Solutions to Norris' **Markov Chains**

Oliver Maynard

June 2026

I was working through *Markov Chains*, Cambridge University Press, 1997 by J.R. Norris and decided that I might provide my solutions in the off-chance that they are of use to others.

Contents

1 Discrete-time Markov chains **2**
1.1 Definition and Basic Properties 2

1 Discrete-time Markov chains

1.1 Definition and Basic Properties

Exercise 1.1.1. *Let $B_1, B_2, \dots$ be disjoint events with $\bigcup_{n=1}^{\infty} B_n = \Omega$. Suppose A is another event with $\mathbb{P}(A \mid B_n) = p$ for every n. Show that $\mathbb{P}(A) = p$. Deduce that if X and Y are discrete random variables, then the following are equivalent:*

(a) X and Y are independent;

(b) the conditional distribution of X given $Y = y$ is independent of y.

Proof. Consider first that

$$\mathbb{P}(A \mid B_n) = p \qquad \text{for } n = 1, 2, \dots,$$

and that by the law of total probability,

$$\mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(A \mid B_n)\, \mathbb{P}(B_n) = \sum_{n=1}^{\infty} p\, \mathbb{P}(B_n).$$

Since the B_n are disjoint with $\bigcup_{n=1}^{\infty} B_n = \Omega$, we have $\sum_{n=1}^{\infty} \mathbb{P}(B_n) = \mathbb{P}(\Omega) = 1$, and

$$\sum_{n=1}^{\infty} p\, \mathbb{P}(B_n) = p \sum_{n=1}^{\infty} \mathbb{P}(B_n) = p(1) = p.$$

So

$$\mathbb{P}(A) = p.$$

□

We now demonstrate the equivalence of (a) and (b).

(a) ⇒ (b). If X and Y are independent, then

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\, \mathbb{P}(Y = y) \qquad \text{for all } x, y.$$

Now consider that, as a direct result,

$$\mathbb{P}(X = x \mid Y = y) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} = \mathbb{P}(X = x).$$

Since this is true for every value x and every y, the conditional distribution of X given $Y = y$ is, for each y, exactly the distribution of X; i.e.,

$$\mathbb{P}(X \mid Y = y) = \mathbb{P}(X) \qquad \text{for all } y,$$

which is statement (b).

(b) ⇒ (a). If the conditional distribution of X given $Y = y$ is independent of y, then

$$\mathbb{P}(X \mid Y = y) = \mu(X) \qquad \text{for all } y \text{ (with } \mathbb{P}(Y = y) > 0),$$

where $\mu(X)$ is some generic distribution. Since $\mu(X)$ is constant for all y, we may use the above result regarding $\mathbb{P}(A \mid B_n) = p$ and obtain

$$\mathbb{P}(X) = \mu(X) \implies \mathbb{P}(X \mid Y = y) = \mathbb{P}(X).$$

So

$$\underbrace{\mathbb{P}(X = x \mid Y = y)}_{= \mathbb{P}(X = x)} = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)} \implies \mathbb{P}(X = x) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}.$$

Therefore,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\, \mathbb{P}(Y = y) \qquad \text{for all } x, y,$$

which is exactly the definition of independence between X and Y.