Math 639: Lecture 8

Limit laws, introduction to random walk

Bob Hough

February 21, 2017

Poisson processes

Theorem

Let $X_{n,m}$, $1 \leqslant m \leqslant n$ be independent nonnegative integer valued random variables with $\text{Prob}(X_{n,m} = 1) = p_{n,m}$, $\text{Prob}(X_{n,m} \geqslant 2) = \epsilon_{n,m}$.

1. $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0, \infty)$
2. $\max_{1 \leqslant m \leqslant n} p_{m,n} \to 0$
3. $\sum_{m=1}^{n} \epsilon_{n,m} \to 0$

If $S_n = X_{n,1} + \cdots + X_{n,n}$ then $S_n \Rightarrow Z$ where Z is $\text{Poisson}(\lambda)$.

Proof.

- Set $X'_{n,m} = 1$ if $X_{n,m} = 1$ and 0 otherwise.
- Let $S'_n = X'_{n,1} + \cdots + X'_{n,n}$.
- The conditions imply $S'_n \Rightarrow Z$, and $\mathrm{Prob}(S_n \neq S'_n) \to 0$.

$\square$

Poisson processes

Theorem

Let $N(s,t)$ be the number of arrivals at a bank in time interval $(s,t]$. Suppose

1. *The numbers of arrivals in disjoint intervals are independent*
2. *The distribution of $N(s,t)$ only depends on $t-s$*
3. $\mathrm{Prob}(N(0,h)=1) = \lambda h + o(h)$
4. $\mathrm{Prob}(N(0,h)\geqslant 2) = o(h)$.

Then $N(0,t)$ has a Poisson distribution with mean λt.

Proof.

Let $X_{n,m} = N\left(\frac{(m-1)t}{n}, \frac{mt}{n}\right)$ for $1 \leqslant m \leqslant n$ and apply the previous theorem. $\square$

Poisson processes

Definition

A family of random variables N_t, $t \geqslant 0$ satisfying

1. If $0 = t_0 < t_1 < \cdots < t_n$, $N(t_k) - N(t_{k-1})$, $1 \leqslant k \leqslant n$ are independent
2. $N(t) - N(s)$ is Poisson$(\lambda(t-s))$.

is called a *Poisson process with rate* λ.

Poisson processes

Theorem

Let $\xi_1, \xi_2, \ldots$ be independent random variables with $\mathrm{Prob}(\xi_i > t) = e^{-\lambda t}$ for $t \geqslant 0$. Let $T_n = \xi_1 + \cdots + \xi_n$ with $T_0 = 0$ and $N_t = \sup\{n : T_n \leqslant t\}$. Then N_t is a Poisson process of parameter λ.

Poisson processes

Proof.

- One may check that T_n has density $f_n(s) = \frac{\lambda^n s^{n-1}}{(n-1)!} e^{-\lambda s}$.
- Now check by induction that

$$\begin{aligned}
\mathrm{Prob}(N_t = 0) &= \mathrm{Prob}(T_1 > t) = e^{-\lambda t}, \\
\mathrm{Prob}(N_t = n) &= \mathrm{Prob}(T_n \leqslant t < T_{n+1}) \\
&= \int_0^t \mathrm{Prob}(T_n = s)\, \mathrm{Prob}(\xi_{n+1} > t - s) ds \\
&= \int_0^t \frac{\lambda^n s^{n-1}}{(n-1)!} e^{-\lambda s} e^{-\lambda(t-s)} ds = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.
\end{aligned}$$

Thus N_t has a Poisson distribution with mean λt.

$\square$

Proof.

- Observe

$$\mathrm{Prob}(T_{n+1} \geqslant u | N_t = n) = \frac{\mathrm{Prob}(T_{n+1} \geqslant u, T_n \leqslant t)}{\mathrm{Prob}(N_t = n)}.$$

- Calculate

$$\begin{aligned}\mathrm{Prob}(T_{n+1} \geqslant u, T_n \leqslant t) &= \int_0^t f_n(s)\,\mathrm{Prob}(\xi_{n+1} \geqslant u - s) ds \\ &= \int_0^t \frac{\lambda^n s^{n-1}}{(n-1)!} e^{-\lambda s} e^{-\lambda(u-s)} ds = e^{-\lambda u} \frac{(\lambda t)^n}{n!}.\end{aligned}$$

Since $\mathrm{Prob}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$,
$\mathrm{Prob}(T_{n+1} \geqslant u | N_t = n) = \frac{e^{-\lambda u}}{e^{-\lambda t}} = e^{-\lambda(u-t)}$.

□

Proof.

- Let $T_1' = T_{N(t)+1} - t$ and $T_k' = T_{N(t)+k} - T_{N(t)+k-1}$ for $k \geqslant 2$. Then $T_1', T_2', \ldots$ are i.i.d. and independent of N_t. Hence the arrivals after time t are independent of N_t and have the same distribution as the original sequence.
- Hence if $0 = t_0 < t_1 < \cdots < t_n$ then $N(t_i) - N(t_{i-1})$, $i = 1, 2, \ldots, n$ are independent.

$\square$

Definition

A function L is said to be *slowly varying* if

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1, \qquad \text{for all } t > 0.$$

One may check that $L(t) = \log t$ is slowly varying, but $L(t) = t^{\epsilon}$ is not for any $\epsilon \neq 0$.

Stable laws

Theorem

Suppose $X_1, X_2, \ldots$ are i.i.d. with a distribution that satisfies

1. $\lim_{x\to\infty} \frac{\mathrm{Prob}(X_1 > x)}{\mathrm{Prob}(|X_1| > x)} = \theta \in [0,1]$.
2. $\mathrm{Prob}(|X_1| > x) = x^{-\alpha} L(x)$

where $0 < \alpha < 2$ and L is slowly varying. Let $S_n = X_1 + \cdots + X_n$,

$$a_n = \inf\{x : \mathrm{Prob}(|X_1| > x) \leqslant n^{-1}\}, \qquad b_n = n\,\mathrm{E}[X_1 \mathbf{1}(|X_1| \leqslant a_n)].$$

As $n \to \infty$, $\frac{S_n - b_n}{a_n} \Rightarrow Y$ where Y has a nondegenerate distribution.

For a proof, see Durrett pp. 161-162.

Stable laws

Definition

A random variable Y is said to have a *stable law* if for every integer $k > 0$ there are constants a_k and b_k so that if $Y_1, \ldots, Y_k$ are i.i.d. and have the same distribution as Y, then $(Y_1 + \cdots + Y_k - b_k)/a_k =_d Y$.

Theorem

Y is the limit of $(X_1 + \cdots + X_k - b_k)/a_k$ for some i.i.d. sequence X_i if and only if Y has a stable law.

Stable laws

Proof.

- If Y has a stable law, we can take $X_1, X_2, \ldots$ i.i.d. with distribution Y.
- Let $Z_n = \frac{1}{a_n}(X_1 + \cdots + X_n - b_n)$ and $S_n^j = X_{(j-1)n+1} + \cdots + X_{jn}$.
- Thus

$$\begin{aligned} Z_{nk} &= (S_n^1 + \cdots + S_n^k - b_{nk})/a_{nk} \\ a_{nk} Z_{nk} &= (S_n^1 - b_n) + \cdots + (S_n^k - b_n) + (kb_n - b_{nk}) \\ a_{nk} Z_{nk}/a_n &= (S_n^1 - b_n)/a_n + \cdots + (S_n^k - b_n)/a_n + (kb_n - b_{nk})/a_n. \end{aligned}$$

- Let $n \to \infty$. The first k terms on the right tend to $Y_1, \ldots, Y_k$ which are independent copies of Y, and $Z_{nk} \Rightarrow Y$, thus the result follows.

□

Infinitely divisible distributions

Definition

A probability distribution μ is *infinitely divisible* if, for each $n \geqslant 1$, there is probability distribution μ_n such that $\mu = \mu_n^{*n}$.

Measures of compound Poisson type

A large family of infinitely divisible measures is given as follows.

Definition

Let μ be a probability measure with characteristic function $\psi(t)$, and let $\lambda \geqslant 0$ be a parameter. Define μ^{*0} to be the point mass at 0. The probability measure of *compound Poisson type* with parameters μ and λ is the probability measure

$$P(\mu, \lambda) = e^{-\lambda} \sum_{n \geqslant 0} \frac{\lambda^n \mu^{*n}}{n!}.$$

It has characteristic function $\chi(t) = \mathrm{E}_{P(\mu,\lambda)}[e^{it}] = e^{\lambda(\phi(t)-1)}$.

Infinitely divisible distributions

The following discussion is taken from Feller, vol 2.

Definition

A measure μ is *canonical* if the integrals

$$M^+(x) = \int_{x^-}^{\infty} \frac{d\mu(y)}{y^2}, \qquad M^-(-x) = \int_{-\infty}^{-x^+} \frac{d\mu(y)}{y^2}$$

converge for all $x > 0$.

A sequence of measures $c_n x^2 d\mu_n(x)$ *converge properly* to $d\mu(x)$ if it converges to $d\mu(x)$ in distribution, and if, for all $\epsilon > 0$, there exists $\tau > 0$ such that for $x > \tau$,

$$\limsup_n c_n \left[1 - \int_{-x}^{x} d\mu_n(x)\right] < \epsilon.$$

Infinitely divisible distributions

Lemma

If $c_n x^2 d\mu_n(x) \to d\mu(x)$ properly, then

$$c_n \int_{-\infty}^{\infty} z(x) d\mu_n(x) \to \int_{-\infty}^{\infty} x^{-2} z(x) d\mu(x).$$

for every bounded continuous function z such that $x^{-2}z(x)$ is continuous at the origin.

Lemma

Consider a sequence of probability measures μ_n, with characteristic functions ϕ_n, and define, for some sequence of constants $\{c_n\}, \{\beta_n\}$,

$$\psi_n(z) = c_n[\phi_n(z) - 1 - i\beta_n z].$$

Suppose that $\psi_n(z) \to \rho(z)$ uniformly in $|z| \leqslant z_0$. Then for $0 < h \leqslant z_0$,

$$c_n \int_{-\infty}^{\infty} \left(1 - \frac{\sin xh}{xh}\right) d\mu_n(x) \to -\frac{1}{2h} \int_{-h}^{h} \rho(z) dz.$$

Infinitely divisible distributions

Proof.

Define $b_n = \int_{-\infty}^{\infty} \sin x d\mu_n(x)$ and write

$$\psi_n(z) = \int_{-\infty}^{\infty} c_n[e^{izx} - 1 - iz \sin x] d\mu_n(x) + ic_n(b_n - \beta_n)z.$$

Divide by $-2h$ and integrate in $|z| \leqslant h$ to obtain the claim. $\square$

Infinitely divisible distributions

Lemma

Under the conditions of the previous lemma, there exists a canonical measure μ and a sequence $n_1, n_2, \ldots \to \infty$ such that $c_{n_k} x^2 d\mu_{n_k}(x) \to d\mu(x)$ properly.

Proof.

- Put $d\nu_n(x) = c_n x^2 d\mu_n(x)$.
- Since $\left(1 - \frac{\sin xh}{xh}\right) \sim \frac{1}{6}x^2h^2$ for x small, and is positive in any case, it follows that $\nu_n(I) < \infty$ for all finite intervals I.
- By Helly's selection theorem, there is a subsequence $\{\nu_{n_k}\}$ converging in distribution to a measure ν.
- To prove the tightness, note that $\rho(0) = 0$. By choosing h sufficiently small, $-\frac{1}{2h}\int_{-h}^{h}\rho(z)dz$ may be made arbitrarily small. Since the integrand on the left is $\geqslant \frac{1}{2}$ for $|xh| \geqslant 2$ and is non-negative otherwise, the result follows.

$\square$

Infinitely divisible distributions

Lemma

For any canonical measure μ the integral defined by

$$\psi(z) = \int_{-\infty}^{\infty} \frac{e^{izx} - 1 - iz\sin x}{x^2} d\mu(x)$$

defines a continuous function, and to distinct canonical measures there correspond distinct functions.

Infinitely divisible distributions

Proof.

- The integral converges by the property of a canonical measure.
- The continuity is immediate.
- Write, for $h > 0$,

$$\psi(z) - \frac{\psi(z+h) + \psi(z-h)}{2} = \int_{-\infty}^{\infty} e^{izx} \frac{1 - \cos xh}{x^2} d\mu(x).$$

  This is the characteristic function of the bounded measure $dA_h(x) = \frac{1-\cos xh}{x^2} d\mu(x)$, and hence determines μ except for possible atoms where $\cos xh = 1$, $x \neq 0$.
- By varying h, μ is determined.

□

Infinitely divisible distributions

Theorem

Let $\{\mu_n\}$ be a sequence of probability measures, with characteristic functions $\{\phi_n\}$, and let $\{c_n\}$, $\{\beta_n\}$ be sequences of real numbers. Set $b_n = \int_{-\infty}^{\infty} \sin x d\mu_n(x)$. A continuous limit

$$\psi_n(z) = c_n[\phi_n(z) - 1 - i\beta_n z] \to \rho(z)$$

exists if and only if there exists a canonical measure μ and a number b such that

$$c_n x^2 d\mu_n(x) \to d\mu(x)$$

properly and $c_n(b_n - \beta_n) \to b$. In this case

$$\rho(z) = \psi(z) + ibz$$

where $\psi(z) = \int_{-\infty}^{\infty} \frac{e^{izx} - 1 - iz\sin x}{x^2} d\mu(x)$. This uniquely determines μ.

Infinitely divisible distributions

Proof.

- The claim in the forward direction holds since $z(x) = e^{i\zeta x} - 1 - i\zeta \sin x$ is bounded as a function of x, and satisfies $\frac{z(x)}{x^2}$ is continuous at the origin.
- For the reverse direction, let $\psi_n(z) \to \rho(z)$ for all z, with ρ continuous.
- $e^{\psi_n(z)}$ is a characteristic function, and it converges uniformly to $e^{\rho(z)}$ in finite intervals.
- By the uniform convergence, there exists a canonical measure μ and a subsequence $\{n_1, n_2, ...\}$ such that $c_{n_k} x^2 d\mu_{n_k}(x) \to d\mu(x)$ properly.
- The proper convergence guarantees

$$\psi_{n_k}(z) = \int_{-\infty}^{\infty} \left[e^{izx} - 1 - iz \sin x\right] c_{n_k} d\mu_{n_k}(x) \to \psi(z).$$

□

Proof.

- Thus $\rho(z) = \psi(z) + ibz$.
- Since $\psi(1)$ is real, $b = \Im\rho(1)$.
- It follows that ψ and b are uniquely determined independent of the sequence $\{n_k\}$. This proves the required convergence.

$\square$

Infinitely divisible distributions

Lemma

Let $\{\phi_n\}$ be a sequence of characteristic functions. If the limit on the right is continuous, the relations

$$\phi_n^n(z) \to \omega(z), \qquad n[\phi_n(z) - 1] \to \rho(z)$$

are equivalent, and if either holds, $\omega(z) = e^{\rho(z)}$.

Infinitely divisible distributions

Proof.

- First assume $n[\phi_n(z) - 1] \to \rho(z)$. Thus $\phi_n(z) \to 1$ and the convergence is uniform in fixed intervals $|z| < z_0$.
- For n sufficiently large, $\log \phi_n(z)$ is well defined, and by Taylor expansion,
$$\log \phi_n^n(z) = n \log \phi_n(z) \to \rho(z),$$
so $\omega(z) = e^{\rho(z)}$.

□

Infinitely divisible distributions

Proof.

- Now suppose $\phi_n^n(z) \to \omega(z)$. Then $\omega(0) = 1$, so $\omega(z) \neq 0$ for $|z| \leqslant z_0$.
- Since the convergence is uniform, $\phi_n(z) \neq 0$ for $|z| < z_0$ for all n sufficiently large, and thus $\log \phi_n(z)$ is well defined in $|z| < z_0$. It follows that $\rho(z) = \log \omega(z)$ for $|z| < z_0$.
- After passing to a subsequence, we can find ρ such that for all z

$$n_k[\phi_{n_k}(z) - 1] \to \rho(z).$$

This implies $\phi_{n_k}^{n_k}(z) \to e^{\rho(z)}$, so $\rho = \log \omega$. But the limit now holds for the full sequence.

□

Infinitely divisible distributions

Theorem

For ω to be an infinitely divisible characteristic function it is necessary and sufficient that there exist a canonical measure μ and a real number b such that $\omega = e^{\rho}$ with

$$\rho(z) = \psi(z) + ibz$$

$$\psi(z) = \int_{-\infty}^{\infty} \frac{e^{izx} - 1 - iz\sin x}{x^2} d\mu(x).$$

Infinitely divisible distributions

Proof.

- First suppose ω is infinitely divisible with $\omega_n^n = \omega$. Then $n[\omega_n(z) - 1] \to \rho(z)$, which is the special case $c_n = n$, $\beta_n = 0$ of the previous theorem. The existence of canonical measure μ with ψ and b as defined there follows.
- Now suppose that $\omega = e^\rho$ is of the described form. First, suppose that the canonical measure μ is concentrated on $|x| > \delta$. Let

$$d\mu(x) = cx^2 d\nu(x).$$

  where ν is a probability distribution with characteristic function γ.
- We have $e^{c(\gamma(z)-1)}$ is the characteristic function of a distribution of compound Poisson type, and hence is infinitely divisible. It differs from e^ρ by the centering factor $e^{i\beta z}$, so that e^ρ is infinitely divisible.

$\square$

Infinitely divisible distributions

Proof.

- Now set, for $\delta > 0$, μ_δ the measure $\mu\mathbf{1}(|x| \geqslant \delta)$ and let $\psi_\delta(z)$ be the corresponding integral.
- Let $\sigma^2 \geqslant 0$ be the mass assigned by μ to 0. Hence, as $\delta \to 0$,

$$-\frac{1}{2}\sigma^2 z^2 + \psi_\delta(z) \to \psi(z).$$

- The left hand side is the logarithm of a characteristic function, hence so is the right. Since $\frac{\psi}{n}$ is obtained by replacing μ with $\frac{\mu}{n}$, the claim follows on setting $\omega_n = e^{\frac{\psi}{n}}$, $\omega = e^{\psi}$.

$\square$

Random walk

Definition

Let $X_1, X_2, \ldots$ be i.i.d. taking values in $\mathbb{R}^d$, and let $S_n = X_1 + \cdots + X_n$. S_n is a *random walk*.

In studying random walk we work on the product probability space $(\Omega, \mathscr{F}, \mathrm{Prob})$ from Kolmogorov's extension theorem,

$$\begin{aligned} \Omega &= \{(\omega_1, \omega_2, \ldots) : \omega_i \in \mathbb{R}^d\} \\ \mathscr{F} &= \mathscr{B} \times \mathscr{B} \times \ldots \\ \mathrm{Prob} &= \mu \times \mu \times \ldots, \qquad \mu \text{ is the distribution of } X_i \\ X_n(\omega) &= \omega_n. \end{aligned}$$

Permutable variables

Definition

A *finite permutation* of $\mathbb{N} = \{1, 2, \ldots\}$ is a map π from $\mathbb{N}$ to $\mathbb{N}$ so that $\pi(i) \neq i$ for only finitely many i. An event A is *permutable* if $\pi^{-1}A = \{\omega : \pi\omega \in A\} = A$ for all finite permutations π. The collection of permutable events is a σ-field, called the *exchangeable σ-field*, $\mathscr{E}$.

The tail σ-algebra is contained in $\mathscr{E}$, as are the events

- $\{\omega : S_n(\omega) \in B \text{ i.o.}\}$.
- $\{\omega : \limsup_{n\to\infty} S_n(\omega)/c_n \geqslant 1\}$.

Hewitt-Savage 0-1 Law

Theorem (Hewitt-Savage 0-1 Law)

If $X_1, X_2, \ldots$ are i.i.d. and $A \in \mathscr{E}$ then $\mathrm{Prob}(A) \in \{0, 1\}$.

Hewitt-Savage 0-1 Law

- Let $A \in \mathscr{E}$
- Let $A_n \in \sigma(X_1, ..., X_n)$ so that $\mathrm{Prob}(A_n \Delta A) \to 0$ as $n \to \infty$. Here $A \Delta B = (A - B) \cup (B - A)$ is the symmetric difference.
- Define $\pi = \pi_n$ by $\pi(j) = n + j$ if $j \leqslant n$, $\pi(j) = j - n$ if $n + 1 \leqslant j \leqslant 2n$ and $\pi(j) = j$ otherwise. Note π^2 is the identity.
- Write $A_n = B_n \times \mathbb{R}^{\mathbb{N}}$ where $B_n \subset \mathbb{R}^n$. We have

$$\mathrm{Prob}(A_n \Delta A) = \mathrm{Prob}(\omega : \pi\omega \in A_n \Delta A)$$

and

$$\{\omega : \pi\omega \in A_n\} = \{\omega : (\omega_{n+1}, ..., \omega_{2n}) \in B_n\}$$

Write A_n' for this event.

Hewitt-Savage 0-1 Law

- Use $\mathrm{Prob}(A_n \Delta A) = \mathrm{Prob}(A_n' \Delta A)$, so

$$\mathrm{Prob}(A_n \Delta A_n') \leqslant \mathrm{Prob}(A_n \Delta A) + \mathrm{Prob}(A_n' \Delta A) \to 0.$$

- The implies

$$\begin{aligned} 0 &\leqslant \mathrm{Prob}(A_n) - \mathrm{Prob}(A_n \cap A_n') \\ &\leqslant \mathrm{Prob}(A_n \cup A_n') - \mathrm{Prob}(A_n \cap A_n') = \mathrm{Prob}(A_n \Delta A_n') \to 0, \end{aligned}$$

  so $\mathrm{Prob}(A_n \cap A_n') \to \mathrm{Prob}(A)$.

- By independence

$$\mathrm{Prob}(A_n \cap A_n') = \mathrm{Prob}(A_n)\,\mathrm{Prob}(A_n') \to \mathrm{Prob}(A)^2.$$

  This shows $\mathrm{Prob}(A) = \mathrm{Prob}(A)^2$ so $\mathrm{Prob}(A) \in \{0, 1\}$.

Hewitt-Savage 0-1 Law

Theorem

For a random walk on $\mathbb{R}$, there are only four possibilities, of which one has probability 1:

- $S_n = 0$ *for all* n
- $\lim_{n\to\infty} S_n = \infty$
- $\lim_{n\to\infty} S_n = -\infty$
- $-\infty = \liminf S_n < \limsup S_n = \infty$.

Hewitt-Savage 0-1 Law

Proof.

- The Hewitt-Savage 0-1 Law implies that $\limsup S_n$ is a constant $c \in [-\infty, \infty]$.
- Let $S'_n = S_{n+1} - X_1$, which has the same distribution. Thus $c = c - X_1$, so that if c is finite, then $X_1 = 0$.
- The remaining cases are obvious.

$\square$

Stopping times

Definition

Let $\mathscr{F}_n = \sigma(X_1, ..., X_n)$. A random variable N taking values in $\{1, 2, ...\} \cup \{\infty\}$ is said to be a *stopping time* or an *optional random variable* if for every $n < \infty$, $\{N = n\} \in \mathscr{F}_n$. The σ-algebra generated by stopping time N is

$$\mathscr{F}_N = \{A \in \sigma(X_1, X_2, ...) : \forall n,\ A \cap \{N = n\} \in \mathscr{F}_n\}.$$

Given a set A, the *hitting time of* A is $N = \inf\{n : S_n \in A\}$. This is a stopping time.

Stopping times

Theorem

Let $X_1, X_2, \ldots$ be i.i.d., $\mathscr{F}_n = \sigma(X_1, \ldots, X_n)$ and let N be a stopping time with $\mathrm{Prob}(N < \infty) > 0$. Conditional on $\{N < \infty\}$, $\{X_{N+n}, n \geqslant 1\}$ is independent of $\mathscr{F}_N$ and has the same distribution as the original sequence.

Stopping times

Proof.

- It suffices to show that if $A \in \mathscr{F}_N$ and $B_j \in \mathscr{B}$ for $1 \leqslant j \leqslant k$, then

$$\begin{aligned}&\mathrm{Prob}(A, N < \infty, X_{N+j} \in B_j, 1 \leqslant j \leqslant k) \\ &= \mathrm{Prob}(A \cap \{N < \infty\}) \prod_{j=1}^{k} \mu(B_j).\end{aligned}$$

- For each fixed n

$$\begin{aligned}&\mathrm{Prob}(A, N = n, X_{N+j} \in B_j, 1 \leqslant j \leqslant k) \\ &= \mathrm{Prob}(A \cap \{N = n\}) \prod_{j=1}^{k} \mu(B_j).\end{aligned}$$

since $A \cap \{N = n\} \in \mathscr{F}_n$. This suffices.

$\square$

Shift

Definition

Let $\Omega = \mathbb{R}^{\mathbb{N}}$, and define the shift $\theta : \Omega \to \Omega$ by

$$(\theta\omega)(n) = \omega(n+1), \qquad n = 1, 2, \ldots.$$

Define, iteratively, $\theta^1 = \theta$ and, for $k > 1$, $\theta^k = \theta \circ \theta^{k-1}$. If N is a stopping time, define

$$\theta^N \omega = \begin{cases} \theta^n \omega & \{N = n\} \\ \Delta & \{N = \infty\} \end{cases}$$

where Δ is an extra point added to Ω.

Stopping times

Example

The stopping time

$$\tau(\omega) = \inf\{n > 0 : \omega_1 + \cdots + \omega_n = 0\}$$

is the time of the first return to 0. Set $\tau(\Delta) = \infty$. Define $\tau_1 = \tau$ and, for $n > 1$,

$$\tau_n(\omega) = \tau_{n-1}(\omega) + \tau(\theta^{\tau_{n-1}}(\omega)).$$

This records the time of the nth return to 0.

Stopping times

In general, the iterates of a stopping time T are defined by $T_0 = 0$ and

$$T_n(\omega) = T_{n-1}(\omega) + T(\theta^{T_{n-1}}\omega), \qquad n \geqslant 1.$$

One can check by induction that $\mathrm{Prob}(T_n < \infty) = \mathrm{Prob}(T < \infty)^n$.
Let $t_n = T(\theta^{T_{n-1}})$.

Theorem

Suppose $\mathrm{Prob}(T < \infty) = 1$. *The random vectors*

$$V_n = (t_n, X_{T_{n-1}+1}, ..., X_{T_n})$$

are independent and identically distributed.

This follows since $V_1, ..., V_{n-1} \in \mathscr{F}(T_{n-1})$.

Wald's equation

Theorem (Wald's equation)

Let $X_1, X_2, \ldots$ be i.i.d. with $\mathrm{E}[|X_i|] < \infty$. If N is a stopping time with $\mathrm{E}[N] < \infty$, then $\mathrm{E}[S_N] = \mathrm{E}[X_1]\,\mathrm{E}[N]$.

Wald's equation

Proof.

First suppose the $X_i \geqslant 0$.

$$\mathrm{E}[S_N] = \int S_N dP = \sum_{n=1}^{\infty} \int S_n \mathbf{1}_{(N=n)} dP = \sum_{n=1}^{\infty} \sum_{m=1}^{n} \int X_m \mathbf{1}_{(N=n)} dP.$$

By Fubini

$$= \sum_{m=1}^{\infty} \sum_{n=m}^{\infty} \int X_m \mathbf{1}_{(N=n)} dP = \sum_{m=1}^{\infty} \int X_m \mathbf{1}_{(N \geqslant m)} dP.$$

Since $\{N \geqslant m\} = \{N \leqslant m-1\}^c \in \mathscr{F}_{m-1}$ it is independent of X_m, the last expression is

$$\sum_{m=1}^{\infty} \mathrm{E}[X_m] \, \mathrm{Prob}(N \geqslant m) = \mathrm{E}[X_1] \, \mathrm{E}[N].$$

$\square$

Wald's equation

Proof.

To handle the general case, use

$$\infty > \sum_{m=1}^{\infty} \mathrm{E}[|X_m|] \, \mathrm{Prob}(N \geqslant m) = \sum_{m=1}^{\infty} \sum_{n=m}^{\infty} \int |X_m| \mathbf{1}_{(N=n)} dP,$$

which justifies the application of Fubini. $\square$

Example

- Let $X_1, X_2, \ldots$ be i.i.d. with $\text{Prob}(X_i = 1) = \text{Prob}(X_i = -1) = \frac{1}{2}$.
- Let $a < 0 < b$ be integers and let $N = \inf\{n : S_n \notin (a, b)\}$.
- Observe that if $x \in (a, b)$, $\text{Prob}(x + S_{b-a} \notin (a, b)) \geqslant 2^{-(b-a)}$, since $b - a$ steps right land outside the interval. Hence

$$\text{Prob}(N > n(b-a)) \leqslant \left(1 - 2^{-(b-a)}\right)^n \qquad \Rightarrow \qquad \text{E}[N] < \infty.$$

- By the previous theorem, $\text{E}[S_N] = 0$, so $b \,\text{Prob}(S_N = b) + a \,\text{Prob}(S_N = a) = 0$ and

$$\text{Prob}(S_N = b) = \frac{-a}{b-a}, \qquad \text{Prob}(S_N = a) = \frac{b}{b-a}.$$

Wald's second equation

Theorem

Let $X_1, X_2, \ldots$ *be i.i.d. with* $\mathrm{E}[X_n] = 0$ *and* $\mathrm{E}[X_n^2] = \sigma^2 < \infty$. *If* T *is a stopping time with* $\mathrm{E}[T] < \infty$, *then* $\mathrm{E}[S_T^2] = \sigma^2 \,\mathrm{E}[T]$.

Wald's second equation

Proof.

Since $\mathrm{E}[X_n] = 0$ and X_n is independent of S_{n-1} and $\mathbf{1}_{(T \geqslant n)} \in \mathscr{F}_{n-1}$,

$$S_{T\wedge n}^2 = S_{T\wedge(n-1)}^2 + (2X_n S_{n-1} + X_n^2)\mathbf{1}_{(T\geqslant n)}$$
$$\mathrm{E}[S_{T\wedge n}^2] = \mathrm{E}[S_{T\wedge(n-1)}^2] + \sigma^2 \,\mathrm{Prob}(T \geqslant n).$$

Thus

$$\mathrm{E}[S_{T\wedge n}^2] = \sigma^2 \sum_{m=1}^{n} \mathrm{Prob}(T \geqslant m)$$
$$\mathrm{E}[(S_{T\wedge n} - S_{T\wedge m})^2] = \sigma^2 \sum_{k=m+1}^{n} \mathrm{Prob}(T \geqslant k).$$

This shows that $S_{T\wedge n}$ is a Cauchy sequence in L^2, so the equality is obtained by letting $n \to \infty$.

□

Small deviations

Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $\mathrm{E}[X_n] = 0$ and $\mathrm{E}[X_n^2] = 1$, and let $T_c = \inf\{n \geqslant 1 : |S_n| > cn^{\frac{1}{2}}\}$.

$$\mathrm{E}[T_c] \begin{cases} < \infty & c < 1 \\ = \infty & c \geqslant 1 \end{cases}.$$

Small deviations

Proof.

- If $\mathrm{E}[T_c] < \infty$ then $\mathrm{E}[T_c] = \mathrm{E}[S_{T_c}^2] > c^2\,\mathrm{E}[T_c]$, a contradiction if $c \geqslant 1$.
- Now suppose $c < 1$ and let $\tau = T_c \wedge n$ and observe $S_{\tau-1}^2 \leqslant c^2(\tau - 1)$, so by Cauchy-Schwarz

$$\begin{aligned}\mathrm{E}[\tau] = \mathrm{E}[S_\tau^2] &= \mathrm{E}[S_{\tau-1}^2 + 2\,\mathrm{E}[S_{\tau-1}X_\tau] + \mathrm{E}[X_\tau^2] \\ &\leqslant c^2\,\mathrm{E}[\tau] + 2c\left(\mathrm{E}[\tau]\,\mathrm{E}[X_\tau^2]\right)^{\frac{1}{2}} + \mathrm{E}[X_\tau^2].\end{aligned}$$

- The proof is completed by the following lemma.

$\square$

Small deviations

Lemma

If T is a stopping time with $\mathrm{E}[T] = \infty$, then

$$\mathrm{E}[X_{T\wedge n}^2]/E[T \wedge n] \to 0$$

as $n \to \infty$.

This suffices to show $\mathrm{E}[T_c] < \infty$, since otherwise, with $0 < \epsilon < 1 - c^2$ and n large one obtains $\mathrm{E}[\tau] \leqslant (c^2 + \epsilon)\,\mathrm{E}[\tau]$.

Small deviations

Proof.

- Write

$$\mathrm{E}[X_{T\wedge n}^2] = \mathrm{E}[X_{T\wedge n}^2 \mathbf{1}(X_{T\wedge n}^2 \leqslant \epsilon(T \wedge n))] + \sum_{j=1}^{n} \mathrm{E}[X_j^2 \mathbf{1}(T \wedge n = j, X_j^2 > \epsilon j)].$$

- Bound the first term by $\leqslant \epsilon\, \mathrm{E}[T \wedge n]$.
- To bound the second, choose $N \geqslant 1$ so that for $n \geqslant N$,

$$\sum_{j=1}^{n} \mathrm{E}[X_j^2 \mathbf{1}(X_j^2 > \epsilon j)] < n\epsilon,$$

which is possible since $\mathrm{E}[X_j^2] < \infty$.

$\square$

Small deviations

Proof.

- Bound
$$\sum_{j=1}^{N} \mathrm{E}[X_j^2 \mathbf{1}(T \wedge n, X_j^2 > \epsilon j)] \leqslant N\,\mathrm{E}[X_1^2],$$
and
$$\sum_{j=N}^{n} \mathrm{E}[X_j^2 \mathbf{1}(T \wedge n, X_j^2 > \epsilon j)] \leqslant \sum_{j=N}^{n} \mathrm{E}[X_j^2 \mathbf{1}(T \wedge n \geqslant j, X_j^2 > \epsilon j)]$$
$$= \sum_{j=N}^{n} \mathrm{Prob}(T \wedge n \geqslant j)\,\mathrm{E}[X_j^2 \mathbf{1}(X_j^2 > \epsilon j)]$$
$$= \sum_{j=N}^{n} \sum_{k=j}^{\infty} \mathrm{Prob}(T \wedge n = k)\,\mathrm{E}[X_j^2 \mathbf{1}(X_j^2 > \epsilon j)]$$

$\square$

Small deviations

Proof.

- Bound the last sum by

$$\leqslant \sum_{k=N}^{\infty} \sum_{j=1}^{k} \mathrm{Prob}(T \wedge n = k)\, \mathrm{E}[X_j^2 \mathbf{1}(X_j^2 > \epsilon j)]$$

$$\leqslant \epsilon \sum_{k=N}^{\infty} k\, \mathrm{Prob}(T \wedge n = k) \leqslant \epsilon\, \mathrm{E}(T \wedge n).$$

- We've checked

$$\mathrm{E}[X_{T \wedge n}^2] \leqslant 2\epsilon\, \mathrm{E}[T \wedge n] + N\, \mathrm{E}[X_1^2].$$

Since $\mathrm{E}[T \wedge n] \to \infty$, the conclusion follows. □