

SUMMARY OF RESULTS ON PATH SPACES AND CONVERGENCE IN DISTRIBUTION FOR STOCHASTIC PROCESSES

RUTH J. WILLIAMS *

April 3, 2014

*Department of Mathematics, University of California, San Diego, 9500 Gilman Drive,
La Jolla CA 92093-0112, USA, email: williams@math.ucsd.edu

Contents

A	Notation	3
B	Stochastic Processes	3
C	Path Spaces	4
C.1	Definitions and Topologies for $\mathbb{C}^d$ and $\mathbb{D}^d$	4
C.2	Compact Sets in $\mathbb{D}^d$	5
D	Weak Convergence of Probability Measures	7
D.1	General Definitions and Results	7
D.2	Tightness of Probability Measures on $\mathbb{D}^d$	8
E	Convergence in Distribution for Stochastic Processes	9
E.1	Types of Convergence	9
E.2	Tightness and Continuity of Limit Processes	9
E.3	Continuous Mapping Theorem	11
E.4	Random Time Change	11
F	Functional Central Limit Theorems	12

A Notation

For each integer $d \geq 1$, let $\mathbb{R}^d$ denote d -dimensional Euclidean space and let $\mathbb{R}_+^d$ denote the d -dimensional non-negative orthant

$$\{x \in \mathbb{R}^d : x = (x_1, \dots, x_d), x_i \geq 0 \text{ for } i = 1, \dots, d\}.$$

We will use the following norm on $\mathbb{R}^d$:

$$|x| \equiv \max_{1 \leq i \leq d} |x_i|, \quad x \in \mathbb{R}^d. \quad (1)$$

We endow $\mathbb{R}^d$ with the σ -algebra of Borel sets. When $d = 1$, we will omit the qualifying superscript d .

B Stochastic Processes

A d -dimensional stochastic process is a collection of random variables $X = \{X(t), t \geq 0\}$ defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathbb{R}^d$. In particular, for each $t \geq 0$, the function $X(t) : \Omega \rightarrow \mathbb{R}^d$ is measurable, where Ω is endowed with the σ -algebra $\mathcal{F}$ and $\mathbb{R}^d$ is endowed with the Borel σ -algebra. We shall often write $X(t, \omega)$ or $X_t(\omega)$ in place of $X(t)(\omega)$ for $t \geq 0, \omega \in \Omega$. Such a stochastic process X is said to have r.c.l.l. (sample) paths if for each $\omega \in \Omega$, the function $t \rightarrow X(t, \omega)$ from $[0, \infty)$ into $\mathbb{R}^d$ is right continuous on $[0, \infty)$ and has finite limits from the left on $(0, \infty)$. Here r.c.l.l. stands for “right continuous with finite left limits”. The acronym càdlàg for the equivalent French phrase is used by some authors instead of r.c.l.l. The stochastic process X is said to have continuous (sample) paths if for each $\omega \in \Omega$, the function $t \rightarrow X(t, \omega)$ is continuous from $[0, \infty)$ into $\mathbb{R}^d$.

We will be concerned with the construction of, and convergence in distribution of, d -dimensional stochastic processes having continuous paths or r.c.l.l. paths. The next few sections summarize some basic definitions and properties needed for this. For more details, the reader is referred to [Billingsley \(1999\)](#) or [Ethier and Kurtz \(1986\)](#) or [Jacod and Shiryaev \(1987\)](#).

C Path Spaces

C.1 Definitions and Topologies for $\mathbb{C}^d$ and $\mathbb{D}^d$

For $d \geq 1$, let $\mathbb{C}^d \equiv C([0, \infty), \mathbb{R}^d)$ denote the space of continuous functions from $[0, \infty)$ into $\mathbb{R}^d$. When $d = 1$, we shall suppress the superscript d . We endow $\mathbb{C}^d$ with the topology of uniform convergence on compact time intervals. Let $\mathbb{D}^d$ denote the space of functions from $[0, \infty)$ into $\mathbb{R}^d$ that are right continuous on $[0, \infty)$ and have finite left limits on $(0, \infty)$. When $d = 1$, we shall suppress the superscript d . An element $x \in \mathbb{D}^d$ only has discontinuities of jump type and there are only countably many points in $(0, \infty)$ where x has a jump discontinuity. We endow $\mathbb{D}^d$ with the (Skorokhod) J_1 -topology. There is a metric m_{J_1} on $\mathbb{D}^d$ which induces this topology and under which the space is a complete, separable metric space (i.e., a Polish space). For our purposes, we do not need to know the precise form of this metric, rather it will suffice to characterize convergence of sequences in the J_1 -topology. For this, let Γ denote the set of functions $\gamma : [0, \infty) \rightarrow [0, \infty)$ that are strictly increasing and continuous with $\gamma(0) = 0$ and $\lim_{t \rightarrow \infty} \gamma(t) = \infty$. (In particular, γ maps $[0, \infty)$ onto $[0, \infty)$.) A sequence $\{x_n\}_{n=1}^\infty$ in $\mathbb{D}^d$ converges to $x \in \mathbb{D}^d$ in the J_1 -topology if and only if for each $T > 0$ there is a sequence $\{\gamma_n\}_{n=1}^\infty$ (possibly depending on T) in Γ such that

$$\sup_{0 \leq t \leq T} |\gamma_n(t) - t| \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad \text{and} \quad (2)$$

$$\sup_{0 \leq t \leq T} |x_n(\gamma_n(t)) - x(t)| \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (3)$$

If x is in $\mathbb{C}^d$, $\{x_n\}_{n=1}^\infty$ converges to x in the J_1 -topology if and only if $\{x_n\}_{n=1}^\infty$ converges to x uniformly on compact time intervals. In particular, with the topologies described above, $\mathbb{C}^d$ is a topological subspace of $\mathbb{D}^d$. For later use, for each $T \geq 0$, we define

$$\|x\|_T = \sup_{t \in [0, T]} |x(t)|, \quad \text{for } x \in \mathbb{D}^d. \quad (4)$$

We note in passing that one may endow $\mathbb{D}^d$ with the topology of uniform convergence on compact time intervals. However, this topology is finer than the J_1 -topology and $\mathbb{D}^d$ with this topology is not separable.

Remark. The reader is cautioned here that for positive integers d and k , the product space $\mathbb{C}^d \times \mathbb{C}^k$ is the same topologically as the space $\mathbb{C}^{d+k}$, because a sequence of functions $\{x^n\}$ in $\mathbb{C}^d$ say, converges uniformly on each compact time interval if and only if each component function $x_i^n, i = 1, \dots, d$, converges uniformly on each compact time interval. On the other hand, the product space $\mathbb{D}^d \times \mathbb{D}^k$ is not the same topologically as the space $\mathbb{D}^{d+k}$, because in defining the J_1 -topology a change of time scale (given by functions γ_n , cf. (2)), is used and this may be different for two sequences taking values in the two spaces $\mathbb{D}^d$ and $\mathbb{D}^k$, respectively.

For the purpose of defining measurability, we endow $\mathbb{C}^d$ with the Borel σ -algebra associated with the topology of uniform convergence on compact time intervals; this agrees with the σ -algebra $\mathcal{C}^d = \sigma\{x(t) : x \in \mathbb{C}^d, 0 \leq t < \infty\}$ which is the smallest σ -algebra on $\mathbb{C}^d$ such that for each $t \geq 0$, the projection mapping $x \rightarrow x(t)$, from $\mathbb{C}^d$ into $\mathbb{R}^d$, is measurable. We endow the space $\mathbb{D}^d$ with the Borel σ -algebra associated with the J_1 -topology; this agrees with the σ -algebra $\mathcal{D}^d = \sigma\{x(t) : x \in \mathbb{D}^d, 0 \leq t < \infty\}$. As usual, when $d = 1$, we shall omit the superscript d .

Consider a d -dimensional stochastic process X defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If X has r.c.l.l. (resp. continuous) paths, then $\omega \rightarrow X(\cdot, \omega)$ defines a measurable mapping from $(\Omega, \mathcal{F})$ into $(\mathbb{D}^d, \mathcal{D}^d)$ (resp. $(\mathbb{C}^d, \mathcal{C}^d)$) and it induces a probability measure π on $(\mathbb{D}^d, \mathcal{D}^d)$ (resp. $(\mathbb{C}^d, \mathcal{C}^d)$) via $\pi(A) = P(\{\omega : X(\cdot, \omega) \in A\})$ for all $A \in \mathcal{D}^d$ (resp. $\mathcal{C}^d$). This probability measure π is called the *law* of X .

C.2 Compact Sets in $\mathbb{D}^d$

To develop a criterion for relative compactness of probability measures associated with d -dimensional stochastic processes having r.c.l.l. paths, we need a

characterization of the (relatively) compact sets in $\mathbb{D}^d$ with the J_1 -topology.

Definition C.1. For each $x \in \mathbb{D}^d$, $\delta > 0$ and $T > 0$, define

$$w'(x, \delta, T) = \inf_{\{t_i\}} \max_i \sup_{s, t \in [t_{i-1}, t_i)} |x(s) - x(t)|$$

where $\{t_i\}$ ranges over all partitions of the form $0 = t_0 < \dots < t_{n-1} < T \leq t_n$ with $\min(t_i - t_{i-1}) > \delta$ and $n \geq 1$.

Remark. This definition is slightly different from that in Billingsley (1999) and is taken from Ethier and Kurtz (1986), p. 122. The notation $w'(x, \delta, T)$ is used rather than simply $w(x, \delta, T)$ because $w(x, \delta, T)$ is commonly used for the usual modulus of continuity. The quantity $w'(x, \delta, T)$ is often called a modified modulus of continuity.

Proposition C.2. For each $x \in \mathbb{D}^d$ and $T > 0$,

$$\lim_{\delta \rightarrow 0} w'(x, \delta, T) = 0.$$

For each $\delta > 0$ and $T > 0$, $w'(\cdot, \delta, T)$ is a Borel measurable function from $\mathbb{D}^d$ into $\mathbb{R}$.

Proof. See Ethier and Kurtz (1986), Lemma 3.6.2. □

Proposition C.3. A set $A \subset \mathbb{D}^d$ is relatively compact if and only if the following two conditions hold for each $T > 0$:

- (i) $\sup_{x \in A} \|x\|_T < \infty$,
- (ii) $\lim_{\delta \rightarrow 0} \sup_{x \in A} w'(x, \delta, T) = 0$.

Proof. See Ethier and Kurtz (1986), Theorem 3.6.3. □

Remark. One can replace condition (i) by the following and then the same result holds:

- (i)' for each rational $t \in [0, T]$, $\sup_{x \in A} |x(t)| < \infty$.

D Weak Convergence of Probability Measures

D.1 General Definitions and Results

Let $(\mathbb{M}, m)$ be a complete separable metric space where $\mathbb{M}$ denotes the set and m is the metric. Let $\mathcal{M}$ denote the σ -algebra of Borel sets associated with the topology induced on $\mathbb{M}$ by m .

Definition D.1. A sequence of probability measures $\{\pi_n\}_{n=1}^\infty$ on $(\mathbb{M}, \mathcal{M})$ converges weakly to a probability measure π on $(\mathbb{M}, \mathcal{M})$ if and only if

$$\int_{\mathbb{M}} f d\pi_n \longrightarrow \int_{\mathbb{M}} f d\pi \quad \text{as } n \rightarrow \infty$$

for each bounded continuous function $f : \mathbb{M} \rightarrow \mathbb{R}$.

Definition D.2. A family of probability measures Π on $(\mathbb{M}, \mathcal{M})$ is (weakly) relatively compact if each sequence $\{\pi_n\}_{n=1}^\infty$ in Π has a subsequence that converges weakly to a probability measure π on $(\mathbb{M}, \mathcal{M})$.

Definition D.3. A family of probability measures Π on $(\mathbb{M}, \mathcal{M})$ is tight if for each $\varepsilon > 0$ there is a compact set A in $\mathbb{M}$ such that

$$\pi(A) > 1 - \varepsilon \quad \text{for all } \pi \in \Pi.$$

Theorem D.4 (Prohorov's Theorem). *A family of probability measures on $(\mathbb{M}, \mathcal{M})$ is tight if and only if it is (weakly) relatively compact.*

Proof. See Billingsley (1999), Theorems 5.1 and 5.2, or Ethier and Kurtz (1986), Theorem 3.2.2. $\square$

Remark. The “if part” of this proposition uses the fact that the metric space $(\mathbb{M}, m)$ is separable and complete.

Corollary D.5. *Suppose that $\{\pi_n\}_{n=1}^\infty$ is a tight family of probability measures on $(\mathbb{M}, \mathcal{M})$ and that there is a probability measure π on $(\mathbb{M}, \mathcal{M})$ such that each weakly convergent subsequence of $\{\pi_n\}_{n=1}^\infty$ has limit π . Then $\{\pi_n\}_{n=1}^\infty$ converges to π .*

Proof. See Billingsley (1999), p. 59. $\square$

The following theorem is often useful for reducing arguments about convergence of probability laws associated with stochastic processes to real analysis arguments based on almost sure convergence of equivalent distributional representatives for those processes.

Theorem D.6 (Skorokhod Representation Theorem). *Suppose that π and $\{\pi_n\}_{n=1}^\infty$ are all probability measures on $(\mathbb{M}, \mathcal{M})$ and that $\{\pi_n\}_{n=1}^\infty$ converges weakly to π . Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which are defined $\mathbb{M}$ -valued random variables X and $\{X^n\}_{n=1}^\infty$ such that X has distribution π , X^n has distribution π_n for $n = 1, 2, \dots$, and $X^n \rightarrow X$ $\mathbb{P}$ -a.s. as $n \rightarrow \infty$.*

Proof. See Ethier and Kurtz (1986), Theorem 3.1.8. $\square$

D.2 Tightness of Probability Measures on $\mathbb{D}^d$

Combining the characterization of relative compactness in $\mathbb{D}^d$ with Prohorov's Theorem yields the following.

Theorem D.7. *A sequence of probability measures $\{\pi_n\}_{n=1}^\infty$ on $(\mathbb{D}^d, \mathcal{D}^d)$ is tight if and only if for each $T > 0$ and $\varepsilon > 0$,*

$$(i) \quad \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \pi_n(\{x \in \mathbb{D}^d : \|x\|_T \geq K\}) = 0,$$

$$(ii) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \pi_n(\{x \in \mathbb{D}^d : w'(x, \delta, T) \geq \varepsilon\}) = 0.$$

Proof. See Ethier and Kurtz (1986), Theorem, 3.7.4, or Billingsley (1999), Theorem 16.8. $\square$

Remark. Condition (i) can be replaced by the following and then the above theorem still holds:

$$(i)' \quad \text{for each rational } t \in [0, T], \lim_{K \rightarrow \infty} \limsup_n \pi_n(\{x \in \mathbb{D}^d : |x(t)| \geq K\}) = 0.$$

E Convergence in Distribution for Stochastic Processes

E.1 Types of Convergence

Suppose that $X, X^n, n = 1, 2, \dots$, are d -dimensional stochastic processes with r.c.l.l. paths (these processes may be defined on different probability spaces). Let π denote the law of X and for each $n = 1, 2, \dots$, let π_n denote the law of X^n . The sequence of processes $\{X^n\}_{n=1}^\infty$ *converges in distribution* to X if and only if $\{\pi_n\}_{n=1}^\infty$ converges weakly to π . Some authors abuse terminology and say that $\{X^n\}_{n=1}^\infty$ converges weakly to X . We denote such convergence in distribution by $X^n \Rightarrow X$ as $n \rightarrow \infty$.

Suppose that X and $\{X^n\}_{n=1}^\infty$ are all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then $\{X^n\}_{n=1}^\infty$ converges to X almost surely (resp. in probability) if and only if $\lim_{n \rightarrow \infty} m_{J_1}(X^n(\omega), X(\omega)) = 0$ for $\mathbb{P}$ -almost every $\omega \in \Omega$ (resp. for each $\varepsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}(\omega \in \Omega : m_{J_1}(X^n(\omega), X(\omega)) \geq \varepsilon) = 0$). (Here m_{J_1} is the metric introduced previously which induces the J_1 -topology on $\mathbb{D}^d$.) We denote the almost sure convergence by $X^n \rightarrow X$ a.s. as $n \rightarrow \infty$, and we denote the convergence in probability by $X^n \rightarrow X$ in prob. as $n \rightarrow \infty$. If $\{X^n\}_{n=1}^\infty$ converges to X almost surely or in probability, then $\{X^n\}_{n=1}^\infty$ converges in distribution to X . Conversely, if $\{X^n\}_{n=1}^\infty$ converges in distribution to X and X is a.s. constant, then $\{X^n\}_{n=1}^\infty$ converges in probability to X as $n \rightarrow \infty$.

E.2 Tightness and Continuity of Limit Processes

The following criterion due to Aldous often provides a convenient mechanism for verifying tightness of the laws associated with d -dimensional stochastic processes having r.c.l.l. paths.

Theorem E.1. *Let $\{X^n\}_{n=1}^\infty$ be a sequence of d -dimensional stochastic processes with r.c.l.l. paths. Then the probability measures induced on $\mathbb{D}^d$ by $\{X^n\}_{n=1}^\infty$ are tight if the following two conditions hold for each $T > 0$:*

$$(i) \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|X^n\|_T \geq K) = 0,$$

(ii) for each $\varepsilon > 0, \eta > 0$, there are positive constants $\delta_{\varepsilon, \eta}$ and $n_{\varepsilon, \eta}$ such that for all $0 < \delta \leq \delta_{\varepsilon, \eta}$ and $n \geq n_{\varepsilon, \eta}$,

$$\sup_{\tau \in \mathcal{T}_{[0, T]}^f} \mathbb{P}(|X^n(\tau + \delta) - X^n(\tau)| \geq \varepsilon) \leq \eta$$

where $\mathcal{T}_{[0, T]}^f$ denotes the set of all stopping times relative to the filtration generated by X^n that take values in a finite subset of $[0, T]$.

Proof. See Billingsley (1999), Theorem 16.10. $\square$

The following proposition provides a useful criterion for checking when a sequence of d -dimensional stochastic processes with r.c.l.l. paths has associated laws that are tight and whose limit points are concentrated on the set of continuous paths $\mathbb{C}^d$. (We say such a sequence of laws is C -tight and (with an abuse of terminology) we sometimes say the sequence of processes is C -tight.)

Proposition E.2. *Let $\{X^n\}_{n=1}^\infty$ be a sequence of d -dimensional processes with r.c.l.l. paths. Then the sequence of probability measures $\{\pi_n\}_{n=1}^\infty$ induced on $\mathbb{D}^d$ by $\{X^n\}_{n=1}^\infty$ is tight and any weak limit point of this sequence is concentrated on $\mathbb{C}^d$ if and only if the following two conditions hold for each $T > 0$ and $\varepsilon > 0$:*

$$(i) \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P(\|X^n\|_T \geq K) = 0,$$

$$(ii) \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P(w(X^n, \delta, T) \geq \varepsilon) = 0,$$

where for $x \in \mathbb{D}^d$,

$$w(x, \delta, T) = \sup \left\{ \sup_{u, v \in [t, t+\delta]} |x(u) - x(v)| : 0 \leq t < t + \delta \leq T \right\}. \quad (5)$$

Proof. See Proposition VI.3.26 in Jacod and Shiryaev (1987). $\square$

E.3 Continuous Mapping Theorem

Theorem E.3 (Continuous Mapping Theorem). *For fixed positive integers d and k , let $h_n, n = 1, 2, \dots, h$ be measurable functions from $(\mathbb{D}^d, \mathcal{D}^d)$ into $(\mathbb{D}^k, \mathcal{D}^k)$. Define*

$$C_h = \{x \in \mathbb{D}^d : h_n(x_n) \rightarrow h(x) \text{ in } \mathbb{D}^k \text{ whenever } x_n \rightarrow x \text{ in } \mathbb{D}^d\}.$$

Let $X, X^n, n = 1, 2, \dots$, be d -dimensional stochastic processes with r.c.l.l. paths such that $X^n \Rightarrow X$ as $n \rightarrow \infty$ and $\mathbb{P}(X \in C_h) = 1$. Then $h_n(X^n) \Rightarrow h(X)$ as $n \rightarrow \infty$.

Exercise E.4. Prove the continuous mapping theorem. Hint: use the Skorohod Representation Theorem.

Exercise E.5. Verify that the following mapping is continuous. Consider $h : \mathbb{D}^d \rightarrow \mathbb{D}$ defined by

$$h(x)(t) = \sup_{0 \leq s \leq t} \|x(s)\|, \quad t \geq 0.$$

E.4 Random Time Change

Let

$$\mathbb{D}_o = \{\phi \in \mathbb{D} : \phi \text{ is non-decreasing and } 0 \leq \phi(t) < \infty \text{ for all } t \geq 0\}.$$

Note that if $x \in \mathbb{D}^d$ and $\phi \in \mathbb{D}_o$, then $x \circ \phi \in \mathbb{D}^d$.

Theorem E.6 (Random time change theorem). *Suppose that $\{X^n\}_{n=1}^\infty, X$ are d -dimensional stochastic processes with r.c.l.l. paths and $\{\Phi^n\}_{n=1}^\infty, \Phi$ are one-dimensional stochastic processes whose paths are in $\mathbb{D}_o$. If $(X^n, \Phi^n) \Rightarrow (X, \Phi)$ as $n \rightarrow \infty$ and $P(X \in \mathbb{C}^d) = 1$, then*

$$X^n \circ \Phi^n \Rightarrow X \circ \Phi \quad \text{as } n \rightarrow \infty.$$

Exercise E.7. Prove this theorem using the continuous mapping theorem by showing that the mapping $\psi : \mathbb{D}^d \times \mathbb{D}_o \rightarrow \mathbb{D}^d$ given by $\psi(x, \phi) = x \circ \phi$ is continuous at $x \in \mathbb{C}^d$.

F Functional Central Limit Theorems

Theorem F.1 (Donsker's Theorem). *Let $v_1, v_2, \dots$ be independent, identically distributed (i.i.d.) real-valued random variables with mean $\mu \in (-\infty, \infty)$ and variance $\sigma^2 \in (0, \infty)$. Let*

$$\widehat{V}^n(t) = \frac{1}{\sigma\sqrt{n}} \left(\sum_{i=1}^{[nt]} v_i - \mu nt \right), \quad t \geq 0. \quad (6)$$

Then $\widehat{V}^n \Rightarrow W$, where W is a standard one-dimensional Brownian motion.

Proof. This is usually shown in two steps: (i) convergence of finite dimensional distributions, and (ii) tightness. See Billingsley (1999), Theorem 14.1, or Ethier and Kurtz (1986), Theorem 5.1.2. $\square$

Remark. This result is also called a “Functional Central Limit Theorem (FCLT)” or “Invariance Principle”.

Remark. There are extensions of Donsker's theorem that relax the i.i.d. requirements.

Theorem F.2. (Functional Central Limit Theorem for Renewal Processes) *Let $u_1, u_2, \dots$ be i.i.d. positive random variables with mean $\lambda^{-1} \in (0, \infty)$ and variance $\sigma^2 \in (0, \infty)$. For each $t \geq 0$, define*

$$N(t) = \sup\{k \geq 0 : u_1 + \dots + u_k \leq t\}$$

and set

$$\widehat{N}^n(t) = \frac{1}{\lambda^{3/2}\sigma\sqrt{n}} (N(nt) - \lambda nt).$$

Then $\widehat{N}^n \Rightarrow W$ as $n \rightarrow \infty$, where W is a standard one-dimensional Brownian motion.

Proof. This follows from Donsker's theorem by a clever inversion argument, see Billingsley (1999), Theorem 14.6. $\square$

References

- P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, second edition, 1999.
- S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, New York, 1986.
- J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, New York, 1987.