Lecture 10

David Aldous

18 September 2015

Ideas used in Lecture 9.

- If $\mu(t)$ (the distribution of $X(t)$) converges to a limit distribution π, then the limit π must be a stationary distribution and (mean proportion of time at i before t) $\to \pi_i$.
- Theorem about existence and formula for stationary distribution.
- Current material treated in sections 4.3, 4.4 of [PK].

Before continuing, some reminders about the distinction between random variables and probability distributions, and the difference between convergence of random variables and convergence of probability distributions.

Given a random variable X, its distribution describes the probabilities of the possible values;

- X die throw; $\pi_1 = 1/6, \pi_2 = 1/6, \ldots, \pi_6 = 1/6$
- Z standard Normal$(0, 1)$: density function $\phi(z)$ or distribution function $\Phi(z)$.

If X_1 and X_2 are independent dice throws then they have the same distribution. But $\mathbb{P}(X_1 = X_2) = 1/6$, so they are not the same random variable.

- Two random variables X and Y are regarded as "the same" if $\mathbb{P}(X = Y) = 1$.

"Convergence in distribution" of random variables means "convergence of their distributions". For real-valued random variables this is illustrated by the central limit theorem.

For i.i.d. (ξ_i) with mean θ and variance σ^2,

$$\bar{S}_n := \frac{\sum_{i-1}^{n} \xi_i - n\theta}{\sigma\sqrt{n}} \to_d Z$$

for standard Normal Z.. The "convergence in distribution" symbol $\to_d$ here means

$$\mathbb{P}(\bar{S}_n \le z) \to \mathbb{P}(Z \le z) = \Phi(z) \text{ for each } -\infty < z < \infty.$$

In a more advanced course there are several related notions of "convergence of random variables". Given $(X_1, X_2, \ldots)$ and also X_∞ then (going back to the formal math set-up, where a RV is a function from a sample space Ω to a range space, here $\mathbb{R}$) there is an event

$$\{\omega : X_n(\omega) \to X_\infty(\omega) \text{ as } n \to \infty\}$$

which has some probability. We say

$$X_n \to X_\infty \text{ a.s.}$$

to mean $\mathbb{P}(X_n \to X_\infty \text{ as } n \to \infty) = 1$. Here "a.s." is an abbreviation for "almost surely". Note that X_∞ might be a constant.

As a basic example, the "law of averages" can be formalized as the **strong law of large numbers**: *For i.i.d. (ξ_i) with mean θ,*

$$n^{-1} \sum_{i=1}^{n} \xi_i \to \theta \text{ a,s..}$$

Recall $T_i = \min\{t \geq 0 : X_t = i\}$ and define also the **return time**

$$T_i^+ = \min\{t \geq 1 : X_t = i\}.$$

Fix a reference state b.

Theorem

Suppose irreducible.
(a) If state space is finite then $\mathbb{E}_b T_b^+ < \infty$.
(b) Suppose $\mathbb{E}_b T_b^+ < \infty$. *Define*

$$a(b,i) = \mathbb{E}_b \sum_{s=1}^{T_b^+} \mathbb{1}_{(X(s)=i)}$$

$=$ *mean number of visits to* i *before returning to* b. *So* $a(b,b) = 1$. *Then*

$$\pi_i = \frac{a(b,i)}{\mathbb{E}_b T_b^+}$$

is a stationary distribution, and is the ***only*** *stationary distribution.*

Discussion.

(a) There is a calculation which checks this π does satisfy $\pi = \pi \mathbf{P}$.

(b) Because π is the same for each choice of b we have another formula

$$\pi_i = \frac{1}{\mathbb{E}_i T_i^+} \text{ for each } i.$$

(c) For an irreducible chain, the properties

$\mathbb{E}_i T_i^+ < \infty$ for some i

$\mathbb{E}_i T_i^+ < \infty$ for all i

are equivalent. When these hold we call the chain **positive-recurrent**.

(d) The theorem implies that every finite-state irreducible chain is positive-recurrent. So every finite-state irreducible chain has a unique stationary distribution.

- -

Limit theory for Markov chains combines this theorem with the next two theorems. The first is about convergence of **distributions**, the second is about convergence of **random variables**.

Theorem

If the chain is irreducible, positive-recurrent and aperiodic, then for any initial distribution

$$\mathbb{P}(X(t) = j) \to \pi_j \text{ as } t \to \infty$$

where π is the unique stationary distribution.

Theorem

Write

$$N_i(t) = \sum_{s=0}^{t-1} \mathbb{1}_{(X(s)=i)} = \text{ number of visits to } i \text{ before } t.$$

If the chain is irreducible and positive-recurrent, then for any initial distribution

$$t^{-1} N_i(t) \to \pi_i \text{ a.s. } \text{ as } t \to \infty.$$

Note this implies but is stronger than previous fact

$$\mathbb{E}[t^{-1} N_i(t)] \to \pi_i \text{ as } t \to \infty.$$

A final result is rather subtle. Note we only need this when the state space is infinite.

Proposition

If irreducible, if there exists a probability distribution π satisfying $\pi = \pi \mathbf{P}$, then the chain is positive-recurrent.

Then we can apply previous theorems and this π is the unique stationary distribution.

See texts for proofs. I want to focus on what these results say, in our specific examples.

Recall that for a doubly-stochastic chain the stationary distribution is the uniform distribution.

Card-shuffling examples. Theory implies that for any "non-stupid" random shuffle model, the distribution will eventually become closer and closer to uniform.

Simple random walk on the n-by-n torus.
[board]

Example: Umbrellas.

- A man owns K umbrellas, which are either at home or at work.
- He goes to work each morning, and goes home each evening.
- If raining, he takes an umbrella, if one is available. If not raining he does not take an umbrella.
- Model (unrealistic) that $\mathbb{P}(\text{ rain }) = p$, independently, each morning and evening.

To set up as a Markov chain, consider

$$X_t = \text{ number of umbrellas at home, end of day } t.$$

States $\{0, 1, \ldots, K\}$.

- -

$$p_{01} = p, \quad p_{00} = 1 - p$$

$$p_{K,K-1} = p(1-p), \quad p_{KK} = 1 - p(1-p)$$

$$p_{i,i+1} = p_{i,i-1} = p(1-p), \ p_{ii} = 1 - 2p(1-p), \ 1 \le i \le K-1.$$