Lecture 25. More on Markov chains.

Now we go to the general theory.

Theorem 25.1. *A sequence* $\xi_0, \xi_1, ..., \xi_n, ...$ *of random elements of the space* X *is a Markov chain with a prescribed initial distribution* $\{q_x\}$, $q_x \ge 0$, $\sum_x q_x = 1$, *on* X, *and a sequence of stochastic matrices* $P_1 = (p^1_{xy})$, $P_2 = (p^2_{xy})$, ..., $P_k = (p^k_{xy})$, ... *with entries indexed by elements of* X *as the transition matrices if and only if for all* n, x_0, x_1, ..., $x_n \in X$

$$P\{\xi_0 = x_0, \, \xi_1 = x_1, ..., \xi_n = x_n\} = q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1} x_n}. \tag{25.1}$$

Proof. The 'only if' part: Let $\xi_0, \xi_1, \xi_2, ..., \xi_n, ...$ be a Markov chain with initial distribution $\{q_x\}$ and transition matrices $P_1, P_2, ..., P_n, ...$; let us prove (25.1).

We have, by formula (23 – 24.2):

$$P\{\xi_0 = x_0, \, \xi_1 = x_1\} = P\{\xi_0 = x_0\} \cdot P\{\xi_1 = x_1 | \xi_0 = x_0\} = q_{x_0} \cdot p^1_{x_0 x_1}. \tag{25.2}$$

This, for $P\{\xi_0 = x_0\} = q_{x_0} > 0$. But if $P\{\xi_0 = x_0\} = 0$, (25.2) is also true, but for a different reason:

$$\{\xi_0 = x_0, \, \xi_1 = x_1\} \subseteq \{\xi_0 = x_0\}, \quad P\{\xi_0 = x_0, \, \xi_1 = x_1\} \le P\{\xi_0 = x_0\} = 0. \tag{25.3}$$

Now to $P\{\xi_0 = x_0, \, \xi_1 = x_1, \, \xi_2 = x_2\}$: by the same formula (23 – 24.2) and formula (23 – 24.15),

$$\begin{aligned} P\{\xi_0 = x_0, \xi_1 = x_1, \, \xi_2 = x_2\} &= P\{\xi_0 = x_0, \, \xi_1 = x_1\} \cdot P\{\xi_2 = x_2 | \xi_0 = x_0, \, \xi_1 = x_1\} \\ &= P\{\xi_0 = x_0, \, \xi_1 = x_1\} \cdot P\{\xi_2 = x_2 | \xi_1 = x_1\} = q_{x_0} \cdot p^1_{x_0 x_1} \cdot p^2_{x_1 x_2}. \end{aligned} \tag{25.4}$$

Again, this for $P\{\xi_0 = x_0, \, \xi_1 = x_1\} = q_{x_0} \cdot p^1_{x_0 x_1} > 0$; if it is equal to 0, by another, simpler reason:

$$P\{\xi_0 = x_0, \, \xi_1 = x_1, \, \xi_2 = x_2\} \le P\{\xi_0 = x_0, \, \xi_1 = x_1\} = 0. \tag{25.5}$$

Proceeding like this step by step, we get (23 – 24.15).

Now the 'if' part: the left-hand side of (23 – 24.15) is clearly $p^{n+1}_{x_n x_{n+1}}$, and the right-hand side

$$\begin{aligned} P\{\xi_{n+1} = x_{n+1} | \xi_n = x_n\} &= \frac{P\{\xi_n = x_n, \, \xi_{n+1} = x_{n+1}\}}{P\{\xi_n = x_n\}} \\ &= \frac{\sum_{x_0, \, x_1, \, ..., \, x_{n-1}} P\{\xi_0 = x_0, \, \xi_1 = x_1, \, ..., \, \xi_{n-1} = x_{n-1}, \, \xi_n = x_n, \, \xi_{n+1} = x_{n+1}\}}{\sum_{x_0, \, x_1, \, ..., \, x_{n-1}} P\{\xi_0 = x_0, \, \xi_1 = x_1, \, ..., \, \xi_{n-1} = x_{n-1}, \, \xi_n = x_n\}} \\ &= \frac{\sum_{x_0, \, x_1, \, ..., \, x_{n-1}} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1} x_n} \cdot p^{n+1}_{x_n x_{n+1}}}{\sum_{x_0, \, x_1, \, ..., \, x_{n-1}} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1} x_n}}, \end{aligned} \tag{25.6}$$

and after taking the factor that does not depend on the summation variables from under the summation sign, we get the same $p^{n+1}_{x_n x_{n+1}}$.

Theorem 25.2. *Let* q_x, $x \in X$, *be a probability distribution on* X; P_1, P_2, ..., P_n, ... *a sequence of stochastic matrices whose rows and columns are indexed by elements of* X. *Then there exists a Markov chain* ξ_0, ξ_1, ..., ξ_n, ... *with these initial distribution and transition matrices.*

Proof. Let us take $\Omega = [0,\,1)$, $\mathcal{F} = \mathcal{B}_{[0,\,1)}$, $P = \lambda_1$; and let us define the random variables ξ_0, ξ_1, ξ_2, ... in the following way. Let us number the elements of X by natural numbers:

$$X = \{x^1,\, x^2,\, ...,\, x^m\ (,\, ...)\} \tag{25.7}$$

(x with *sub*scripts is reserved for uses like (23 – 24.15): in particular, x_3 can be equal to x_2, whereas x^k is the numbering without repetitions).

Divide the interval [0, 1) into a finite or a countable (possibly finite) number of smaller intervals of lengths q_{x^1}, ..., q_{x^m} (, ...):

$$[\,0,\,1) = \bigcup_i [\,q_{x^1} + ... + q_{x^{i-1}},\ q_{x^1} + ... + q_{x^{i-1}} + q_{x^i}\,). \tag{25.8}$$

Here we are using the fact that $\sum_x q_x = 1$. (For a finite X, we could have used, with the same construction, also intervals $(0,\,1]$, $(0,\,q_{x^1}]$, $(q_{x^1},\,q_{x^1}+q_{x^2}]$, ..., $(q_{x^1}+...+q_{x^{m-1}},\,q_{x^1}+...+q_{x^{m-1}}+q_{x^m}=1]$; but for an infinite countable X using intervals with their right ends would require arranging small intervals from the right to the left.)

Now we define $\xi_0(\omega) = x^i$ for $\omega \in [\,q_{x^1} + ... + q_{x^{i-1}},\, q_{x^1} + ... + q_{x^{i-1}} + q_{x^i}\,)$.

Then we subdivide the intervals into smaller ones of lengths $q_{x^i} \cdot p^1_{x^i x^j}$:

$$\begin{aligned}
&[\,q_{x^1} + ... + q_{x^{i-1}},\, q_{x^1} + ... + q_{x^{i-1}} + q_{x^i}\,) \\
&\qquad = \bigcup_j [\,q_{x^1} + ... + q_{x^{i-1}} + q_{x^i}(p^1_{x^i x^1} + ... + p^1_{x^i x^{j-1}}), \\
&\qquad\qquad\qquad q_{x^1} + ... + q_{x^{i-1}} + q_{x^i}(p^1_{x^i x^1} + ... + p^1_{x^i x^{j-1}} + p^1_{x^i x^j})\big)
\end{aligned} \tag{25.9}$$

(it is possible because $\sum_j p^1_{x^i x^j} = 1$), and define $\xi_1(\omega)$ as being equal to x^j on the union of all intervals with x^j in the second place: $\xi_1(\omega) = x^j$ for $\omega \in \bigcup_i \big[q_{x^1} + ... + q_{x^{i-1}} + q_{x^i}(p^1_{x^i x^1} + ... + p^1_{x^i x^{j-1}}),\ q_{x^1} + ... + q_{x^{i-1}} + q_{x^i}(p^1_{x^i x^1} + ... + p^1_{x^i x^{j-1}} + p^1_{x^i x^j})\big)$. After that we subdivide the intervals of length $q_{x^i} \cdot p^1_{x_i x_j}$ into smaller still intervals of length $q_{x^i} \cdot p^1_{x^i x^j} \cdot p^2_{x^j x^k}$, and take $\xi_2 = x^k$ on every k-th small interval; etc.

It is easy to see that for the sequence thus constructed the equality (25.1) is satisfied.

Note that this construction has already been used in our course: in the proof of Theorem 10.3; only the random variables were numbered starting with 1 rather than with 0; and it was about *independent* random variables taking values 1 and 0 with probability 1/2. As we know (see formula (23 – 24.17)), a sequence of such random variables is a Markov chain with transition matrix $\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ (and with initial probabilities $q_0 = q_1 = 1/2$), so instead of dividing intervals according to some odd proportions, we divide each time in halves; and take $\xi_1(\omega) = 0$ for $0 \le \omega < 1/2$, $\xi_1(\omega) = 1$ for $1/2 \le \omega < 1$; $\xi_2(\omega) = 0$ for $\omega \in [0, 1/4) \cup [1/2, 3/4)$, and $\xi_2(\omega) = 1$ otherwise; $\xi_3(\omega) = 0$ for $\omega \in [0, 1/8) \cup [1/4, 3/8) \cup [1/2, 5/8) \cup [3/4, 7/8)$, etc. Make a picture, probably it will be better than what I drew on the blackboard.

I have told you that the kind of dependence between the random variables ξ_i in a Markov chain is an expression of some general property: *the future depends on the past only through the present.* In the definition of a Markov chain, the *past* was the event $\{\xi_0 = x_0, \xi_1 = x_1, ..., \xi_{n-1} = x_{n-1}\}$, the present, the event $\{\xi_n = x_n\}$, and the future, $\{\xi_{n+1} = x_{n+1}\}$. So the future was taken only one step from the present.

It turns out that this general Markov principle still holds for Markov chains if we consider a "longer" future.

Theorem 25.3. *Let $\xi_0, \xi_1, \xi_2, ..., \xi_n, ...$ be a Markov chain with initial probabilities q_x and transition matrices at the n-th step $P_n = (p^n_{xy})_{x,\,y \in X}$. Then for all natural n and $m > n$ and for all $x_0, x_1, x_2, ..., x_{n-1}, x_n, x_{n+1}, ..., x_m \in X$*

$$\begin{aligned} P\{\xi_{n+1} = x_{n+1}, ..., \xi_m = x_m | \xi_0 = x_0, \xi_1 = x_1, ..., \xi_{n-1} = x_{n-1}, \xi_n = x_n\} \\ = P\{\xi_{n+1} = x_{n+1}, ..., \xi_m = x_m | \, \xi_n = x_n\}, \end{aligned} \tag{25.10}$$

if only the left-hand side makes sense (i. e. if $P\{\xi_0 = x_0, \xi_1 = x_1, ..., \xi_{n-1} = x_{n-1}, \xi_n = x_n\} \neq 0$).

Proof. The left-hand side is

$$\frac{P\{\xi_0 = x_0, \xi_1 = x_1, ..., \xi_{n-1} = x_{n-1}, \xi_n = x_n, \xi_{n+1} = x_{n+1}, ..., \xi_m = x_m\}}{P\{\xi_0 = x_0, \xi_1 = x_1, ..., \xi_{n-1} = x_{n-1}, \xi_n = x_n\}}. \tag{25.11}$$

According to Theorem 25.1, this is equal to

$$\frac{q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1},\, x_n} \cdot p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1},\, x_m}}{q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1},\, x_n}} = p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1},\, x_m}. \tag{25.12}$$

The right-hand side in (25.10) is equal to

$$\frac{P\{\xi_n = x_n, \xi_{n+1} = x_{n+1}, ..., \xi_m = x_m\}}{P\{\xi_n = x_n\}}. \tag{25.13}$$

The events in the numerator and in the denominator here are unions of the corresponding disjoint events in (25.11) over all $x_0, x_1, x_2, ..., x_{n-1} \in X$; so the fraction (25.13) is equal to

$$\frac{\sum\limits_{x_0,\, x_1,\, ...,\, x_{n-1} \in X} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1},\, x_n} \cdot p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1},\, x_m}}{\sum\limits_{x_0,\, x_1,\, ...,\, x_{n-1} \in X} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1},\, x_n}}. \tag{25.14}$$

The factor $p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1},\, x_m}$ is common for all summands in the sum in the numerator, it can be taken from under the summation sign. The remaining sums cancel, and we get that the conditional probability $P\{\xi_{n+1} = x_{n+1}, ..., \xi_m = x_m | \xi_n = x_n\}$ is equal to $p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1}, x_m}$, which is the same as (25.12).

The theorem is proved.

It turns out that we can take a wider class of events as "the past" and as "the future" (keeping, though, the "present" as $\{\xi_n = x_n\}$).

Theorem 25.4. *Let* $\xi_0, \xi_1, \xi_2, ..., \xi_n, ...$ *be a Markov chain*; n *and* $m > n$ *natural numbers*, C *a subset of* X^n, *and* $D \subseteq X^{m-n}$; *and* x_n *an element of* X.

Then

$$P\{(\xi_{n+1}, ..., \xi_m) \in D | (\xi_0, \xi_1, ..., \xi_{n-1}) \in C, \xi_n = x_n\} = P\{(\xi_{n+1}, ..., \xi_m) \in D | \xi_n = x_n\} \tag{25.15}$$

Proof. The left-hand side is equal to the ratio of two unconditional probabilities, and each of them can be represented as a sum of probabilities of disjoint events: this left-hand side is equal to

$$\frac{\sum\limits_{\substack{(x_0, x_1, ..., x_{n-1}) \in C \\ (x_{n+1}, ..., x_m) \in D}} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1}, x_n} \cdot p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1}, x_m}}{\sum\limits_{(x_0, x_1, ..., x_{n-1}) \in C} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1}, x_n}}. \tag{25.16}$$

The sum in the numerator is clearly equal to

$$\Big[\sum_{(x_0, x_1, ..., x_{n-1}) \in C} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1}, x_n} \Big] \cdot \Big[\sum_{(x_{n+1}, ..., x_m) \in D} p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1}, x_m} \Big], \tag{25.17}$$

the first factor cancels, and the left-hand side of (25.15) is equal to

$$\sum_{(x_{n+1}, ..., x_m) \in D} p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1}, x_m}. \tag{25.18}$$

The right-hand side of (25.15) is equal to

$$\frac{\sum\limits_{\substack{(x_0, x_1, ..., x_{n-1}) \in X^n \\ (x_{n+1}, ..., x_m) \in D}} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1}, x_n} \cdot p^{n+1}_{x_n x_{n+1}} \cdot ... \cdot p^m_{x_{m-1}, x_m}}{\sum\limits_{(x_0, x_1, ..., x_{n-1}) \in X^n} q_{x_0} \cdot p^1_{x_0 x_1} \cdot ... \cdot p^n_{x_{n-1}, x_n}}, \tag{25.19}$$

and it is handled the same way, with the same result.

In particular, if we take $x_n = x \in X$, $D = \{(x_{n+1}, ..., x_{m-1}, x_m) : x_m = y\} = X \times ... \times X \times \{y\}$, we have:

$$P\{\xi_m = y | (\xi_0, \xi_1, ..., \xi_{n-1}) \in C, \xi_n = x\} = P\{\xi_m = y | \xi_n = x\}. \tag{25.20}$$

We'll denote this conditional probability p^{nm}_{xy}, and will refer to it as *the transition probability from* x *after the* n*-th step to* y *after the* m*-th step* (but of course, this expression is so long that we'll not be using it very often). The corresponding matrix $(p_{xy}^{nm})_{x,\,y\in X}$ will be denoted P^{nm}, and called *the transition matrix from step number n to step number m.*

Using formula (25.18), we get an expression for p_{xy}^{nm}:

$$p_{xy}^{nm} = \sum_{x_{n+1},\,...,\,x_{m-1}\in X} p_{xx_{n+1}}^{n+1} \cdot p_{x_{n+1},\,x_{n+2}}^{n+2} \cdot ... \cdot p_{x_{m-2},\,x_{m-1}}^{m-1} \cdot p_{x_{m-1},\,y}^{m}. \tag{25.21}$$

This formula allows us to evaluate the more-than one step transition probabilities p_{xy}^{nm} (starting from one-step transition matrices P_k) *in principle*; but in fact the calculations using this formula are sometimes not feasible. E. g., if the space X consists of two elements, and $m-n=1000$ (1000 steps), the sum (25.21) contains $2^{999}>10^{300}$ summands, which we cannot handle with or without computers.

Formula (25.20) and the notations p_{xy}^{nm}, P^{nm} were introduced only for $m>n$; but we can also consider the corresponding conditional probabilities and notations for $m=n$: it is much simpler:

$$P\{\xi_n=y|(\xi_0,\,\xi_1,\,...,\,\xi_{n-1})\in C,\,\xi_n=x\} = P\{\xi_n=y|\xi_n=x\} = p_{xy}^{nn} = \delta_{xy} = \begin{cases} 1, & y=x, \\ 0, & y\neq x; \end{cases} \tag{25.22}$$

the corresponding matrix is the identity matrix: $P^{nn}=I$.

It is clear that the transition matrix from step n to step number $n+1$ is $P^{n,\,n+1}=P_{n+1}$, our $(n+1)$-th original transition matrix.

Now let us have a theorem about the properties of the transition matrices P^{nm}, which will allow us to evaluate these matrices:

Theorem 25.5. *For $0\le n\le m\le r$ we have*:

$$P^{nr}=P^{nm}\cdot P^{mr}. \tag{25.23}$$

Proof. Remembering how we multiply matrices, we see that (25.22) means that for all $x,\,y\in X$

$$p_{xy}^{nr}=\sum_{z\in X} p_{xz}^{nm}\cdot p_{zy}^{mr}. \tag{25.24}$$

We can prove it using formula (25.21); but I better give another proof.

By Theorem 23 − 24.4 (the total conditional probability formula), with $C=\{\xi_r=y\}$, $B=\{\xi_n=x\}$, $k=z$, $A_z=\{\xi_m=z\}$, we have:

$$p_{xy}^{nr}=P\{\xi_r=y|\xi_n=x\}=\sum_{z\in X}P\{\xi_m=z|\xi_n=x\}\cdot P\{\xi_r=y|\xi_n=x,\,\xi_m=z\}. \tag{25.25}$$

The first factor in the z-th summand in the right-hand side is p_{xz}^{nm}; the second factor is, by Theorem 25.4, equal to $P\{\xi_r=y|\xi_m=z\}=p_{zy}^{mr}$. This proves (25.24), and (25.23) with it.

Theorem 25.6. *We have*:

$$P^{nm} = \prod_{k=n+1}^{m} P_k = P_{n+1} \cdot P_{n+2} \cdot ... \cdot P_m \tag{25.26}$$

(I wrote the product with dots because the product of matrices depends on the order in which they are taken, and I had to explain the convention as for in which order the product $\prod$ should be taken).

Proof. By (25.23) with m instead of r and $m-1$ instead of m, we get:

$$P^{nm} = P^{n,\, m-1} \cdot P^{m-1,\, m} = P^{n,\, m-1} \cdot P_m; \tag{25.27}$$

applying the same a second time, we obtain

$$P^{nm} = P^{n,\, m-2} \cdot P_{m-1} \cdot P_m, \tag{25.28}$$

etc. Finally we come to (25.26).

This is how p_{xy}^{nm} can be evaluated. In our example of X consisting of two elements, $m-n = 1000$, multiplying one thousand 2×2-matrices takes, for each multiplication, 8 multiplications and 4 additions, which leads to approximately $12,000$ operations: quite feasible.

By the way, formula (25.26) holds also for $m = n$ if we take, as it is reasonable to do, that the product $\prod_{k=n+1}^{n}$ of an empty collection of matrices is the identity matrix I.

The following material was not in Lecture 25, but I am including it here, because this is the place for it.

Theorem 25.7 (a little generalization of Theorem 25.4). *Let* $\xi_0, \xi_1, \xi_2, ..., \xi_n, ...$ *be a Markov chain*; n *and* $m > n$ *natural numbers*, C *a subset of* X^{n+1}, *and* $D \subseteq X^{m-n+1}$; *and* x *an element of* X.

Then

$$P\{(\xi_n, ..., \xi_m) \in D | (\xi_0, \xi_1, ..., \xi_n) \in C, \xi_n = x\} = P\{(\xi_n, ..., \xi_m) \in D | \xi_n = x\} \tag{25.29}$$

Proof. Let us introduce the $(m-n)$-dimensional set $D_x = \{(x_{n+1}, ..., x_m) : (x, x_{n+1}, ..., x_m) \in D\}$, and the n-dimensional $C_x = \{(x_0, ..., x_{n-1}) : (x_0, ..., x_{n-1}, x) \in C\}$. The conditioning event $\{(\xi_0, \xi_1, ..., \xi_n) \in C, \xi_n = x\}$ in the left-hand side is the same as $\{(\xi_0, \xi_1, ..., \xi_{n-1}) \in C_x, \xi_n = x\}$; and also $\{\xi_n = x\} \cap \{(\xi_n, ..., \xi_m) \in D\} = \{\xi_n = x\} \cap \{(\xi_{n+1}, ..., \xi_m) \in D_x\}$, $\{(\xi_0, \xi_1, ..., \xi_n) \in C, \xi_n = x\} \cap \{(\xi_n, ..., \xi_m) \in D\} = \{(\xi_0, \xi_1, ..., \xi_{n-1}) \in C_x, \xi_n = x\} \cap \{(\xi_{n+1}, ..., \xi_m) \in D_x\}$. So by Theorem 25.4 we obtain (25.29).

This theorem means that the present can be *included* both in the past and in the future.

Theorem 25.7 can be reformulated as follows:

For $0 \leq k \leq l$, let $\mathcal{F}_{[k,\, l]}$ be the σ-algebra in the space Ω generated by the random variables $\xi_k, ..., \xi_l$ (that is, the σ-algebra of all events of the form $\{(\xi_k, ..., \xi_l) \in E\}$,

E being an arbitrary subset of X^{l-k+1}; we should have introduced these σ-algebras in Lecture 11: it's in the same spirit as the σ-algebras introduced there).

For every event $A \in \mathcal{F}_{[0,\,n]}$ $(= \mathcal{F}_{\le n})$, *and every event* $B \in \mathcal{F}_{[n,\,m]}$, *and every* $x \in X$ *we have*:

$$P(B|A \cap \{\xi_n = x\}) = P\{B|\xi_n = x\}. \tag{25.30}$$

Theorem 25.8. *For every event* $A \in \mathcal{F}_{\le n}$, *and every event* $B \in \mathcal{F}_{\ge n}$, *and every* $x \in X$ *formula* (25.30) *holds.*

Proof. Both sides of (25.30) are measures as functions of B. By Theorem 25.7, these measures coincide with one another on the *algebra* $\bigcup_{m=n}^{\infty} F_{[n,\,m]}$; so by the uniqueness theorem they have to coincide on the σ-algebra $\sigma\big(\bigcup_{m=n}^{\infty} F_{[n,\,m]}\big) = \mathcal{F}_{\ge n}$.

So our Markovian principle of the future depending on the past only through the present holds if we interpret the past as all events in $\mathcal{F}_{\le n}$, as the future all events $\in \mathcal{F}_{\ge n}$, and as the present, still the event $\{\xi_n = x\}$.