Math 639: Lecture 10

Intro to Martingales

Bob Hough

February 28, 2017

Conditional expectation

Recall the definition of conditional expectation.

Definition

Given a probability space $(\Omega, \mathscr{F}_0, \text{Prob})$, a σ-field $\mathscr{F} \subset \mathscr{F}_0$, and a random variable $X \in \mathscr{F}_0$ with $\mathrm{E}[|X|] < \infty$, the *conditional expectation of* X *given* $\mathscr{F}$, $\mathrm{E}[X|\mathscr{F}]$ is a $\mathscr{F}$-measurable random variable such that, for all $A \in \mathscr{F}$,

$$\int_A X dP = \int_A Y dP.$$

Conditional expectation

Example

Suppose X is independent of $\mathscr{F}$, that is, for all $B \in \mathscr{B}$ and $A \in \mathscr{F}$,

$$\mathrm{Prob}(\{X \in B\} \cap A) = \mathrm{Prob}(X \in B)\,\mathrm{Prob}(A).$$

Then $\mathrm{E}[X|\mathscr{F}] = \mathrm{E}[X]$, since if $A \in \mathscr{F}$,

$$\int_A X dP = \mathrm{E}[X\mathbf{1}_A] = \mathrm{E}[X]\,\mathrm{E}[1_A] = \int_A \mathrm{E}[X] dP.$$

Conditional expectation

Example

- Suppose X and Y are independent. Let ϕ be a function with $\mathrm{E}[|\phi(X,Y)|] < \infty$ and let $g(x) = \mathrm{E}[\phi(x,Y)]$. Then $\mathrm{E}[\phi(X,Y)|X] = g(X)$.
- To check this, let $A \in \sigma(X)$, then $A = \{X \in C\}$ for a measureable set C, and

$$\begin{aligned}\int_A \phi(X,Y)dP &= \mathrm{E}[\phi(X,Y)\mathbf{1}_C(X)] \\ &= \int\int \phi(x,y)\mathbf{1}_C(x)\nu(dy)\mu(dx) \\ &= \int \mathbf{1}_C(x)g(x)\mu(dx) = \int_A g(X)dP.\end{aligned}$$

Properties of conditional expectation

Theorem

Conditional expectation satisfies the following properties.

1. *Linearity*
$$\mathrm{E}[aX + Y|\mathscr{F}] = a\,\mathrm{E}[X|\mathscr{F}] + \mathrm{E}[Y|\mathscr{F}].$$
2. *If $X \leqslant Y$ then*
$$\mathrm{E}[X|\mathscr{F}] \leqslant \mathrm{E}[Y|\mathscr{F}]$$
3. *If $X_n \geqslant 0$ and $X_n \uparrow X$ with $\mathrm{E}[X] < \infty$, then*
$$\mathrm{E}[X_n|\mathscr{F}] \uparrow \mathrm{E}[X|\mathscr{F}].$$

Properties of conditional expectation

Proof.

For the first item, let $A \in \mathscr{F}$ and write

$$\begin{aligned}\int_A a\,\mathrm{E}[X|\mathscr{F}] + \mathrm{E}[Y|\mathscr{F}]dP &= a\int_A \mathrm{E}[X|\mathscr{F}]dP + \int_A \mathrm{E}[Y|\mathscr{F}]dP \\ &= a\int_A XdP + \int_A YdP = \int_A (aX+Y)dP.\end{aligned}$$

For the second item,

$$\int_A \mathrm{E}[X|\mathscr{F}]dP = \int_A XdP \leqslant \int_A YdP = \int_A \mathrm{E}[Y|\mathscr{F}]dP.$$

Let $A = \{\mathrm{E}[X|\mathscr{F}] - \mathrm{E}[Y|\mathscr{F}] > \epsilon\}$ to get the claim. $\square$

Properties of conditional expectation

Proof.

Let $Y_n = X - X_n$. Since Y_n decreases, $Z_n = \mathrm{E}[Y_n|\mathscr{F}]$ decreases to a limit Z_∞. For $A \in \mathscr{F}$,

$$\int_A Z_n dP = \int_A Y_n dP.$$

Since $Y_n \downarrow 0$, dominated convergence gives $\int_A Z_\infty dP = 0$ for all A, so $Z_\infty = 0$. □

Properties of conditional expectation

Theorem

If ϕ is convex and $\mathrm{E}[|X|], \mathrm{E}[|\phi(X)|] < \infty$, then

$$\phi(\mathrm{E}[X|\mathscr{F}]) \leqslant \mathrm{E}[\phi(X)|\mathscr{F}].$$

Properties of conditional expectation

Proof.

Let $S = \{(a, b) : a, b \in \mathbb{Q}, ax + b \leqslant \phi(x)\}$. Then

$$\phi(x) = \sup\{ax + b : (a, b) \in S\}.$$

For all $a, b \in S$,

$$\mathrm{E}[\phi(X)|\mathscr{F}] \geqslant a\,\mathrm{E}[X|\mathscr{F}] + b$$

so $\mathrm{E}[\phi(X)|\mathscr{F}] \geqslant \phi(\mathrm{E}[X|\mathscr{F}])$. $\square$

Properties of conditional expectation

Theorem

Conditional expectation is a contraction in L^p, $p \geqslant 1$.

Properties of conditional expectation

Proof.

By convexity, $|\operatorname{E}[X|\mathscr{F}]|^p \leqslant \operatorname{E}[|X|^p|\mathscr{F}]$. Hence, taking expectation,

$$\operatorname{E}[|\operatorname{E}[X|\mathscr{F}]|^p] \leqslant \operatorname{E}[\operatorname{E}[|X|^p|\mathscr{F}]] = \operatorname{E}[|X|^p].$$

$\square$

Properties of conditional expectation

Theorem

If $\mathscr{F} \subset \mathscr{G}$ and $\mathrm{E}[X|\mathscr{G}] \in \mathscr{F}$, then $\mathrm{E}[X|\mathscr{F}] = \mathrm{E}[X|\mathscr{G}]$.

Proof.

If $A \in \mathscr{F} \subset \mathscr{G}$, then

$$\int_A X dP = \int_A \mathrm{E}[X|\mathscr{G}] dP.$$

$\square$

Properties of conditional expectation

Theorem

If $\mathscr{F}_1 \subset \mathscr{F}_2$ *then*

1. $\mathrm{E}[\mathrm{E}[X|\mathscr{F}_1]|\mathscr{F}_2] = \mathrm{E}[X|\mathscr{F}_1]$
2. $\mathrm{E}[\mathrm{E}[X|\mathscr{F}_2]|\mathscr{F}_1] = \mathrm{E}[X|\mathscr{F}_1]$.

Properties of conditional expectation

Proof.

The first item follows because $\mathrm{E}[X|\mathscr{F}_1]$ is $\mathscr{F}_2$-measurable. To prove the second item, note that both sides are $\mathscr{F}_1$ measurable. Given $A \in \mathscr{F}_1 \subset \mathscr{F}_2$,

$$\int_A \mathrm{E}[X|\mathscr{F}_1]dP = \int_A XdP = \int_A \mathrm{E}[X|\mathscr{F}_2]dP.$$

$\square$

Properties of conditional expectation

Theorem

If $X \in \mathscr{F}$ *and* $\mathrm{E}[|Y|], \mathrm{E}[|XY|] < \infty$, *then*

$$\mathrm{E}[XY|\mathscr{F}] = X\,\mathrm{E}[Y|\mathscr{F}].$$

Properties of conditional expectation

Proof.

First let $B \in \mathscr{F}$ and let $X = \mathbf{1}_B$ with $B \in \mathscr{F}$. For $A \in \mathscr{F}$,

$$\int_A \mathbf{1}_B \, \mathrm{E}[Y|\mathscr{F}] dP = \int_{A\cap B} \mathrm{E}[Y|\mathscr{F}] dP = \int_{A \cap B} Y dP = \int_A \mathbf{1}_B Y dP.$$

The same holds for simple X by linearity, then for positive variables by monotone convergence, and finally in general by splitting into positive and negative parts. $\square$

Theorem

Suppose $\mathrm{E}[X^2] < \infty$. $\mathrm{E}[X|\mathscr{F}]$ *is the variable* $Y \in \mathscr{F}$ *that minimizes the mean square error* $\mathrm{E}[(X - Y)^2]$.

Properties of conditional expectation

Proof.

For $Z \in L^2(\mathscr{F})$,

$$Z\,\mathrm{E}[X|\mathscr{F}] = \mathrm{E}[ZX|\mathscr{F}].$$

Hence

$$\mathrm{E}[Z\,\mathrm{E}[X|\mathscr{F}]] = \mathrm{E}[\mathrm{E}[ZX|\mathscr{F}]] = \mathrm{E}[ZX],$$

or

$$\mathrm{E}[Z(X - \mathrm{E}[X|\mathscr{F}])] = 0, \qquad \forall Z \in L^2(\mathscr{F}).$$

If $Y \in L^2(\mathscr{F})$ and $Z = \mathrm{E}[X|\mathscr{F}] - Y$, then

$$\mathrm{E}[(X-Y)^2] = \mathrm{E}[(X - \mathrm{E}[X|\mathscr{F}])^2] + \mathrm{E}[Z^2].$$

□

Martingales

Definition

A *filtration* is an increasing sequence of σ-algebras $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots$. A sequence $\{X_n\}$ is said to be *adapted* to $\mathscr{F}_n$ if $X_n \in \mathscr{F}_n$ for all n. If $\{X_n\}$ satisfies

- $\mathrm{E}[|X_n|] < \infty$ for all n
- $\mathrm{E}[X_{n+1}|\mathscr{F}_n] = X_n$ for all n

then X is a *martingale* with respect to $\mathscr{F}_n$. If instead $\mathrm{E}[X_{n+1}|\mathscr{F}_n] \leqslant X_n$ then X is a *supermartingale*. If instead $\mathrm{E}[X_{n+1}|\mathscr{F}_n] \geqslant X_n$ then X is a *submartingale*.

Simple random walk

Example

Let $\xi_1, \xi_2, ...$ be i.i.d. ± 1 with equal probability, and let $X_n = \xi_1 + \cdots + \xi_n$. Set $\mathscr{F}_n = \sigma(\xi_1, ..., \xi_n)$. Then

$$\mathrm{E}[X_{n+1}|\mathscr{F}_n] = \mathrm{E}[X_n|\mathscr{F}_n] + \mathrm{E}[\xi_{n+1}|\mathscr{F}_n] = X_n + \mathrm{E}[\xi_{n+1}] = X_n.$$

Superharmonic functions

Example

The name supermartingale comes from the fact that a superharmonic function, which satisfies $\Delta f \leqslant 0$, has

$$f(x) \geqslant \frac{1}{|B(0,r)|}\int_{B(x,r)} f(y)dy.$$

Theorem

If X_n is a supermartingale then for $n > m$, $\mathrm{E}[X_n|\mathscr{F}_m] \leqslant X_m$.

Proof.

This holds for $n = m + 1$ by definition. For $n = m + k$,

$$\mathrm{E}[X_{m+k}|\mathscr{F}_m] = \mathrm{E}[\mathrm{E}[X_{m+k}|\mathscr{F}_{m+k-1}]|\mathscr{F}_m] \leqslant \mathrm{E}[X_{m+k-1}|\mathscr{F}_m].$$

The claim in general now follows by induction. □

Supermartingales and submartingales

Theorem

If X_n is a submartingale, then for $n > m$, $\mathrm{E}[X_n|\mathscr{F}_m] \geqslant X_m$. If X_n is a martingale then for $n > m$, $\mathrm{E}[X_n|\mathscr{F}_m] = X_m$.

Proof.

If X_n is a submartingale, then $-X_n$ is a supermartingale, from which the first claim follows. The second follows since a martingale is both a submartingale and a supermartingale. □

Supermartingales and submartingales

Theorem

If X_n is a martingale with respect to filtration $\mathscr{F}_n$ and ϕ is a convex function with $\mathrm{E}[|\phi(X_n)|] < \infty$ for all n, then $\phi(X_n)$ is a submartingale with respect to $\mathscr{F}_n$. In particular, if $p \geqslant 1$ and $\mathrm{E}[|X_n|^p] < \infty$ for all n, then $|X_n|^p$ is a submartingale with respect to $\mathscr{F}_n$.

Proof.

By Jensen's inequality,

$$\mathrm{E}[\phi(X_{n+1})|\mathscr{F}_n] \geqslant \phi(\mathrm{E}[X_{n+1}|\mathscr{F}_n]) = \phi(X_n).$$

$\square$

Supermartingales and submartingales

Theorem

If X_n is a submartingale with respect to $\mathscr{F}_n$ and ϕ is an increasing convex function with $\mathrm{E}[|\phi(X_n)|] < \infty$ for all n, then $\phi(X_n)$ is a submartingale with respect to $\mathscr{F}_n$. Consequently

1. *If X_n is a submartingale, then $(X_n - a)^+$ is a submartingale.*
2. *If X_n is a supermartingale, then $\min(X_n, a)$ is a supermartingale.*

Proof.

By Jensen's inequality, and the fact that ϕ is increasing,

$$\mathrm{E}[\phi(X_{n+1})|\mathscr{F}_n] \geqslant \phi(\mathrm{E}[X_{n+1}|\mathscr{F}_n]) \geqslant \phi(X_n).$$

□

Predictable sequences

Definition

Let $\mathscr{F}_n$, $n \geqslant 0$ be a filtration. H_n, $n \geqslant 1$ is a *predictable sequence* if $H_n \in \mathscr{F}_{n-1}$ for all $n \geqslant 1$. The *martingale transform* of H_n with respect to the sequence of sub or super martingales $(X_n, \mathscr{F}_n)$ is

$$Y_0 = 0, \qquad Y_n = \sum_{k=1}^{n} H_k(X_k - X_{k-1}), \; n \geqslant 1.$$

Predictable sequences

Theorem

Suppose $\{Y_n\}$ is the martingale transform of $\mathscr{F}_n$-predictable $\{V_n\}$ with respect to a sub or super martingale $(X_n, \mathscr{F}_n)$.

- *If Y_n is integrable and $(X_n, \mathscr{F}_n)$ is a martingale, then $(Y_n, \mathscr{F}_n)$ is also a martingale.*
- *If Y_n is integrable, $V_n \geqslant 0$ and $(X_n, \mathscr{F}_n)$ is a sub or super martingale, then Y_n is a sub or super martingale.*

Proof.

Check

$$\mathrm{E}[Y_{n+1} - Y_n|\mathscr{F}_n] = E[V_{n+1}(X_{n+1} - X_n)|\mathscr{F}_n] = V_{n+1}\,\mathrm{E}[X_{n+1} - X_n|\mathscr{F}_n],$$

from which the claims follows.

□

Stopping times

Theorem

If $(X_n, \mathscr{F}_n)$ is a sub-martingale, or sup-martingale and $\theta \leqslant \tau$ are stopping times for $\{\mathscr{F}_n\}$ then $(X_{n\wedge\tau} - X_{n\wedge\theta}, \mathscr{F}_n)$ is also a sub or sup-martingale. In particular, taking $\theta = 0$, $(X_{n\wedge\tau}, \mathscr{F}_n)$ is a sub or sup-martingale.

Proof.

- Suppose X_n is a sub-martingale, otherwise replace it with $-X_n$.
- Let $V_k = \mathbf{1}(\theta < k \leqslant \tau)$. Thus V_k is $\mathscr{F}_{k-1}$-measurable.
- Since

$$X_{n\wedge\tau} - X_{n\wedge\theta} = \sum_{k=1}^{n} V_k(X_k - X_{k-1})$$

is a martingale transform, it is again a sub-martingale.

□

Upcrossing inequality

Example (Upcrossings)

Let $a < b$ and let $N_0 = -1$. For $k \geqslant 1$,

$$\begin{aligned} N_{2k-1} &= \inf\{m > N_{2k-2} : X_m \leqslant a\} \\ N_{2k} &= \inf\{m > N_{2k-1} : X_m \geqslant b\}. \end{aligned}$$

The N_j are stopping times, and

$$H_m = \begin{cases} 1 & N_{2k-1} < m \leqslant N_{2k}, \text{ some } k \\ 0 & \text{otherwise} \end{cases}$$

is a predictable sequence.

Upcrossing inequality

Define $U_n = \sup\{k : N_{2k} \leqslant n\}$ to be the number of upcrossings to time n.

Theorem (Upcrossing inequality)

If X_m, $m \geqslant 0$, is a submartingale, then

$$(b-a)\,\mathrm{E}[U_n] \leqslant \mathrm{E}[(X_n-a)^+] - \mathrm{E}[(X_0-a)^+].$$

Upcrossing inequality

Proof.

- Let $Y_m = a + (X_m - a)^+$.
- Y_m is a submartingale, and it upcrosses $[a, b]$ the same number of times that X_m does.
- One has $(H \cdot Y)_n \geqslant (b - a)U_n$.
- Set $K = 1 - H$, and note that $\mathrm{E}[K \cdot Y_n] \geqslant \mathrm{E}[K \cdot Y_0] = 0$. Hence $\mathrm{E}[H \cdot Y_n] \leqslant \mathrm{E}[Y_n - Y_0]$.

$\square$

Martingale convergence theorem

Theorem (Martingale convergence theorem)

If X_n is a submartingale with $\sup \mathrm{E}[X_n^+] < \infty$, then as $n \to \infty$, X_n converges a.s. to a limit X, $\mathrm{E}[|X|] < \infty$.

Martingale convergence theorem

Proof.

- Since $(X-a)^+ \leqslant X^+ + |a|$,

$$\mathrm{E}[U_n] \leqslant \frac{|a| + \mathrm{E}[X_n^+]}{b-a}.$$

- As $n \uparrow \infty$, $U_n \uparrow U$ the number of upcrossings of $[a, b]$ by the whole sequence.
- If $\sup \mathrm{E}[X_n^+] < \infty$ then $\mathrm{E}[U] < \infty$, so $U < \infty$ a.s., so for all rational a, b,

$$\bigcup_{a,b \in \mathbb{Q}} \{\liminf X_n < a < b < \limsup X_n\}$$

has probability 0. Hence $\lim X_n$ exists with probability 1.

$\square$

Martingale convergence theorem

Proof.

- We have $\mathrm{E}[X^+] \leqslant \liminf \mathrm{E}[X_n^+] < \infty$.
- Also, $\mathrm{E}[X_n^-] = \mathrm{E}[X_n^+] - \mathrm{E}[X_n] \leqslant \mathrm{E}[X_n^+] - \mathrm{E}[X_0]$
- $\mathrm{E}[X^-] \leqslant \liminf_{n\to\infty} \mathrm{E}[X_n^-] \leqslant \sup_n \mathrm{E}[X_n^+] - \mathrm{E}[X_0] < \infty$.

$\square$

Supermartingale version

Theorem

If $X_n \geqslant 0$ is a supermartingale, then as $n \to \infty$, $X_n \to X$ a.s. and $\mathrm{E}[X] \leqslant \mathrm{E}[X_0]$.

Proof.

$-X_n \leqslant 0$ is a submartingale. $\square$

Examples

Example

- Let $S_0 = 1$, $S_n = 1 + \xi_1 + \cdots + \xi_n$ be simple random walk.
- Let $N = \inf\{n : S_n = 0\}$ and $X_n = S_{N \wedge n}$.
- X_n is a non-negative martingale, which converges a.s. to a finite limit, which is zero.
- Since $\mathrm{E}[X_n] = \mathrm{E}[X_0] = 1$ for all n, the convergence is not in L^1.

Doob's decomposition

Theorem (Doob's decomposition)

Any submartingale X_n, $n \geqslant 0$, can be written in a unique way as $X_n = M_n + A_n$, where M_n is a martingale and A_n is a predictable increasing sequence with $A_0 = 0$.

Doob's decomposition

Proof.

- Let $A_0 = 0$ and for $n \geqslant 1$,

$$A_n = A_{n-1} + \mathrm{E}[X_n - X_{n-1} | \mathscr{F}_{n-1}].$$

  By construction, $\{A_n\}$ is $\mathscr{F}_{n-1}$-measurable.

- To check that $Y_n = X_n - A_n$ is a martingale, calculate

$$\begin{aligned} \mathrm{E}[Y_n - Y_{n-1} | \mathscr{F}_{n-1}] &= \mathrm{E}[X_n - X_{n-1} - (A_n - A_{n-1}) | \mathscr{F}_{n-1}] \\ &= \mathrm{E}[X_n - X_{n-1} | \mathscr{F}_{n-1}] - (A_n - A_{n-1}) = 0. \end{aligned}$$

□

Bounded increments

Theorem

Let $X_1, X_2, \ldots$ *be a martingale with* $|X_{n+1} - X_n| \leqslant M < \infty$. *Let*

$$C = \{\lim X_n \text{ exists and is finite}\}$$
$$D = \{\limsup X_n = \infty, \; \liminf X_n = -\infty\}.$$

Then $\mathrm{Prob}(C \cup D) = 1$.

Bounded increments

Proof.

- We can assume $X_0 = 0$ by replacing X_n with $X_n - X_0$.
- Let $N = \inf\{n : X_n \leqslant -K\}$. Then $X_{n \wedge N}$ is bounded below, so converges, and hence X_n converges on $\{N = \infty\}$.
- Letting $K \to \infty$ the limit exists on $\{\liminf X_n > -\infty\}$. Replacing X_n with $-X_n$, the claim follows.

$\square$

Borel-Cantelli revisited

Theorem (Second Borel-Cantelli lemma)

Let $\mathscr{F}_n$, $n \geqslant 0$ be a filtration with $\mathscr{F}_0 = \{\varnothing, \Omega\}$ and A_n, $n \geqslant 1$ a sequence of events with $A_n \in \mathscr{F}_n$. Then

$$\{A_n \text{ i.o.}\} = \left\{ \sum_{n=1}^{\infty} \mathrm{Prob}(A_n | \mathscr{F}_{n-1}) = \infty \right\}.$$

Borel-Cantelli revisited

Proof.

- Let $X_0 = 0$ and $X_n = \sum_{m=1}^n (\mathbf{1}_{A_m} - \text{Prob}(A_m|\mathscr{F}_{m-1}))$ for $n \geqslant 1$. Thus $|X_n - X_{n-1}| \leqslant 1$.
- Using the decomposition $C \cup D$ of the previous theorem, on C where the limit exists,

$$\sum_{n=1}^{\infty} \mathbf{1}_{A_n} = \infty \Leftrightarrow \sum_{n=1}^{\infty} \text{Prob}(A_n|\mathscr{F}_{n-1}) = \infty.$$

On D, where the lim sup is ∞ and the lim inf is $-\infty$

$$\sum_{n=1}^{\infty} \mathbf{1}_{A_n} = \infty \text{ and } \sum_{n=1}^{\infty} \text{Prob}(A_n|\mathscr{F}_{n-1}) = \infty.$$

□

Radon-Nikodym derivatives

Lemma

Let μ be a finite measure and ν a probability measure on $(\Omega, \mathscr{F})$. Let $\mathscr{F}_n \uparrow \mathscr{F}$ be σ-algebras. Let μ_n and ν_n be the restrictions of μ and ν to $\mathscr{F}_n$. Suppose $\mu_n \ll \nu_n$ for all n, and let $X_n = \frac{d\mu_n}{d\nu_n}$ is a martingale with respect to $\mathscr{F}_n$.

Radon-Nikodym derivatives

Proof.

- Let $A \in \mathscr{F}_n$. Calculate
$$\int_A X_n d\nu = \int_A X_n d\nu_n = \mu_n(A) = \mu(A).$$
- Hence if $A \in \mathscr{F}_{m-1}$
$$\int_A X_m d\nu = \mu(A) = \int_A X_{m-1} d\nu$$
so $\mathrm{E}[X_m|\mathscr{F}_{m-1}] = X_{m-1}$.

□

Radon-Nikodym derivatives

Theorem

With the set-up as in the previous lemma, let $X = \limsup X_n$. *Then*

$$\mu(A) = \int_A X d\nu + \mu(A \cap \{X = \infty\}).$$

For a proof, see Durrett pp. 242–243.

Branching processes

Definition

Let ξ_i^n, $i, n \geqslant 1$ be i.i.d. nonnegative integer-valued random variables. The *Galton-Watson process* is a sequence Z_n, $n \geqslant 0$ by $Z_0 = 1$ and

$$Z_{n+1} = \begin{cases} \xi_1^{n+1} + \cdots + \xi_{Z_n}^{n+1} & Z_n > 0 \\ 0 & Z_n = 0 \end{cases}.$$

$p_k = \mathrm{Prob}(\xi_i^n = k)$ is called the *offspring distribution*.

Branching processes

Lemma

Let $\mathscr{F}_n = \sigma(\xi_i^m : i \geqslant 1, 1 \leqslant m \leqslant n)$ and $\mu = \mathrm{E}[\xi_i^m] \in (0, \infty)$. Then $\frac{Z_n}{\mu^n}$ is a martingale with respect to $\mathscr{F}_n$.

Branching processes

Proof.

Calculate

$$\begin{aligned}
\mathrm{E}[Z_{n+1}|\mathscr{F}_n] &= \sum_{k=1}^{\infty} \mathrm{E}[Z_{n+1}\mathbf{1}(Z_n = k)|\mathscr{F}_n] \\
&= \sum_{k=1}^{\infty} \mathrm{E}[(\xi_1^{n+1} + \cdots + \xi_k^{n+1})\mathbf{1}(Z_n = k)|\mathscr{F}_n] \\
&= \sum_{k=1}^{\infty} \mathbf{1}(Z_n = k)\, \mathrm{E}[\xi_1^{n+1} + \cdots + \xi_k^{n+1}|\mathscr{F}_n] \\
&= \mu \sum_{k=1}^{\infty} \mathbf{1}(Z_n = k)k = \mu Z_n.
\end{aligned}$$

□

Branching processes

Theorem

If $\mu < 1$ then $Z_n = 0$ for all n sufficiently large, so $\frac{Z_n}{\mu^n} \to 0$.

Proof.

$\mathrm{E}\left[\frac{Z_n}{\mu^n}\right] = \mathrm{E}[Z_0] = 1$, so $\mathrm{E}[Z_n] = \mu^n$. Since $Z_n \geqslant 1$ when $Z_n \neq 0$, $\mathrm{Prob}(Z_n \neq 0) \leqslant \mu^n \to 0$. $\square$

Branching processes

Theorem

If $\mu = 1$ *and* $\mathrm{Prob}(\xi_i^m = 1) < 1$ *then* $Z_n = 0$ *for all* n *sufficiently large.*

Branching processes

Proof.

- When $\mu = 1$, Z_n is a non-negative martingale.
- Z_n has an almost sure finite limit Z_∞, and since Z_n is integer valued, $Z_n = Z_\infty$ for all n sufficiently large.
- Since $\text{Prob}(\xi_i^m = 1) < 1$, the only possibility is $Z_\infty = 0$.

$\square$

For $s \in [0,1]$, let $\phi(s) = \sum_{k=0}^{\infty} p_k s^k$ where $p_k = \mathrm{Prob}(\xi_i^m = k)$.

Theorem

If $\mu = \mathrm{E}[\xi_i^m] > 1$ then $\mathrm{Prob}(Z_n = 0$ for some $n) = \rho$, the unique fixed point of ϕ in $[0,1)$.

Proof.

Calculate

$$\phi'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1} \geqslant 0$$

$$\phi''(s) = \sum_{k=2}^{\infty} k(k-1) p_k s^{k-2} \geqslant 0.$$

Thus ϕ is increasing and convex and $\lim_{s\uparrow 1} \phi'(s) = \sum_{k=1}^{\infty} k p_k = \mu$. $\square$

Branching processes

Proof.

- Let $\theta_m = \text{Prob}(Z_m = 0)$. By conditioning on Z_1, $\theta_m = \sum_{k=0}^{\infty} p_k \theta_{m-1}^k$, since each child of the first generation must die out.
- We check that there is a unique $0 \leqslant \rho < 1$ such that $\phi(\rho) = \rho$. Indeed, $\phi(0) \geqslant 0$, and $\phi(1) = 1$, $\phi'(1) = \mu > 1$ implies that $\phi(1 - \epsilon) < 1 - \epsilon$ for some $\epsilon > 0$. This proves the existence of a fixed point less than 1. The fixed point is unique since ϕ is strictly convex.
- $\theta_m \uparrow \rho$ follows since $\theta_0 = 0$, ϕ is increasing, and $\phi(\rho) = \rho$, so that θ_m is increasing and $\theta_m \leqslant \rho$ for all m.

$\square$