (December 1986). Some statistical methods for random process data from seismology and neurophysiology.

85. DIACONIS, P. and FREEDMAN, D. (December 1986). A dozen de Finetti-style results in search of a theory. Ann. Inst. Henri Poincaré, 1987, 23, 397-423.

86. DABROWSKA, D.M. (January 1987). Uniform consistency of nearest neighbour and kernel conditional Kaplan - Meier estimates.

87. FREEDMAN, D.A., NAVIDI, W. and PETERS, S.C. (February 1987). On the impact of variable selection in fitting regression equations.

88. ALDOUS, D. (February 1987, revised April 1987). Hashing with linear probing, under non-uniform probabilities.

89. DABROWSKA, D.M. and DOKSUM, K.A. (March 1987, revised January 1988). Estimating and testing in a two sample generalized odds rate model. J. Amer. Statist. Assoc., 1988, 83, 744-749.

90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.

91. DIACONIS, P and FREEDMAN, D.A. (April 1988). Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. To appear in the Journal of Applied Probability.

92. DABROWSKA, D.M. (April 1987, revised September 1987). Kaplan-Meier estimate on the plane.

92a. ALDOUS, D. (April 1987). The Harmonic mean formula for probabilities of Unions: Applications to sparse random graphs.

93. DABROWSKA, D.M. (June 1987, revised Feb 1988). Nonparametric quantile regression with censored data.

94. DONOHO, D.L. & STARK, P.B. (June 1987). Uncertainty principles and signal recovery.

95. CANCELLED

96. BRILLINGER, D.R. (June 1987). Some examples of the statistical analysis of seismological data. To appear in Proceedings, Centennial Anniversary Symposium, Seismographic Stations, University of California, Berkeley.

97. FREEDMAN, D.A. and NAVIDI, W. (June 1987). On the multi-stage model for carcinogenesis. To appear in Environmental Health Perspectives.

98. O'SULLIVAN, F. and WONG, T. (June 1987). Determining a function diffusion coefficient in the heat equation.

99. O'SULLIVAN, F. (June 1987). Constrained non-linear regularization with application to some system identification problems.

100. LE CAM, L. (July 1987, revised Nov 1987). On the standard asymptotic confidence ellipsoids of Wald.

101. DONOHO, D.L. and LIU, R.C. (July 1987). Pathologies of some minimum distance estimators. Annals of Statistics, June, 1988.

102. BRILLINGER, D.R., DOWNING, K.H. and GLAESER, R.M. (July 1987). Some statistical aspects of low-dose electron imaging of crystals.

103. LE CAM, L. (August 1987). Harald Cramér and sums of independent random variables.

104. DONOHO, A.W., DONOHO, D.L. and GASKO, M. (August 1987). Macspin: Dynamic graphics on a desktop computer. IEEE Computer Graphics and applications, June, 1988.

105. DONOHO, D.L. and LIU, R.C. (August 1987). On minimax estimation of linear functionals.

106. DABROWSKA, D.M. (August 1987). Kaplan-Meier estimate on the plane: weak convergence, LIL and the bootstrap.

107. CHENG, C-S. (Aug 1987, revised Oct 1988). Some orthogonal main-effect plans for asymmetrical factorials.

108. CHENG, C-S. and JACROUX, M. (August 1987). On the construction of trend-free run orders of two-level factorial designs.

109. KLASS, M.J. (August 1987). Maximizing $E \max_{1 \le k \le n} S_k^+ / ES_n^+$: A prophet inequality for sums of I.I.D. mean zero variates.

110. DONOHO, D.L. and LIU, R.C. (August 1987). The "automatic" robustness of minimum distance functionals. Annals of Statistics, June, 1988.

111. BICKEL, P.J. and GHOSH, J.K. (August 1987, revised June 1988). A decomposition for the likelihood ratio statistic and the Bartlett correction — a Bayesian argument.

112. BURDZY, K., PITMAN, J.W. and YOR, M. (September 1987). Some asymptotic laws for crossings and excursions.

113. ADHIKARI, A. and PITMAN, J. (September 1987). The shortest planar arc of width 1.

114. RITOV, Y. (September 1987). Estimation in a linear regression model with censored data.

115. BICKEL, P.J. and RITOV, Y. (Sept. 1987, revised Aug 1988). Large sample theory of estimation in biased sampling regression models I.

116. RITOV, Y. and BICKEL, P.J. (Sept.1987, revised Aug. 1988). Achieving information bounds in non and semiparametric models.

117. RITOV, Y. (October 1987). On the convergence of a maximal correlation algorithm with alternating projections.

118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.

119. HESSE, C.H. (October 1987). An asymptotic expansion for the mean of the passage-time distribution of integrated Brownian Motion.

120. DONOHO, D. and LIU, R. (Oct. 1987, revised Mar. 1988, Oct. 1988). Geometrizing rates of convergence, II.

121. BRILLINGER, D.R. (October 1987). Estimating the chances of large earthquakes by radiocarbon dating and statistical modelling. Statistics a Guide to the Unknown, pp. 249-260 (Eds. J.M. Tanur et al.) Wadsworth, Pacific Grove.

122. ALDOUS, D., FLANNERY, B. and PALACIOS, J.L. (November 1987). Two applications of urn processes: The fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains.

123. DONOHO, D.L., MACGIBBON, B. and LIU, R.C. (Nov.1987, revised July 1988). Minimax risk for hyperrectangles.

124. ALDOUS, D. (November 1987). Stopping times and tightness II.

125. HESSE, C.H. (November 1987). The present state of a stochastic model for sedimentation.

126. DALANG, R.C. (December 1987, revised June 1988). Optimal stopping of two-parameter processes on nonstandard probability spaces.

127. Same as No. 133.

128. DONOHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.

129. SMITH, D.L. (December 1987). Exponential bounds in Vapnik-Červonenkis classes of index 1.

130. STONE, C.J. (Nov.1987, revised Sept. 1988). Uniform error bounds involving logspline models.

131. Same as No. 140

132. HESSE, C.H. (Dec. 1987, revised June 1989). A Bahadur - Type representation for empirical quantiles of a large class of stationary, possibly infinite - variance, linear processes

133. DONOHO, D.L. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean, I.

134. CANCELLED

135. FREEDMAN, D.A. and NAVIDI, W. (December 1987). On the risk of lung cancer for ex-smokers.

136. LE CAM, L. (January 1988). On some stochastic models of the effects of radiation on cell survival.

137. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the uniform consistency of Bayes estimates for multinomial probabilities.

137a. DONOHO, D.L. and LIU, R.C. (1987). Geometrizing rates of convergence, I.

138. DONOHO, D.L. and LIU, R.C. (January 1988). Geometrizing rates of convergence, III.

139. BERAN, R. (January 1988). Refining simultaneous confidence sets.

140. HESSE, C.H. (December 1987). Numerical and statistical aspects of neural networks.

141. BRILLINGER, D.R. (Jan. 1988). Two reports on trend analysis: a) An elementary trend analysis of Rio negro levels at Manaus, 1903-1985. b) Consistent detection of a monotonic trend superposed on a stationary time series.

142. DONOHO, D.L. (Jan. 1985, revised Jan. 1988). One-sided inference about functionals of a density.

143. DALANG, R.C. (Feb. 1988, revised Nov. 1988). Randomization in the two-armed bandit problem.
144. DABROWSKA, D.M., DOKSUM, K.A. and SONG, J.K. (February 1988). Graphical comparisons of cumulative hazards for two populations.
145. ALDOUS, D.J. (February 1988). Lower bounds for covering times for reversible Markov Chains and random walks on graphs.
146. BICKEL, P.J. and RITOV, Y. (Feb.1988, revised August 1988). Estimating integrated squared density derivatives.
147. STARK, P.B. (March 1988). Strict bounds and applications.
148. DONOHO, D.L. and STARK, P.B. (March 1988). Rearrangements and smoothing.
149. NOLAN, D. (March 1988). Asymptotics for a multivariate location estimator.
150. SEILLIER, F. (March 1988). Sequential probability forecasts and the probability integral transform.
151. NOLAN, D. (Mar. 1988, revised May 1989). Asymptotics for multivariate trimming.
152. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On a theorem of Kuchler and Lauritzen.
153. DIACONIS, P. and FREEDMAN, D.A. (April 1988). On the problem of types.
154. DOKSUM, K.A. (May 1988). On the correspondence between models in binary regression analysis and survival analysis.
155. LEHMANN, E.L. (May 1988). Jerzy Neyman, 1894-1981.
156. ALDOUS, D.J. (May 1988). Stein's method in a two-dimensional coverage problem.
157. FAN, J. (June 1988). On the optimal rates of convergence for nonparametric deconvolution problem.
158. DABROWSKA, D. (June 1988). Signed-rank tests for censored matched pairs.
159. BERAN, R.J. and MILLAR, P.W. (June 1988). Multivariate symmetry models.
160. BERAN, R.J. and MILLAR, P.W. (June 1988). Tests of fit for logistic models.
161. BREIMAN, L. and PETERS, S. (June 1988). Comparing automatic bivariate smoothers (A public service enterprise).
162. FAN, J. (June 1988). Optimal global rates of convergence for nonparametric deconvolution problem.
163. DIACONIS, P. and FREEDMAN, D.A. (June 1988). A singular measure which is locally uniform. (Revised by Tech Report No. 180).
164. BICKEL, P.J. and KRIEGER, A.M. (July 1988). Confidence bands for a distribution function using the bootstrap.
165. HESSE, C.H. (July 1988). New methods in the analysis of economic time series I.
166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.
167. BREIMAN, L., STONE, C.J. and KOOPERBERG, C. (August 1988). Confidence bounds for extreme quantiles.
168. LE CAM, L. (August 1988). Maximum likelihood an introduction.
169. BREIMAN, L. (Aug.1988, revised Feb. 1989). Submodel selection and evaluation in regression I. The X-fixed case and little bootstrap.
170. LE CAM, L. (September 1988). On the Prokhorov distance between the empirical process and the associated Gaussian bridge.
171. STONE, C.J. (September 1988). Large-sample inference for logspline models.
172. ADLER, R.J. and EPSTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.
173. MILLAR, P.W. (October 1988). Optimal estimation in the non-parametric multiplicative intensity model.
174. YOR, M. (October 1988). Interwinings of Bessel processes.
175. ROJO, J. (October 1988). On the concept of tail-heaviness.
176. ABRAHAMS, D.M. and RIZZARDI, F. (September 1988). BLSS - The Berkeley interactive statistical system: An overview.

177. MILLAR, P.W. (October 1988). Gamma-funnels in the domain of a probability, with statistical implications.

178. DONOHO, D.L. and LIU, R.C. (October 1988). Hardest one-dimensional subproblems.

179. DONOHO, D.L. and STARK, P.B. (October 1988). Recovery of sparse signal when the low frequency information is missing.

180. FREEDMAN, D.A. and PITMAN, J.A. (Nov. 1988). A measure which is singular and uniformly locally uniform. (Revision of Tech Report No. 163).

181. DOKSUM, K.A. and HOYLAND, ARNLJOT (Nov. 1988, revised Jan. 1989, Aug. '89). Models for variable stress accelerated life testing experiments based on Wiener processes and the inverse Gaussian distribution.

182. DALANG, R.C., MORTON, A. and WILLINGER, W. (November 1988). Equivalent martingale measures and no-arbitrage in stochastic securities market models.

183. BERAN, R. (November 1988). Calibrating prediction regions.

184. BARLOW, M.T., PITMAN, J. and YOR, M. (Feb. 1989). On Walsh's Brownian Motions.

185. DALANG, R.C. and WALSH, J.B. (Dec. 1988). Almost-equivalence of the germ-field Markov property and the sharp Markov property of the Brownian sheet.

186. HESSE, C.H. (Dec. 1988). Level-Crossing of integrated Ornstein-Uhlenbeck processes

187. NEVEU, J. and PITMAN, J.W. (Feb. 1989). Renewal property of the extrema and tree property of the excursion of a one-dimensional brownian motion.

188. NEVEU, J. and PITMAN, J.W. (Feb. 1989). The branching process in a brownian excursion.

189. PITMAN, J.W. and YOR, M. (Mar. 1989). Some extensions of the arcsine law.

190. STARK, P.B. (Dec. 1988). Duality and discretization in linear inverse problems.

191. LEHMANN, E.L. and SCHOLZ, F.W. (Jan. 1989). Ancillarity.

192. PEMANTLE, R. (Feb. 1989). A time-dependent version of Pólya's urn.

193. PEMANTLE, R. (Feb. 1989). Nonconvergence to unstable points in urn models and stochastic approximations.

194. PEMANTLE, R. (Feb. 1989, revised May 1989). When are touchpoints limits for generalized Pólya urns.

195. PEMANTLE, R. (Feb. 1989). Random walk in a random environment and first-passage percolation on trees.

196. BARLOW, M., PITMAN, J. and YOR, M. (Feb. 1989). Une extension multidimensionnelle de la loi de l'arc sinus.

197. BREIMAN, L. and SPECTOR, P. (Mar. 1989). Submodel selection and evaluation in regression — the X-random case.

198. BREIMAN, L., TSUR, Y. and ZEMEL, A. (Mar. 1989). A simple estimation procedure for censored regression models with known error distribution.

199. BRILLINGER, D.R. (Mar. 1989). Two papers on bilinear systems: a) A study of second- and third-order spectral procedures and maximum likelihood identification of a bilinear system. b) Some statistical aspects of NMR spectroscopy, *Actas del 2° congreso lantinoamericano de probabilidad y estadistica matematica*, Caracas, 1985.

200. BRILLINGER, D.R. (Mar. 1989). Two papers on higher-order spectra: a) Parameter estimation for nonGaussian processes via second and third order spectra with an application to some endocrine data. b) Some history of the study of higher-order moments and spectra.

201. DE LA PENA, V. and KLASS, M.J. (April 1989). L bounds for quadratic forms of independent random variables.

202. FREEDMAN, D.A. and NAVIDI, W.C. (April 1989). Testing the independence of competing risks.

203. TERDIK, G. (May 1989). Bilinear state space realization for polynomial stochastic systems.

204. DONOHO, D.L. and JOHNSTONE, I.M. (May 1989). Minimax risk over l_p-Balls.

205. PEMANTLE, R., PROPP, J. and ULLMAN, D. (May 1989). On tensor powers of integer programs.

206. MILASEVIC, P. and NOLAN, D. (May 1989). Estimation on the sphere: A geometric approach.

207. SPEED, T.P. and YU, B. (July 1989). Stochastic complexity and model selection: normal regression.

208. DUBINS, L.E. (June 1989). A group decision device: Its pareto-like optimality.

209. BREIMAN, L. (July 1989). Fitting additive models to regression data.

210. PEMANTLE, R. (July 1989). Vertex-reinforced random walk.

211. LE CAM, L. (August 1989). On measurability and convergence in distribution.

212. FELDMAN, R.E. (July 1989). Autoregressive processes and first-hit probabilities for randomized random walks.

213. DONOHO, D.L., JOHNSTONE, I.M., HOCH, J.C. and STERN, A.S. (August 1989). Maximum entropy and the nearly black object.

Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:

Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.