Lecture 8 · 7.5.2021 · 10:15–12:00 via Zoom

Conditional expectation (end)

Example. Let $(X_{n,m})_{n,m\geqslant 0}$ a double sequence of i.i.d. r.v. with values in $\mathbb{N}_{\geqslant 0}=\{0,1,2,\ldots\}$ (in particular $\mathbb{P}(X_{1,1}<\infty)=1$). Let $Z_0=1$ and then I define recursively

$$Z_n = X_{n,1}+\cdots+X_{n,Z_{n-1}},$$

for $n\geqslant 1$. Note that Z_n is a integer valued random variable.

The r.v.s $(Z_n)_{n\geqslant 0}$ model the evolution of the size of population of individuals where at every step n the m-th individual give rise to $X_{n,m}\sim X_{1,1}$ individuals which go the next generation.

We want to compute the generating function $f_n(\theta):=\mathbb{E}[\theta^{Z_n}]$ of Z_n for every $\theta\in(0,1)$.

Note that we can more rigorously define

$$Z_n=\sum_{m\geqslant 1}\mathbb{1}_{Z_{n-1}\geqslant m}X_{n,m},$$

and observe that Z_0 it is a.s. finite and if Z_{n-1} is a.s. finite then the above sum is also made a.s. of finitely many summands and therefore Z_n is a.s. finite if $\mathbb{P}(X_{1,1}<\infty)=1$, i.e. this proves that $\mathbb{P}(Z_n<\infty)=1$ for all $n\geqslant 0$ by induction, assuming $\mathbb{P}(X_{1,1}<\infty)=1$.

We know that $f_0(\theta)=\mathbb{E}[\theta^{Z_0}]=\theta$. Let us call $f(\theta):=\mathbb{E}[\theta^{X_{1,1}}]$. How we compute

$$f_n(\theta)=\mathbb{E}[\theta^{Z_n}]=\mathbb{E}[\theta^{X_{n,1}+\cdots+X_{n,Z_{n-1}}}]?$$

We note that $(X_{n,m})_{m\geqslant 1}$ is independent of $Z_{n-1}\hat{\in}\sigma(X_{k,m}:0\leqslant k\leqslant n-1,m\geqslant 1)$ since the family $(X_{n,m})_{n,m\geqslant 0}$ is iid. In this case we can condition the averange on the value of Z_{n-1}: by the theorem we proved in the last lecture (Prop 14 in Note 2)

$$\mathbb{E}[\theta^{X_{n,1}+\cdots+X_{n,Z_{n-1}}}]=\mathbb{E}[\mathbb{E}[\theta^{X_{n,1}+\cdots+X_{n,Z_{n-1}}}|Z_{n-1}]]=\mathbb{E}[\varphi(Z_{n-1})]$$

with

$$\varphi(z)=\mathbb{E}[\theta^{X_{n,1}+\cdots+X_{n,z}}],\qquad z=0,1,2,\ldots$$

But now we can compute this easily since cond. exp. has disappeared thanks to independence, for $z\geqslant 0$

$$\varphi(z)=\mathbb{E}[\theta^{X_{n,1}}]\cdots\mathbb{E}[\theta^{X_{n,z}}]=(\mathbb{E}[\theta^{X_{1,1}}])^z=(f(\theta))^z$$

which means that

$$f_n(\theta)=\mathbb{E}[\theta^{X_{n,1}+\cdots+X_{n,Z_{n-1}}}]=\mathbb{E}[\varphi(Z_{n-1})]=\mathbb{E}[(f(\theta))^{Z_{n-1}}]=f_{n-1}(f(\theta)).$$

Therefore we have shown that f_n solve the recursive equation

$$f_0(\theta) = \theta, \qquad f_n(\theta) = f_{n-1}(f(\theta))$$

which has unique solution $f_n(\theta) = f^{\circ n}(\theta)$. This is very useful if one would like to understand what happens to Z_n as $n \to \infty$, i.e. how the population behave on long time. Intersting question: does it become extinct with probability 1. Note that if $Z_n = 0$ then $Z_k = 0$ for all $k \geqslant n$.

Regular conditional probabilities

Given a σ-algebra $\mathscr{G} \subseteq \mathscr{F}$ we can consider the family of r.v.

$$A \in \mathscr{F} \mapsto \mathbb{P}(A|\mathscr{G}) := \mathbb{E}[\mathbb{1}_A|\mathscr{G}] \in L^\infty(\Omega, \mathscr{G}, \mathbb{P}) \subseteq L^\infty(\Omega, \mathscr{F}, \mathbb{P})$$

These r.v. satisfy **almost surely** the following equalities:

a) $\mathbb{P}(\emptyset|\mathscr{G}) = 0$, $\mathbb{P}(A^c|\mathscr{G}) = 1 - \mathbb{P}(A|\mathscr{G})$

b) $\mathbb{P}(\cup_n A_n|\mathscr{G}) = \sum_n \mathbb{P}(A_n|\mathscr{G})$ for a family $(A_n)_{n\geqslant 1} \subseteq \mathscr{F}$ of pairwise disjoint events.

These relations shows that the map $A \in \mathscr{F} \mapsto \mathbb{P}(A|\mathscr{G})$ behaves like a probability measure. So we would like to think to $\mathbb{P}(\cdot|\mathscr{G})$ as a random probability measure $\mathbb{P}_{\mathscr{G}}$, i.e. something like

$$\mathbb{P}_{\mathscr{G}}: \omega \in \Omega \mapsto \mathbb{P}_{\mathscr{G}}(\omega) \in \Pi(\Omega, \mathscr{F})$$

where $\Pi(\Omega, \mathscr{F})$ is the set of all probability measures on $(\Omega, \mathscr{F})$. This is in this generality not possible, because the proprerties a),b) are true only a.s., that is for any choice of A or of $(A_n)_n$ one has different exceptional set $\mathscr{N}_A$, $\mathscr{N}_{(A_n)_n}$ in which the property is not satisfied and unfortunately one cannot construct an exceptional universal measurable set valid for all the possible choices of A, (A_n), because we have uncountably many choices here.

In some situations however this is possible. In this case we say that the family $(\mathbb{P}(A|\mathscr{G}))_{A\in\mathscr{F}}$ admits a **regular conditional version**, or that we have a **regular conditional probability** for $\mathbb{P}$ given $\mathscr{G}$. This means that there exists a map

$$\mathbb{P}_{\mathscr{G}}: \Omega \to \Pi(\Omega, \mathscr{F}) \subseteq (\Omega \times \mathscr{F} \to [0,1])$$

such that for all $A \in \mathscr{F}$ it holds

$$\mathbb{P}_{\mathscr{G}}(\omega, A) = \mathbb{P}(A|\mathscr{G})(\omega), \qquad \mathbb{P} - a.e. \, \omega \in \Omega.$$

In case we have a regular conditional probability, then for any $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ we can express the conditional expectation wrt. $\mathscr{G}$ as an integral:

$$\mathbb{E}[X|\mathscr{G}](\omega) = \int X(\omega')\mathbb{P}_{\mathscr{G}}(\omega, \mathrm{d}\omega'), \qquad \mathbb{P} - a.e. \, \omega \in \Omega.$$

(this is not difficult to prove from the definition of reg. cond. prob. and cond. exp.)

Regular conditional probabilities are guaranteed to exist when $(\Omega, \mathscr{F})$ is a **Polish space**: i.e. complete, metrisable topological space endowed with the Borel σ-algebra. Example: $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{\mathbb{N}}$, $C([0,1],\mathbb{R}^n)$.

Filtrations & stopping times

Definition. *A (discrete time) stochastic process is just a family $(X_n)_{n\geqslant 0}$ of random variables indexed by $\mathbb{N}_{\geqslant 0}$ (or $\mathbb{N}_{\geqslant 1}$).*

We think to the index n at *time* and to the sequence $X_1, X_2, \ldots$ as the description of random phenoma which evolves in time. Time take with it a notion of "past", "present" and "future".

This is encoded in the notion of *filtration*:

Definition. *A filtration $(\mathscr{F}_n)_{n\geqslant 0}$ is an increasing family of sub-σ-algebras of $\mathscr{F}$, i.e.*

$$\mathscr{F}_n \subseteq \mathscr{F}_{n+1}$$

for all $n \geqslant 0$. We let $\mathscr{F}_\infty = \sigma(\mathscr{F}_n : n \geqslant 0)$, that is the smallest σ-algebra which contains all the $(\mathscr{F}_n)_n$.

A filtration represents the flow of time, in the sense that $\mathscr{F}_n$ is the information I dispose at time n.

Example. If $X = (X_n)_{n\geqslant 0}$ is a stochastic process then we can always consider its *natural filtration* $(\mathscr{F}_n^X)_{n\geqslant 0}$

$$\mathscr{F}_n^X = \sigma(X_0, X_1, \ldots, X_n).$$

It is easy to check that indeed $\mathscr{F}_n^X \subseteq \mathscr{F}_{n+1}^X$. This filtration encode the information given by the observation of the process X as time passes.

Example. Let $\Omega = (0,1]$ and define

$$\mathscr{F}_n = \{(k/2^n, (k+1)/2^n] : k = 0, \ldots, 2^n - 1\} \subseteq \mathscr{F} = \mathscr{B}([0,1])$$

then $(\mathscr{F}_n)_{n\geqslant 0}$ is a filtration and $\mathscr{F}_\infty = \mathscr{F}$. Here n represent the precision of our observation of a point in $[0,1]$.

Definition. *We say that a process $X = (X_n)_{n\geqslant 0}$ is adapted to the filtration $(\mathscr{F}_n)_{n\geqslant 0}$ iff*

$$X_n \mathrel{\hat{\in}} \mathscr{F}_n$$

for all $n \geqslant 0$. A process X is <u>*previsible*</u> *to the filtration $(\mathscr{F}_n)_{n\geqslant 0}$ iff*

$$X_{n+1} \hat{\in} \mathscr{F}_n$$

for all $n \geqslant 0$.

The natural filtration $\mathscr{F}^X$ of a process X is the smallest filtration for which the process is adapted.

I want to define now <u>stopping times</u>. A stopping time is a rule to determine how to stop given what happened in the past.

Definition. *A stopping time $T\colon \Omega \to \mathbb{N}^* := \mathbb{N} \cup \{+\infty\}$ for the filtration $(\mathscr{F}_n)_{n\geqslant 0}$ is a r.v. with values in $\mathbb{N}^*$ such that*

$$\{T \leqslant n\} \in \mathscr{F}_n,$$

for all $n \geqslant 0$.

This is equivalent to require that $\{T = n\} \in \mathscr{F}_n$ for all $n \geqslant 0$.

Example. The first time T we observe "head" in a repeated launch of a coin, is a stopping time wrt. the natural filtration of this problem. However, the last time S I observe "head" is not a stopping time.

The notion of stopping time encode a "fair" stopping rule, i.e. a rule which does not use information from future to make a decision (to stop or no).

Definition. *The σ-algebra $\mathscr{F}_T$ of the stopping time T wrt. the filtration $(\mathscr{F}_n)_{n\geqslant 0}$ is defined as*

$$\mathscr{F}_T := \{A \in \mathscr{F} : A \cap \{T \leqslant n\} \in \mathscr{F}_n \text{ for all } n \in \mathbb{N}^*\}$$

Exercise. Show that $\mathscr{F}_T$ is a σ-algebra.
