Lecture note 27. More theory of martingales. Discrete time.

Microtheorem 27.1. *Let Y_t, $t \in T$, be a martingale with respect to a stochastic process X_t, $t \in T$. Let τ be a stopping time taking finitely many values $t_1 < t_2 < ... < t_n$* (all in T, not $+\infty$).

Then for an arbitrary $t_* \in T$

$$E(Y_\tau) = E(Y_{t_*}). \tag{27.1}$$

By the way, I did not mention this before, but the expectation (unconditional) of a martingale Y_t does not depend on t: for $t < s$ we have:

$$E(Y_s) = E\big(E(Y_s \| X_u, u \le t)\big) = E(Y_t). \tag{27.2}$$

Proof of Microtheorem 27.1. We have:

$$E(Y_\tau) = \sum_{i=1}^{n} E\big(I_{\{\tau = t_i\}}(\omega) \cdot Y_\tau\big) = \sum_{i=1}^{n} E\big(I_{\{\tau = t_i\}}(\omega) \cdot Y_{t_i}\big). \tag{27.3}$$

For $i > 1$ the event $\{\tau = t_i\}$ whose indicator is used in this formula is equal to the difference of two:

$$\{\tau = t_i\} = \{\tau \le t_i\} \setminus \{\tau \le t_{i-1}\}; \tag{27.4}$$

and for $i = 1$ it is just

$$\{\tau = t_1\} = \{\tau \le t_1\}. \tag{27.5}$$

So the i-th summand in (27.3) can be written as

$$E\big(\big(I_{\{\tau \le t_i\}}(\omega) - I_{\{\tau \le t_{i-1}\}}(\omega)\big) \cdot Y_{t_i}\big) \tag{27.6}$$

for $i > 1$, and as

$$E\big(I_{\{\tau \le t_1\}}(\omega) \cdot Y_{t_1}\big) \tag{27.7}$$

for $i = 1$. So we get:

$$\begin{aligned} E(Y_\tau) = {} & E\big(I_{\{\tau \le t_1\}}(\omega) \cdot Y_{t_1}\big) - E\big(I_{\{\tau \le t_1\}}(\omega) \cdot Y_{t_2}\big) + E\big(I_{\{\tau \le t_2\}}(\omega) \cdot Y_{t_2}\big) \\ & - E\big(I_{\{\tau \le t_2\}}(\omega) \cdot Y_{t_3}\big) + ... + E\big(I_{\{\tau \le t_{n-1}\}}(\omega) \cdot Y_{t_{n-1}}\big) \\ & - E\big(I_{\{\tau \le t_{n-1}\}}(\omega) \cdot Y_{t_n}\big) + E\big(I_{\{\tau \le t_n\}}(\omega) \cdot Y_{t_n}\big) \\ = {} & \sum_{i=1}^{n-1} E\big(I_{\{\tau \le t_i\}}(\omega) \cdot (Y_{t_i} - Y_{t_{i+1}})\big) + E\big(I_{\{\tau \le t_n\}}(\omega) \cdot Y_{t_n}\big). \end{aligned} \tag{27.8}$$

We represent the i-th summand ($i < n$) as the expectation of a conditional expectation:

$$E\big(I_{\{\tau \le t_i\}}(\omega) \cdot (Y_{t_{i+1}} - Y_{t_i})\big) = E\big(E\big(I_{\{\tau \le t_i\}}(\omega) \cdot (Y_{t_i} - Y_{t_{i+1}}) \big\| X_u,\, u \le t_i\big)\big). \tag{27.9}$$

The event $\{\tau \leq t_i\}$ is represented as $\{(X_u,\, u \leq t_i) \in C\}$ (i. e., is determined by the values of X_u with $u \leq t_i$), so its indicator can be taken from under the sign of the conditional expectation:

$$E\big(I_{\{\tau \leq t_i\}}(\omega) \cdot (Y_{t_i} - Y_{t_{i+1}})\big) = E\big(I_{\{\tau \leq t_i\}}(\omega) \cdot E(Y_{t_i} - Y_{t_{i+1}} \| X_u,\, u \leq t_i)\big); \qquad (27.10)$$

and by the martingale property, the conditional expectation is equal to 0.

The last summand in (27.8) is equal to $E(Y_{t_n})$, because the event $\{\tau \leq t_n\}$ is the whole Ω, and its indicator is identically 1. By the way, $E(Y_{t_n})$ is equal to $E(Y_{t_*})$ for any other $t_* \in T$.

The same way we prove

Microtheorem 27.2. *Let Y_t, $t \in T$, be a submartingale with respect to a stochastic process X_t, $t \in T$. Let τ be a stopping time taking finitely many values $t_1 < t_2 < \ldots < t_n = t_{\max}$.*

Then

$$E(Y_\tau) \leq E(Y_{t_{\max}}) \qquad (27.11)$$

(for a *sub*martingale, $E(Y_t)$ is no longer a constant, but rather a nondecreasing function).

Proof. Formulas (27.8), (27.9), (27.10) still hold; but instead of saying "by the martingale property the conditional expectation is equal to 0" we say "by the submartingale property the conditional expectation is ≤ 0".

For random times τ that are not stopping times, the statements of Microtheorems 27.1, 27.2 are not necessarily true. E. g., if $X_0 \equiv 0$, and X_1 is a random variables with density $p(x)$ with $E(X_1) = 0$, the sequence of two random variables X_0, X_1 is a martingale. Let us define the random time τ by

$$\tau = I_{[0,\,\infty)}(X_1). \qquad (27.12)$$

This is *not* a stopping time: to know whether $\tau \leq 0$ (which means $= 0$ in this case), it is not enough to watch X_0 (as if there were anything to watch here!), but we need to know X_1. We have:

$$E(X_\tau) = E(I_{\{\tau=0\}} \cdot X_0 + I_{\{\tau=1\}} \cdot X_1) = E\big(I_{[0,\,\infty)}(X_1) \cdot X_1\big) = \int_0^\infty x \cdot p(x)\, dx > 0, \quad (27.13)$$

while $E(X_1) = \displaystyle\int_{-\infty}^{\infty} x \cdot p(x)\, dx = 0$.

Let Y_t, $t \in T$, be a random function; τ a random variable taking values in $T \cup \{\infty\}$. The random function $\hat{Y}_t$, $t \in T$, obtained by stopping Y_t at the time τ is defined as

$$\hat{Y}_t = \begin{cases} Y_t, & t < \tau, \\ Y_\tau, & t \geq \tau. \end{cases} \qquad (27.14)$$

Microtheorem 27.3. *If Y_t, $t \in T$, is a martingale (a submartingale), and τ a stopping time taking finitely many values, the stopped random function $\hat{Y}_t$, $t \in T$, is also a martingale (a submartingale).*

Proof. We have to prove that for $t < s$

$$E(\hat{Y}_s - \hat{Y}_t \| X_u,\, u \le t) = 0 \qquad \text{(or } \ \ge 0). \tag{27.15}$$

Without restriction of generality, we can assume that t and s are among the values t_i that the stopping time τ takes: $t = t_m$, $s = t_k$. If not, we can add t and s to these values, and state that $\tau = t$ or s only with probability 0.

So we have to prove that

$$E(\hat{Y}_{t_k} - \hat{Y}_{t_m} \| X_u,\, u \le t) = \sum_{i=m+1}^{k} E(\hat{Y}_{t_i} - \hat{Y}_{t_{i-1}} \| X_u,\, u \le t) = 0 \qquad \text{(or } \ \ge 0). \tag{27.16}$$

It's enough to prove that every summand is $= 0$ (or ≥ 0). Using the property 3) of conditional expectations, we get:

$$E(\hat{Y}_{t_i} - \hat{Y}_{t_{i-1}} \| X_u,\, u \le t) = E\big(E(\hat{Y}_{t_i} - \hat{Y}_{t_{i-1}} \| X_u,\, u \le t_{i-1}) \big\| X_u,\, u \le t\big). \tag{27.17}$$

Again, it is enough to prove that

$$E(\hat{Y}_{t_i} - \hat{Y}_{t_{i-1}} \| X_u,\, u \le t_{i-1}) = 0 \qquad \text{(or } \ \ge 0). \tag{27.18}$$

Formula (27.18) means that for every set C in the space of functions x_u, $u \le t_{i-1}$,

$$E\big(I_C(X_u,\, u \le t_{i-1}) \cdot (\hat{Y}_{t_i} - \hat{Y}_{t_{i-1}})\big) = 0 \qquad \text{(or } \ \ge 0). \tag{27.19}$$

We haven't used the fact that τ is a stopping time. And this means, in particular, that there exists a set D in the space of functions x_u, $u \le t_{i-1}$, such that

$$\{\tau \le t_{i-1}\} = \{(X_u,\, u \le t_{i-1}) \in D\} \tag{27.20}$$

(I seem again to run out of letters). The opposite event

$$\{\tau > t_{i-1}\} = \{(X_u,\, u \le t_{i-1}) \in D^c\}, \tag{27.21}$$

where D^c is the complement to the set D in the space of functions x_u, $u \le t_{i-1}$.

The expectation (27.19) can be represented as

$$\begin{aligned} &E\big(I_C(X_u,\, u \le t_{i-1}) \cdot I_D(X_u,\, u \le t_{i-1}) \cdot (\hat{Y}_{t_i} - \hat{Y}_{t_{i-1}})\big) \\ &\qquad + E\big(I_C(X_u,\, u \le t_{i-1}) \cdot I_{D^c}(X_u,\, u \le t_{i-1}) \cdot (\hat{Y}_{t_i} - \hat{Y}_{t_{i-1}})\big). \end{aligned} \tag{27.22}$$

In the first summand the second indicator is the indicator of the event $\{\tau \le t_{i-1}\}$, so in the first summand $\hat{Y}_{t_i} = \hat{Y}_{t_{i-1}} = Y_\tau$, and this first summand is equal to 0. In the second summand we have $\tau > t_{i-1}$, $\tau \ge t_i$, $\hat{Y}_{t_{i-1}} = Y_{t_{i-1}}$, $\hat{Y}_{t_i} = Y_{t_i}$, and the second summand is equal to

$$E\big(I_{C \cap D^c}(X_u,\, u \le t_{i-1}) \cdot (Y_{t_i} - Y_{t_{i-1}})\big), \tag{27.23}$$

and by the martingale (or the submartingale) property this expectation is $= 0$ (or ≥ 0).

Microtheorem 27.4 (the Kolmogorov-type inequality). *If* Y_t, $t \in T$, *is a nonnegative submartingale* (with respect to X_t, $t \in T$), *and* $t_1 < t_2 < ... < t_n = t_{\max}$ *are points in* T, *then for every* $a > 0$

$$P\{\max(Y_{t_1}, Y_{t_2}, ..., Y_{t_n}) \geq a\} \leq \frac{E(Y_{t_{\max}})}{a}. \qquad (27.24)$$

This is a strengthening of the following Chebyshev-type inequality:

$$P\{Y_{t_i} \geq a\} \leq \frac{E(Y_{t_i})}{a} \leq \frac{E(Y_{t_{\max}})}{a}, \qquad (27.25)$$

only the Chebyshev-type inequality is about *one* random variable, and (27.24) is about the maximum of an arbitrarily large number of random variables.

What Kolmogorov started this with was a particular case:

Microtheorem 27.5. *Let* X_1, X_2, ..., X_n *be independent random variables with* $E(X_i) = 0$ *and finite variances* $\mathrm{Var}(X_i) = \sigma_i^2$. *Then for every* $b > 0$

$$P\{\max(|X_1|, |X_1 + X_2|, ..., |X_1 + X_2 + ... + X_n|) \geq b\} \leq \frac{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}{b^2}. \qquad (27.26)$$

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

$$P\{|X_1 + ... + X_i| \geq b\} \leq \frac{\mathrm{Var}(X_1 + ... + X_i)}{b^2} = \frac{\sigma_1^2 + ... + \sigma_i^2}{b^2} \leq \frac{\sigma_1^2 + ... + \sigma_n^2}{b^2}. \qquad (27.27)$$

Let me show how the classical Kolmogorov's inequality (27.26) is a particular case of the Kolmogorov-type inequality (27.24). The sequence $Y_i = X_1 + ... + X_i$, $i = 1, ..., n$, is a martingale with respect to the sequence of the random variables X_i (see Example 23.2, generalization). By Example 24.1, the sequence $Z_i = (X_1 + ... + X_i)^2$, $i = 1, ..., n$, is a *sub*martingale; and by inequality (27.24) we have:

$$P\{\max(|X_1|, |X_1 + X_2|, ..., |X_1 + ... + X_n|) \geq b\} = P\{\max(Z_1, Z_2, ..., Z_n) \geq b^2\} \leq \frac{E(Z_n^2)}{b^2}, \qquad (27.28)$$

which is equal to the right-hand side of (27.26).

Proof of Microtheorem 27.4. Let us introduce the random time

$$\tau = \begin{cases} \min\{t_i \colon 1 \leq i \leq n, Y_{t_i} \geq a\} & \text{if there is at least one such } t_i, \\ \infty & \text{if there is no such } t_i. \end{cases} \qquad (27.29)$$

Similarly to Example 25.3, τ is a stopping time. By Example 26.2, $\min(\tau, t_{\max})$ is also a stopping time – and this one, taking finitely many finite values.

The event $\{\max(Y_{t_1}, Y_{t_2}, ..., Y_{t_n}) \geq a\}$ clearly coincides with the event

$$\{Y_{\min(\tau,\, t_{\max})} \geq a\}. \qquad (27.30)$$

Applying a Chebyshev-type inequality to the nonnegative random variable Y_τ, we get:

$$P\{\max(Y_{t_1}, Y_{t_2}, ..., Y_{t_n}) \geq a\} = P\{Y_{\min(\tau,\, t_{\max})} \geq a\} \leq \frac{E(Y_{\min(\tau,\, t_{\max})})}{a}, \qquad (27.31)$$

and by Microtheorem 27.2 this is $\leq \dfrac{E(Y_{t_{\max}})}{a}$.