UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Discussion 4
Fall 2021

1. Reversible Distributions

Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain with state space $\mathcal{S}$.

(a) Show that for every $m, k \in \mathbb{N}$, with $m \geq 1$, we have

$$P(X_k = i_0 \mid X_{k+1} = i_1, \dots, X_{k+m} = i_m) = P(X_k = i_0 \mid X_{k+1} = i_1),$$

for all states $i_0, i_1, \dots, i_m \in \mathcal{S}$. This is the *backwards Markov property.*

(b) In general, is the reversed chain (i.e. the chain $Y_k := X_{-k}$ for $k \in -\mathbb{N}$) a temporally homogeneous Markov chain? If not, provide an example.

(c) Show that if, in addition, the chain is reversible and started from a stationary distribution $X_0 \sim \pi$, then

$$(X_0, \dots, X_n) \overset{d}{=} (X_n, \dots, X_0).$$

Solution:

(a) By definition of conditional probability we can write

$$P(X_k = i_0 \mid X_{k+1} = i_1, \dots, X_{k+m} = i_m) = \frac{P(X_k = i_0, X_{k+1} = i_1, \dots, X_{k+m} = i_m)}{P(X_{k+1} = i_1, \dots, X_{k+m} = i_m)}.$$

Now using the Markov property the numerator can be written as

$$P(X_k = i_0, X_{k+1} = i_1) \prod_{j=2}^{m} P(X_{k+j} = i_j \mid X_{k+j-1} = i_{j-1}),$$

and the denominator can be written as

$$P(X_{k+1} = i_1) \prod_{j=2}^{m} P(X_{k+j} = i_j \mid X_{k+j-1} = i_{j-1}).$$

So the products cancel out and we obtain

$$\begin{aligned} P(X_k = i_0 \mid X_{k+1} = i_1, \dots, X_{k+m} = i_m) &= \frac{P(X_k = i_0, X_{k+1} = i_1)}{P(X_{k+1} = i_1)} \\ &= P(X_k = i_0 \mid X_{k+1} = i_1). \end{aligned}$$

(b) No. Any example given must not be a reversible (due to part (c)). Consider a three state chain with transition matrix

$$\begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

and initial distribution $[1, 0, 0]$. Then a Baye's rule computation gives $P(X_0 = 0 | X_1 = 1) = 1$, but $P(X_1 = 0 | X_2 = 1) = 1/2$. So indeed, these backward transition probaiblities depend on the initial distribution and the location in time.

(c) Consider the point $x_{0:n} = [x_0, x_1, \ldots, x_n]$ in the state space of $(X_0, \ldots, X_n)$. Then we have

$$\begin{aligned}
P(X_{0:n} = x_{0:n}) &= P(X_n = x_n) \prod_{k=0}^{n-1} P(X_k = x_k | X_{k+1} = x_{k+1}) \\
&= \pi_{x_n} \prod_{k=0}^{n-1} \frac{\pi_{x_k} P(X_{k+1} = x_{k+1} | X_k = x_k)}{\pi_{x_{k+1}}} \\
&= \pi_{x_n} \prod_{k=0}^{n-1} P_{x_{k+1} x_k} \\
&= P(X_0 = x_n) \prod_{k=0}^{n-1} P(X_{n-k} = x_k | X_{n-k-1} = x_{k+1}) \\
&= P(X_{n:0} = x_{0:n})
\end{aligned}$$

where in the first line we used the backwards Markov property proved in the last part, in the second line we used Bayes, in the third line we used reversibilitiy, and in the fifth line we used the Markov property. Thus we've shown that

$$(X_0, \ldots, X_n) \stackrel{d}{=} (X_n, \ldots, X_0).$$

2. Markov Chain Practice

Consider a Markov chain with three states 0, 1, and 2. The transition probabilities are $P(0,1) = P(0,2) = 1/2$, $P(1,0) = P(1,1) = 1/2$, and $P(2,0) = 2/3$, $P(2,2) = 1/3$.

(b) In the long run, what fraction of time does the chain spend in state 1?

(c) Suppose that X_0 is chosen according to the steady state distribution. What is $P(X_0 = 0 \mid X_2 = 2)$?

Solution:

(a) By solving $\pi P = \pi$, we have

$$\pi = \frac{1}{11} \begin{bmatrix} 4 & 4 & 3 \end{bmatrix}.$$

Thus, $\pi(1) = 4/11$.

(b) By the definition of conditional probability,

$$\begin{aligned}
P(X_0 = 0 \mid X_2 = 2) &= \frac{P(X_0 = 0, X_2 = 2)}{P(X_2 = 2)} \\
&= \frac{P(X_0 = 0, X_1 = 2, X_2 = 2)}{P(X_2 = 2)},
\end{aligned}$$

where we exploited the structure of the Markov chain in the last equality. Note that $P(X_2 = 2) = P(X_0 = 2)$ as X_0 is chosen according to π. Thus,

$$\frac{P(X_0 = 0, X_1 = 2, X_2 = 2)}{P(X_2 = 2)} = \frac{\pi(0) \cdot (1/2) \cdot (1/3)}{\pi(2)} = \frac{2}{9}.$$

3. **More Almost Sure Convergence**

(a) Suppose that, with probability 1, the sequence $(X_n)_{n\in\mathbb{N}}$ oscillates between two values $a \neq b$ infinitely often. Is this enough to prove that $(X_n)_{n\in\mathbb{N}}$ does *not* converge almost surely? Justify your answer.

(b) Suppose that Y is uniform on $[-1, 1]$, and X_n has distribution

$$P\big(X_n = (y + n^{-1})^{-1} \bigm| Y = y\big) = 1.$$

Does $(X_n)_{n=1}^{\infty}$ converge a.s.?

(c) Define random variables $(X_n)_{n\in\mathbb{N}}$ in the following way: first, set each X_n to 0. Then, for each $k \in \mathbb{N}$, pick j uniformly randomly in $\{2^k, \ldots, 2^{k+1} - 1\}$ and set $X_j = 2^k$. Does the sequence $(X_n)_{n\in\mathbb{N}}$ converge a.s.?

(d) Does the sequence $(X_n)_{n\in\mathbb{N}}$ from the previous part converge in probability to some X? If so, is it true that $\mathbb{E}[X_n] \to \mathbb{E}[X]$ as $n \to \infty$?

Solution:

(a) Yes. If a sequence oscillates between two values infinitely often, then it does not converge. Here, we have a sequence that oscillates between two values infinitely often (with probability 1), which means that the sequence does not converge (with probability 1). (Perhaps we could name this "almost surely not converging"!)

The above paragraph was very cumbersome to read, which is why we often abbreviate "with probability 1" with a.s. With this abbreviation, here is how the above justification reads: $(X_n)_{n\in\mathbb{N}}$ oscillates between two values infinitely often a.s., so $(X_n)_{n\in\mathbb{N}}$ does not converge a.s.

(b) Yes. Observe that when $Y = y \neq 0$, $(X_n)_{n\in\mathbb{N}}$ will converge to y^{-1}. When $Y = 0$, $(X_n)_{n\in\mathbb{N}}$ does not converge; however, $P(Y = 0) = 0$ since Y is a continuous random variable. In other words,

$$P(X_n \text{ does not converge as } n \to \infty) = P(Y = 0) = 0,$$
$$P(X_n \text{ converges as } n \to \infty) = P(Y \neq 0) = 1,$$

so $(X_n)_{n\in\mathbb{N}}$ converges a.s.

(c) No. The sequence $(X_n)_{n\in\mathbb{N}}$ oscillates between 0 and successively higher powers of two infinitely often a.s., so it does not converge a.s.

(d) Yes. Fix $\varepsilon > 0$. For $n \in \mathbb{Z}_+$, one has

$$P(|X_n| > \varepsilon) = \frac{1}{2^k},$$

where $k = \lfloor \log_2 n \rfloor$. As $n \to \infty$, the above probability goes to 0, so $X_n \to 0$ in probability. Intuitively, $(X_n)_{n\in\mathbb{N}}$ has infinitely many oscillations, so it cannot converge a.s. However, the probability of each oscillation shrinks to 0, so $(X_n)_{n\in\mathbb{N}}$ converges in probability.

The expectations do not converge. For all n, one has $\mathbb{E}[X_n] = 1$, so it is not the case that $\mathbb{E}[X_n] \to 0$ as $n \to \infty$. Hence, convergence in probability is not sufficient to imply that the expectations converge (in fact, almost sure convergence is not sufficient either).