Lecture 28. Martingales and stopping times. Application to diffusion processes and Dirichlet problem for elliptic differential equations.

Proof of Theorem 27.6. Let us introduce the random time

$$\tau = \begin{cases} \min\{t_i\colon 1 \le i \le n, \eta_{t_i} \ge a\} & \text{if there is at least one such } t_i, \\ t_{\max} & \text{if there is no such } t_i; \end{cases} \tag{28.1}$$

τ is a stopping time – taking finitely many finite values.

The event $\{\max(\eta_{t_1}, \eta_{t_2}, ..., \eta_{t_n}) \ge a\}$ clearly coincides with the event $\{\eta_\tau \ge a\}$. Applying a Chebyshev-type inequality to the nonnegative random variable η_τ, we get:

$$P\{\max(\eta_{t_1}, \eta_{t_2}, ..., \eta_{t_n}) \ge a\} = P\{\eta_\tau \ge a\} \le \frac{E\eta_\tau}{a}, \tag{28.2}$$

and by Theorem 27.4 this is $\le \dfrac{E\eta_{t_{\max}}}{a}$.

What Kolmogorov started this with was a particular case:

Theorem 28.1. *Let* $\xi_1, \xi_2, ..., \xi_n$ *be independent random variables with* $E\xi_i = 0$ *and finite variances* $\mathrm{Var}(\xi_i) = \sigma_i^2$. *Then for every* $b > 0$

$$P\{\max(|\xi_1|, |\xi_1 + \xi_2|, ..., |\xi_1 + \xi_2 + ... + \xi_n|) \ge b\} \le \frac{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}{b^2}. \tag{28.3}$$

The classical Kolmogorov inequality is a strengthening of the classical Chebyshev's inequality:

$$P\{|\xi_1 + ... + \xi_i| \ge b\} \le \frac{\mathrm{Var}(\xi_1 + ... + \xi_i)}{b^2} = \frac{\sigma_1^2 + ... + \sigma_i^2}{b^2} \le \frac{\sigma_1^2 + ... + \sigma_n^2}{b^2}. \tag{28.4}$$

Let me show how the classical Kolmogorov's inequality (28.3) is a particular case of the Kolmogorov-type inequality (27.29). First we prove an auxiliary theorem:

Theorem 28.2. *Suppose* η_t, $t \in T$, *is a martingale with respect to* $(\mathcal{F}_t)$; *and suppose* $E\eta_t^2 < \infty$. *Then the random function* $\zeta_t = \eta_t^2$, $t \in T$, *is a submartingale.*

Proof. The measurability requirement is quite simple. Requirement (26.10): for $s \ge t$

$$\begin{aligned} E(\eta_s^2 \| \mathcal{F}_{\le t}) &= E\big(\big(\eta_t + (\eta_s - \eta_t)\big)^2 \big\| \mathcal{F}_{\le t}\big) \\ &= E\big(\eta_t^2 + 2\eta_t(\eta_s - \eta_t) + (\eta_s - \eta_t)^2 \big\| \mathcal{F}_{\le t}\big) \\ &= E(\eta_t^2 \| \mathcal{F}_{\le t}) + 2E\big(\eta_t(\eta_s - \eta_t)\big\| \mathcal{F}_{\le t}\big) + E\big((\eta_s - \eta_t)^2 \big\| \mathcal{F}_{\le t}\big). \end{aligned} \tag{28.5}$$

The first summand is equal to η_t^2. In the second, you can take η_t from under the sign of the conditional expectation:

$$E\big(\eta_t(\eta_s - \eta_t)\big\| \mathcal{F}_{\le t}\big) = \eta_t \cdot E(\eta_s - \eta_t \| \mathcal{F}_{\le t}) = 0. \tag{28.6}$$

The third is nonnegative; so

$$E(\eta_s^2 \| \mathcal{F}_{\le t}) \ge \eta_t^2, \tag{28.7}$$

which was to be proved.

It can be proved that if η_t is a martingale, and a function $f(y)$ is concave upwards, then $f(\eta_t)$ is a submartingale (provided $E|f(\eta_t)| < \infty$). The proof uses the fact that for every y_0 there exists a constant $m(y_0)$ such that $f(y) \ge f(y_0) + m(y_0) \cdot (y - y_0)$ (for $f(y) = y^2$ this is the inequality $y^2 \ge y_0^2 + 2y_0 \cdot (y - y_0)$; if $f(y)$ is concave and *differentiable* at $y = y_0$, this is the fact that an upward-concave function lies above its tangent line).

In Theorem 28.1 the sequence $\eta_i = \xi_1 + ... + \xi_i$, $i = 1, ..., n$, is a martingale with respect to the σ-algebras $\mathcal{F}_{\le i} = \sigma(\xi_1, ..., \xi_i)$. By Theorem 28.2, the sequence $\zeta_i = (\xi_1 + ... + \xi_i)^2$, $i = 1, ..., n$, is a *sub*martingale; and by inequality (27.29) we have:

$$P\{\max(|\xi_1|, |\xi_1 + \xi_2|, ..., |\xi_1 + ... + \xi_n|) \ge b\} = P\{\max(\zeta_1, \zeta_2, ..., \zeta_n) \ge b^2\} \le \frac{E\zeta_n^2}{b^2}, \tag{28.8}$$

which is equal to the right-hand side of (28.3).

Theorem 28.3. *Let* η_t, $t \in T$, *be a nonnegative submartingale with sample functions that are continuous from the right* (or from the left). *Then for every* $a > 0$

$$P\{\sup_{t \in T} \eta_t \ge a\} \le \frac{\sup_{t \in T} E\eta_t}{a}. \tag{28.9}$$

This is a continuous-time counterpart of Theorem 27.6.

Proof. Clearly there exists a sequence of finite subsets $T_1 \subseteq T_2 \subseteq ... \subseteq T_n \subseteq ... \subseteq T$ such that

$$\sup_{t \in T} \eta_t = \lim_{n \to \infty} \max_{t \in T_n} \eta_t. \tag{28.10}$$

It would seem that the event

$$\{\sup_{t \in T} \eta_t \ge a\} = \bigcup_{n=1}^{\infty} \{\max_{t \in T_n} \eta_t \ge a\}; \tag{28.11}$$

but this is not true: it is possible that all $\max_{t \in T_n} Y_t$ are less than a, and $\sup_{t \in T} Y_t = a$. What is true, however, is that for every $\varepsilon \in (0, a)$

$$\{\sup_{t \in T} \eta_t \ge a\} \subseteq \bigcup_{n=1}^{\infty} \{\max_{t \in \mathfrak{T}_n} \eta_t \ge a - \varepsilon\}. \tag{28.12}$$

So we have (using Theorem 27.6):

$$\begin{aligned} P\{\sup_{t \in T} \eta_t \ge a\} &\le P\Big(\bigcup_{n=1}^{\infty} \{\max_{t \in T_n} \eta_t \ge a - \varepsilon\}\Big) = \lim_{n \to \infty} P\{\max_{t \in T_n} \eta_t \ge a - \varepsilon\} \\ &\le \frac{\lim_{n \to \infty} \max_{t \in T_n} E\eta_t}{a - \varepsilon} = \frac{\sup_{t \in T} E\eta_t}{a - \varepsilon}. \end{aligned} \tag{28.13}$$

Since the positive ε is arbitrary, we come to (28.9).

Theorem 28.4. *Let η_t, $t \in T$, be a (sub)martingale with sample functions that are continuous on the right; let τ be a stopping time that is bounded by some $t_{\max} \in T$: $\tau \le t_{\max}$. Suppose all $\eta_t(\omega)$ are bounded in absolute value by the same random variable $\zeta(\omega)$, $E\zeta < \infty$.*

Then

$$E\eta_\tau = E\eta_{t_{\max}} \qquad (\text{or} \quad \le E\eta_{t_{\max}} \text{ if } \eta_t \text{ is a } \textit{sub}\text{martingale}). \tag{28.14}$$

Note that, while in Theorem 28.3 it is all the same for us whether the sample function is continuous on the right or on the left, here it is important: with right continuity this is true, and with left continuity it is not, generally (the same as with the strong Markov property).

Proof. For definiteness, let $T = [0, \infty)$. For a positive h, let us define the random time τ^h by

$$\tau^h = \begin{cases} kh & \text{for } (k-1)h < \tau \le kh < t_{\max}, \\ t_{\max} & \text{for } (n-1)h < \tau \le t_{\max} \le nh \end{cases} \tag{28.15}$$

($\tau^h = 0$ if $\tau = 0$). This τ^h is a measurable function of τ, and $\tau^h \ge \tau$, so the random variable τ^h is also a stopping time; it takes only finitely many values (not more than $t_{\max}/h + 1$) and is $\le t_{\max}$, so we can apply Theorem 27.3 or 27.4:

$$E\eta_{\tau^h} = E\eta_{t_{\max}} \qquad (\text{or} \quad \le E\eta_{t_{\max}}). \tag{28.16}$$

By the dominated-convergence theorem we have:

$$E\eta_\tau = E \lim_{h \to 0^+} \eta_{\tau_h} = \lim_{h \to 0^+} E\eta_{\tau_h} = E\eta_{t_{\max}} \qquad (\text{or} \quad \le E\eta_{t_{\max}}). \tag{28.17}$$

In fact, Theorem 28.4 remains true without supposing that all random variables η_t, $t \in T$, are dominated by the same random variable with finite expectation: $|\eta_t(\omega)| \le \zeta(\omega)$ for all t, ω, where $E\zeta < \infty$; but the proof is much more complicated – and we really will be using the result only in the situation of domination.

Let us apply the theory of martingales to diffusion processes; i. e. to Markov processes $\boldsymbol{\xi}_t$ in $\mathbb{R}^d$ with continuous trajectories, whose infinitesimal operator is given, for functions $f \in \mathbf{C}^2_{\text{unif.}}(\mathbb{R}^d)$, by

$$Af(\boldsymbol{x}) = Lf(\boldsymbol{x}) = \frac{1}{2} \sum_{i,\, j=1}^{d} a_{ij}(\boldsymbol{x}) \cdot \frac{\partial^2 f}{\partial x^i \partial x^j}(\boldsymbol{x}) + \sum_{i=1}^{d} b_i(\boldsymbol{x}) \cdot \frac{\partial f}{\partial x^i}(\boldsymbol{x}) \tag{28.18}$$

(with the coefficients satisfying the requirements of Lecture 21).

Let us speak about linear partial differential equations of the elliptic type: equations of the form

$$\mathcal{L}u(\boldsymbol{x}) = \frac{1}{2} \sum_{i,\, j=1}^{d} a_{ij}(\boldsymbol{x}) \cdot \frac{\partial^2 u}{\partial x^i \partial x^j}(\boldsymbol{x}) + \sum_{i=1}^{d} b_i(\boldsymbol{x}) \cdot \frac{\partial u}{\partial x^i}(\boldsymbol{x}) + c(\boldsymbol{x}) \cdot u(\boldsymbol{x}) = g(\boldsymbol{x}), \tag{28.19}$$

where the matrix $\big(a_{ij}(\boldsymbol{x})\big)$ is, for every $\boldsymbol{x}$, positive definite.

For elliptic equations we consider *boundary-value problems.*

The *Dirichlet problem* for the equation (28.19) is formulated as follows: given a function $g(\boldsymbol{x})$ in a region $G \subseteq \mathbb{R}^d$ and a function $\varphi(\boldsymbol{x})$ on its boundary ∂G (make a picture of G and its boundary), find a continuous function $u(\boldsymbol{x})$, $\boldsymbol{x} \in G \cup \partial G$, that is twice continuously differentiable in G, satisfies equation (28.19) for $\boldsymbol{x} \in G$ and satisfies the boundary condition

$$u(\boldsymbol{x}) = \varphi(\boldsymbol{x}), \qquad \boldsymbol{x} \in \partial G. \tag{28.20}$$

The same can be reformulated as follows: find a twice continuously differentiable function $u(\boldsymbol{x})$ in G such that

$$\begin{aligned} \mathcal{L}u(\boldsymbol{x}) &= g(\boldsymbol{x}), \qquad && \boldsymbol{x} \in G, \\ \lim_{\boldsymbol{y}\to\boldsymbol{x}} u(\boldsymbol{y}) &= \varphi(\boldsymbol{x}), \qquad && \boldsymbol{x} \in \partial G. \end{aligned} \tag{28.21}$$

For shortness, we just write the boundary condition as $u(\boldsymbol{x}) = \varphi(\boldsymbol{x})$, $\boldsymbol{x} \in \partial G$ (but meaning that the limit condition in (28.21) is satisfied, or that the function $u(\boldsymbol{x})$ is continuous in the closed region $G \cup \partial G$).

In the case of the dimension $d = 1$, elliptic equations are second-order ordinary differential equations; and the Dirichlet problem becomes a boundary-value problem for an ordinary differential equation in an interval, with two boundary conditions at the two ends of the interval.

Let us consider the case of equation without the term $c(\boldsymbol{x}) \cdot u(\boldsymbol{x})$. It turns out that the solution $u(\boldsymbol{x})$ of the Dirichlet problem

$$\begin{aligned} Lu(\boldsymbol{x}) &= g(\boldsymbol{x}), \qquad && \boldsymbol{x} \in G, \\ u(\boldsymbol{x}) &= \varphi(\boldsymbol{x}), \qquad && \boldsymbol{x} \in \partial G, \end{aligned} \tag{28.22}$$

(without the term $c(\boldsymbol{x}) \cdot u(\boldsymbol{x})$) has the representation

$$u(\boldsymbol{x}) = E_{\boldsymbol{x}}\Big(\varphi(\boldsymbol{\xi}_\tau) - \int_0^\tau g(\boldsymbol{\xi}_s)\, ds\Big), \tag{28.23}$$

where $\tau = \tau_G$ is the first time at which the process $\boldsymbol{\xi}_t$ leaves the region G:

$$\tau = \tau_G = \begin{cases} \min\{t\colon \boldsymbol{\xi}_t \notin G\} & \text{if there are such } t, \\ \infty & \text{otherwise.} \end{cases} \tag{28.24}$$

Particular cases:

1) $g(\boldsymbol{x}) \equiv 0$: the solution $u(\boldsymbol{x})$ of the Dirichlet problem

$$\begin{aligned} Lu(\boldsymbol{x}) &= 0, \qquad && \boldsymbol{x} \in G, \\ u(\boldsymbol{x}) &= \varphi(\boldsymbol{x}), \qquad && x \in \partial G, \end{aligned} \tag{28.25}$$

is represented as

$$u(\boldsymbol{x}) = E_{\boldsymbol{x}} \varphi(\boldsymbol{\xi}_\tau); \tag{28.26}$$

2) $g(\boldsymbol{x}) \equiv -1$, $\varphi(\boldsymbol{x}) \equiv 0$ for $\boldsymbol{x} \in \partial G$: the expectation

$$m(\boldsymbol{x}) = E_{\boldsymbol{x}} \tau_G \tag{28.27}$$

is the solution of the Dirichlet problem

$$\begin{aligned} Lm(\boldsymbol{x}) &= -1, \qquad \boldsymbol{x} \in G, \\ m(\boldsymbol{x}) &= 0, \qquad \boldsymbol{x} \in \partial G. \end{aligned} \tag{28.28}$$

However, this is not as simple as that: for (28.23), (28.26), or (28.28) to hold it is necessary that $E_{\boldsymbol{x}} \tau_G$ be finite; or at least that $\tau = \tau_G < \infty$ almost surely (otherwise $\boldsymbol{\xi}_\tau$ makes no sense). And we were speaking of *the* solution of the Dirichlet problem; but is there such a thing: do we know that this solution is *unique*? Also questions of *existence* of a solution arise. If $E_{\boldsymbol{x}} \tau_G < \infty$, and the functions g and φ are bounded, the expectation (28.23) does exist; but will the function $u(\boldsymbol{x})$ defined by this formula be a solution of (28.22): will it be *smooth* inside G? will it assume the boundary values $\varphi(\boldsymbol{x})$ at the boundary ∂G?

These questions are best addressed with some combination of usual partial-differential-equations methods and probabilistic methods based on martingales.

And it is clear that we need some order.

Theorem 28.5. *Let $m(\boldsymbol{x})$ be a bounded solution of the Dirichlet problem* (28.28). *Suppose that $m(\boldsymbol{x})$, defined, originally, only for $\boldsymbol{x} \in G \cup \partial G$, can be extended to the whole $\mathbb{R}^d$ as a twice continuously differentiable function $u(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^d$ ($u(\boldsymbol{x}) = m(\boldsymbol{x})$ for $\boldsymbol{x} \in G \cup \partial G$), that grows not faster than exponentially together with its first and second derivatives.*

Then $E_{\boldsymbol{x}} \tau_G < \infty$ for every $\boldsymbol{x} \in G$, and $m(\boldsymbol{x}) = E_{\boldsymbol{x}} \tau_G$.

Proof. As we know,

$$\eta_t = u(\boldsymbol{\xi}_t) - \int_0^t Lu(\boldsymbol{\xi}_s)\, ds \tag{28.29}$$

is a martingale. The random variable $\tau = \tau_G$ is a stopping time (the first time that a continuous stochastic process $\boldsymbol{\xi}_t$ hits the closed set $\mathbb{R}^d \setminus G$); and for every $t_* \in (0, \infty)$ the minimum $\min(\tau, t_*)$ is a *bounded* stopping time (bounded by the constant t_*). So by Theorem 28.4,

$$E_{\boldsymbol{x}} \eta_{\min(\tau_G,\, t_*)} = E_{\boldsymbol{x}} \eta_0; \tag{28.30}$$

or, by (28.29),

$$E_{\boldsymbol{x}} \Big(u(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) - \int_0^{\min(\tau,\, t_*)} Lu(\boldsymbol{\xi}_s)\, ds \Big) = E_{\boldsymbol{x}} u(\boldsymbol{\xi}_0) = u(\boldsymbol{x}). \tag{28.31}$$

We have $u(\boldsymbol{x}) = m(\boldsymbol{x})$ because $\boldsymbol{x} \in G$; since $\min(\tau, t_*) \le \tau$, we have $u(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) = m(\boldsymbol{\xi}_{\min(\tau,\, t_*)})$ and $Lu(\boldsymbol{\xi}_s) = Lm(\boldsymbol{\xi}_s) = -1$ for $s < \min(\tau, t_*)$; so

$$m(\boldsymbol{x}) = E_{\boldsymbol{x}} \big(m(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) + \min(\tau, t_*) \big), \tag{28.32}$$

$$E_{\boldsymbol{x}} \min(\tau,\, t_*) = m(\boldsymbol{x}) - E_{\boldsymbol{x}} m(\boldsymbol{\xi}_{\min(\tau,\, t_*)}). \tag{28.33}$$

It is clear that $\min(\tau,\, t_*)$ for every $\omega \in \Omega$ is a non-decreasing function of t_* (it is *increasing* up to $t_* = \tau_G$, and after that it remains constant; if $\tau_G = \infty$, this minimum is equal to t_*, and keeps on growing for all t_*). Clearly the limit

$$\lim_{t_* \to \infty} \min(\tau_G,\, t_*) = \tau_G. \tag{28.34}$$

By the monotone-convergence theorem we have:

$$E_{\boldsymbol{x}} \tau_G = \lim_{t_* \to \infty} E_{\boldsymbol{x}} \min(\tau_G,\, t_*) = m(\boldsymbol{x}) - \lim_{t_* \to \infty} E_{\boldsymbol{x}} m(\boldsymbol{\xi}_{\min(\tau,\, t_*)}). \tag{28.35}$$

The function m is bounded, so the limit (28.35) is finite, $E_{\boldsymbol{x}} \tau_G < \infty$.

Now, using the fact that $\tau_G < \infty$ almost surely, and the continuity of $\boldsymbol{\xi}_t$, we have:

$$\lim_{t_* \to \infty} \boldsymbol{\xi}_{\min(\tau,\, t_*)} = \boldsymbol{\xi}_\tau; \tag{28.36}$$

and using the fact that the function m is continuous in $G \cup \partial G$, we get:

$$\lim_{t_* \to \infty} m(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) = m(\boldsymbol{\xi}_\tau) = 0, \tag{28.37}$$

because $\boldsymbol{\xi}_\tau \in \partial G$.

The random variables $m(\boldsymbol{\xi}_{\min(\tau,\, t_*)})$ are all bounded by the same constant $\sup\{|m(\boldsymbol{x})|\colon \boldsymbol{x} \in G \cup \partial G\} < \infty$; so by the dominated-convergence theorem we have:

$$E_{\boldsymbol{x}} \tau_G = m(\boldsymbol{x}) - \lim_{t_* \to \infty} E_{\boldsymbol{x}} m(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) = m(\boldsymbol{x}) - E_{\boldsymbol{x}} \lim_{t_* \to \infty} m(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) = m(\boldsymbol{x}). \tag{28.38}$$

Theorem 28.6. *Let G be a region in $\mathbb{R}^d$ such that $E_{\boldsymbol{x}} \tau_G < \infty$. Let $g(\boldsymbol{x})$, $\boldsymbol{x} \in G$, and $\varphi(\boldsymbol{x})$, $\boldsymbol{x} \in \partial G$, be bounded functions. Suppose $u(\boldsymbol{x})$ is a bounded solution of the Dirichlet problem*

$$\begin{aligned} &Lu(\boldsymbol{x}) = \frac{1}{2} \sum_{i,\, j=1}^{d} a_{ij}(\boldsymbol{x}) \cdot \frac{\partial^2 u}{\partial x^i \partial x^j}(\boldsymbol{x}) + \sum_{i=1}^{d} b_i(\boldsymbol{x}) \cdot \frac{\partial u}{\partial x^i}(\boldsymbol{x}) = g(\boldsymbol{x}), \qquad \boldsymbol{x} \in G, \\ &\quad u(\boldsymbol{x}) = \varphi(\boldsymbol{x}), \qquad \boldsymbol{x} \in \partial G. \end{aligned} \tag{28.39}$$

Suppose $u(\boldsymbol{x})$ can be extended to the whole $\mathbb{R}^d$ as a twice continuously differentiable function, growing, together with its derivatives, not faster than exponentially.

Then for $x \in G \cup \partial G$

$$u(\boldsymbol{x}) = E_{\boldsymbol{x}} \Bigl(\varphi(\boldsymbol{\xi}_\tau) - \int_0^\tau g(\boldsymbol{\xi}_s)\, ds \Bigr). \tag{28.40}$$

Proof. Again take the martingale

$$\eta_t = u(\boldsymbol{\xi}_t) - \int_0^t Lu(\boldsymbol{\xi}_s)\, ds, \tag{28.41}$$

and again take it at the stopping time $\min(\tau,\, t_*)$:

$$
\begin{aligned}
&E_{\boldsymbol{x}}\Big(u(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) - \int_0^{\min(\tau,\, t_*)} Lu(\boldsymbol{\xi}_s)\ ds\Big) \\
&\qquad = E_{\boldsymbol{x}}\Big(u(\boldsymbol{\xi}_{\min(\tau,\, t_*)}) - \int_0^{\min(\tau,\, t_*)} g(\boldsymbol{\xi}_s)\ ds\Big) = E_{\boldsymbol{x}}\eta_0 = u(\boldsymbol{x}).
\end{aligned} \tag{28.42}
$$

The random variables under the expectation sign are for every t_* dominated by the random variable

$$
\sup_{\boldsymbol{x}\in G\cup\partial G} |u(\boldsymbol{x})| + \sup_{\boldsymbol{x}\in\partial G} |g(\boldsymbol{x})| \cdot \tau, \tag{28.43}
$$

which has a finite expectation; and by the dominated-convergence theorem, letting $t_* \to \infty$, we get

$$
u(\boldsymbol{x}) = E_{\boldsymbol{x}}\Big(u(\boldsymbol{\xi}_\tau) - \int_0^\tau g(\boldsymbol{\xi}_s)\ ds\Big) = E_{\boldsymbol{x}}\Big(\varphi(\boldsymbol{\xi}_\tau) - \int_0^\tau g(\boldsymbol{\xi}_s)\ ds\Big). \tag{28.44}
$$

In the next lecture we'll consider concrete applications of all this to the Wiener process.