Lecture 3. Discrete Markov chains. The random variables τ_k and τ_y.

We can write many things (mostly old things that were written in a different form) using the notations P_x.

In particular: $p_{xy}^{(n)} = P_x\{\xi_n = y\}$. Also:

$$P_x\{\xi_1 = x_1,\, \xi_2 = x_2,\, ...,\, \xi_n = x_n\} = p_{x\,x_1} \cdot p_{x_1\,x_2} \cdot ... \cdot p_{x_{n-1}\,x_n}; \tag{3.1}$$

$$p_{x\,y}^{(n)} = \sum_{x_1,\,...,\,x_{n-1}} p_{x\,x_1} \cdot p_{x_1 x_2} \cdot ... \cdot p_{x_{n-2}\,x_{n-1}} \cdot p_{x_{n-1}\,y}. \tag{3.2}$$

Let us introduce the time τ_1 at which our process changes its state for the first time:

$$\tau_1 = \begin{cases} \min\{n > 0\colon\ \xi_n \neq \xi_0\} & \text{if there is such an } n, \\ +\infty & \text{if } \xi_0 = \xi_1 = \xi_2 = ... = \xi_n = ... \end{cases} \tag{3.3}$$

(that is if the process never changes its state). The random variable τ_1 takes positive integer values, and possibly, also the value ∞. This random variable is associated with infinitely many random variables ξ_i, $i = 0, 1, 2, ...$ (in fact, with all of them).

Of course, if a state $x \in X$ is such that $p_{xx} = 1$, then, starting from the point x, the state will almost surely never change:

$$P_x\{\tau_1 = \infty\} = P_x\{\xi_0 = \xi_1 = \xi_2 = ... = \xi_n = ...\} = 1. \tag{3.4}$$

Indeed,

$$P_x\{\xi_0 = \xi_1 = \xi_2 = ... = \xi_n\} = P_x\{x = \xi_1 = \xi_2 = ... = \xi_n\} = p_{xx} \cdot p_{xx} \cdot ... \cdot p_{xx} = 1 \cdot 1 \cdot ... \cdot 1 = 1; \tag{3.5}$$

and the probability (3.4) is obtained from this by limit passage as $n \to \infty$.

Theorem 3.1. *If $p_{xx} < 1$, then the random variable τ_1 is, almost surely with respect to the probability P_x (P_x-almost surely) finite, and it has the geometric distribution with parameter $1 - p_{xx}$; that is, it takes positive integer values k with probabilities*

$$P_x\{\tau_1 = k\} = (1 - p_{xx}) \cdot p_{xx}^{k-1}. \tag{3.6}$$

Proof. We have:

$$\{\tau_1 = k\} = \{\xi_0 = \xi_1 = ... = \xi_{k-1},\, \xi_k \neq \xi_{k-1}\}. \tag{3.7}$$

With respect to the measure P_x, the random variable ξ_0 is almost surely equal to x; so

$$\begin{aligned} P_x\{\tau_1 = k\} &= P_x\{\xi_1 = ... = \xi_{k-1} = x,\, \xi_k \neq x\} = P_x\Big(\bigcup_{y \neq x} \{\{\xi_1 = ... = \xi_{k-1} = x,\, \xi_k = y\}\Big) \\ &= \sum_{y \neq x} P_x\{\xi_1 = ... = \xi_{k-1} = x,\, \xi_k = y\} = \sum_{y \neq x} p_{xx}^{k-1} \cdot p_{xy} = p_{xx}^{k-1} \cdot \sum_{y \neq x} p_{xy} \\ &= (1 - p_{xx}) \cdot p_{xx}^{k-1}. \end{aligned} \tag{3.8}$$

We have, obviously:

$$P_x\{\tau_1 < \infty\} = P_x\big(\bigcup_{k=1}^{\infty}\{\tau_1 = k\}\big) = \sum_{k=1}^{\infty} P_x\{\tau_1 = k\} = \sum_{k=1}^{\infty}(1 - p_{xx}) \cdot p_{xx}^{k-1} = \frac{1 - p_{xx}}{1 - p_{xx}} = 1: \tag{3.9}$$

P_x-almost surely finite.

What do we know about the geometric distribution? Perhaps only one thing: that its expectation is equal to the inverse of the parameter. In our case:

$$E_x\tau_1 = \frac{1}{1 - p_{xx}}, \tag{3.10}$$

where E_x denotes the expectation corresponding to the probability P_x.

Now let us introduce the random variable $\eta_1 = \xi_{\tau_1}$: the point where we go after the first change of our position.

Theorem 3.2. *If $p_{xx} < 1$, the random variables τ_1 and η_1 are independent with respect to the probability P_x, and the distribution of the random variable η_1 is given by*

$$\pi_{xy} = P_x\{\eta_1 = y\} = \begin{cases} 0, & y = x, \\ \dfrac{p_{xy}}{1 - p_{xx}}, & y \neq x. \end{cases} \tag{3.11}$$

Proof. To be absolutely precise, we should say how the random variable η_1 is defined in the case when $\tau_1 = \infty$: there is no such thing as ξ_∞. True, by the previous theorem it occurs only with zero probability; but by our definition, random variables are functions that are defined on the whole Ω rather than only *almost everywhere*. We can do it like this: take a symbol, say, $*$, that does not belong to the set X; and define

$$\eta_1(\omega) = \begin{cases} \xi_{\tau_1(\omega)}(\omega) & \text{if } \tau_1(\omega) < \infty, \\ * & \text{if } \tau_1(\omega) = \infty. \end{cases} \tag{3.12}$$

Of course, the distribution will not depend on how we defined η_1 on the event $\{\tau_1 = \infty\}$ if this event has zero probability.

We have to find the joint probability mass function of τ_1, η_1:

$$P_x\{\tau_1 = k, \, \eta_1 = y\} = \begin{cases} 0, & y = x, \\ P_x\{\xi_1 = ... = \xi_{k-1} = x, \, \xi_k = y\}, & y \neq x. \end{cases} \tag{3.13}$$

The last probability is equal to

$$p_{xx}^{k-1} \cdot p_{xy} = (1 - p_{xx})\, p_{xx}^{k-1} \cdot \frac{p_{xy}}{1 - p_{xx}}. \tag{3.14}$$

From this we obtain at once two things: that the distribution of the random variable η_1 is

$$P_x\{\eta_1 = y\} = \begin{cases} 0, & y = x, \\ \dfrac{p_{xy}}{1 - p_{xx}}, & y \neq x \end{cases} \tag{3.15}$$

(of course it is very easy to check that the sum of these probabilities is equal to 1), and that the random variables τ_1 and η_1 are independent.

Let us define the random variables τ_k, $k = 0, 1, 2, 3, \ldots$, by $\tau_0 = 0$, and then recurrently:

$$\tau_k = \begin{cases} \min\{n > \tau_{k-1}\colon \xi_n \ne \xi_{\tau_{k-1}}\} & \text{if there are such } n, \\ \infty & \text{if there are no such } n. \end{cases} \tag{3.16}$$

The random variable τ_k is the time of the k-th change in the sequence $\xi_0, \xi_1, \xi_2, \ldots, \xi_n, \ldots$. If some $\tau_k = \infty$, then all subsequent $\tau_{k+1}, \tau_{k+2}, \ldots$ also are equal to $+\infty$.

Also we introduce the random variables η_k, $k = 0, 1, 2, 3, \ldots$, taking values in the space $X \cup \{*\}$, by

$$\eta_k = \begin{cases} \xi_{\tau_k}, & \tau_k < \infty, \\ *, & \tau_k = \infty. \end{cases} \tag{3.17}$$

Theorem 3.3. *The sequence $\eta_0, \eta_1, \eta_2, \ldots, \eta_k, \ldots$ is a Markov chain with respect to the probability P_x, with transition probabilities given by*

$$\pi_{xy} = \begin{cases} \dfrac{p_{xy}}{1 - p_{xx}}, & y \in X,\ y \ne x, \\ 0, & y = x \text{ or } y = * \end{cases} \tag{3.18}$$

if $x \in X$, $p_{xx} < 1$; if $x \in X$, $p_{xx} = 1$ or $x = *$, then

$$\pi_{xy} = \begin{cases} 0, & y \ne *, \\ 1, & y = *. \end{cases} \tag{3.19}$$

For example, if the transition matrix of the chain ξ_i is

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}, \tag{3.20}$$

the transition matrix $\Pi = (\pi_{xy})$ of the chain $\eta_0, \eta_1, \eta_2, \ldots, \eta_k, \ldots$ is

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.21}$$

(I am numbering the supplementary state $*$ the last; the entries at the main diagonal π_{xx} are all equal to 0 except $p_{**} = 1$).

Proof. We have to prove that for $x_0 = x$, and arbitrary $x_1, x_2, \ldots, x_k \in X \cup \{*\}$

$$P_x\{\eta_1 = x_1, \eta_2 = x_2, \ldots, \eta_k = x_k\} = \prod_{i=1}^{k} \pi_{x_{i-1} x_i}. \tag{3.22}$$

Both sides of this formula are equal to 0 if there is an $x_i = x_{i-1} \in X$, or if there is an $x_{i-1} \in X$ with $p_{x_{i-1}x_{i-1}} = 1$, but $x_i \neq *$, or if $x_{i-1} = *$, but $x_i \in X$.

Apart from these cases, we have:

$$
\begin{aligned}
P_x\{\eta_1 = x_1,\, \eta_2 = x_2,\, ...,\, \eta_k = x_k\} \\
&= \sum_{n_1=1}^{\infty} \sum_{n_2=1}^{\infty} ... \sum_{n_k=1}^{\infty} P\{\tau_t = n_1,\, \xi_{n_1} = x_1,\, \tau_2 = n_1 + n_2,\, \xi_{n_1+n_2} = x_2,\, ..., \\
&\qquad\qquad \tau_k = n_1 + ... + n_k,\, \xi_{n_1+...+n_k} = x_k\} \\
&= \sum_{n_1=1}^{\infty} p_{x_0x_0}^{n_1-1} p_{x_0x_1} \cdot \sum_{n_2=1}^{\infty} p_{x_1x_1}^{n_2-1} p_{x_1x_2} \cdot ... \cdot \sum_{n_k=1}^{\infty} p_{x_{k-1}x_{k-1}}^{n_k-1} p_{x_{k-1}x_k} = \prod_{i=1}^{k} \frac{p_{x_{i-1}x_i}}{1 - p_{x_{i-1}x_{i-1}}}.
\end{aligned} \tag{3.23}
$$

We cannot state, by analogy with Theorem 3.2, that the random variables $\tau_1, ..., \tau_k$ are independent from $\eta_1, ..., \eta_k$; but we have the following result:

Theorem 3.4. *Under the condition* $\eta_0 = \xi_0 = x_0,\, \eta_1 = x_1,\, \eta_2 = x_2,\, ...,\, \eta_k = x_k$, *the random variables* $\tau_1,\, \tau_2 - \tau_1,\, ...,\, \tau_k - \tau_{k-1}$ *are independent, and they have geometric distributions with parameters* $1 - p_{x_0x_0},\, 1 - p_{x_1x_1},\, ...,\, 1 - p_{x_{k-1}x_{k-1}}$:

$$
\begin{aligned}
P_{x_0}\{\tau_1 = n_1,\, \tau_2 - \tau_1 = n_2,\, ...,\, \tau_k - \tau_{k-1} = n_k | \eta_1 = x_1,\, ...,\, \eta_k = x_k\} \\
= \prod_{i=1}^{k} (1 - p_{x_{i-1}x_{i-1}}) p_{x_{i-1}x_{i-1}}^{n_i-1}.
\end{aligned} \tag{3.24}
$$

Proof. The conditional probability in the left-hand side of (3.24) is equal to the ratio of the probability of the intersection and the probability of the condition. The latter is given by (3.23), and the probability of the intersection is one summand in the sum (3.23). Dividing, we get (3.24).

Now we introduce another random time having to do with infinitely many random variables ξ_i; even a whole *family* of random times. We take, for $y \in X$,

$$
\tau_y = \begin{cases} \min\{n > 0\colon\, \xi_n = y\} & \text{if there are such } n, \\ \infty & \text{if there are no such } n. \end{cases} \tag{3.25}
$$

This is the first time of the chain at the point y – if $\xi_0 \neq y$, and if $x_0 = t$, the time of the first *return* to this state.

The random variable τ_y takes values 1, 2, ..., n, ..., and ∞ (possibly not all of these values); the distribution of this random variable is known if we know the probabilities

$$
f_{xy}^{(n)} = P_x\{\tau_y = n\} \tag{3.26}
$$

for integer values of n; the probability $P_x\{\tau_y = \infty\}$ is then obtained as $1 - \sum_{n=1}^{\infty} f_{xy}^{(n)}$.

Theorem 3.5. *For every natural* n *and* $x \in X$ *we have*:

$$
p_{xy}^{(n)} = f_{xy}^{(1)} \cdot p_{yy}^{(n-1)} + f_{xy}^{(2)} \cdot p_{yy}^{(n-1)} + ... + f_{xy}^{(n-1)} \cdot p_{yy}^{(1)} + f_{xy}^{(n)} \cdot 1. \tag{3.27}
$$

Proof. We have:

$$
\begin{aligned}
p_{xy}^{(n)} &= P_x\{\xi_n = y\} = P_x\Big(\bigcup_{k=1}^{n}(\{\tau_y = k\} \cap \{\xi_n = y\})\Big) \\
&= \sum_{k=1}^{n} P_x\{\xi_i \neq y \ \text{ for } \ 1 \leq i < k,\, \xi_k = y,\, \xi_n = y\} \qquad (3.28)\\
&= \sum_{k=1}^{n} \sum_{x_1,\,\dots,\,x_{k-1} \neq x} \ \sum_{x_{k+1},\,\dots,\,x_{n-1}} p_{x x_1} p_{x_1 x_2} \cdots p_{x_{k-1} y} \cdot p_{y x_{k+1}} p_{x_{k+1} x_{k+2}} \cdots p_{x_{n-1} y}.
\end{aligned}
$$

The k-th summand here is factorized as

$$
\sum_{x_1,\,\dots,\,x_{k-1} \neq x} p_{x x_1} p_{x_1 x_2} \cdots p_{x_{k-1} y} \cdot \sum_{x_{k+1},\,\dots,\,x_{n-1}} p_{y x_{k+1}} p_{x_{k+1} x_{k+2}} \cdots p_{x_{n-1} y}. \qquad (3.29)
$$

The first sum is nothing but $f_{xy}^{(k)}$; in the second we can change the summation indices:

$$
\sum_{x_1,\,\dots,\,x_{n-k-1}} p_{y x_1} p_{x_1 x_2} \cdots p_{x_{n-k-1} y}, \qquad (3.30)
$$

and we see that this is nothing but $P_y\{\xi_{n-k} = y\} = p_{yy}^{(n-k)}$. This proves (3.27) if we take into account that $p_{yy}^{(0)} = 1$.

The equality (3.27) can also be rewritten in the form

$$
p_{xy}^{(n)} = \sum_{k=0}^{n} f_{xy}^{(k)} \cdot p_{yy}^{(n-k)}, \qquad (3.31)
$$

because $f_{xy}^{(0)} = P_x\{\tau_y = 0\} = 0$. Equality (3.31) is true also for $n = 0$ if $y \neq x$: $0 = \sum_{k=0}^{0} 0$; but not for $n = 0$, $y = x$: the left-hand side is equal to 1, while the right-hand side is equal to 0.