RAP 412 14.01.2009

Wykład 11: Martyngały: definicja, twierdzenia o zbieżności

Wykładowca: Andrzej Ruciński *Pisarz:Mirosława Jańczak*

1 Wstęp

Do tej pory zajmowaliśmy się ciągami zmiennych losowych (X_n) o pewnej strukturze zależności. Ciąg zmiennych losowych niezależnych stanowi tu szczególny przypadek. Badaliśmy pewne własności łańcuchów Markowa, tj. takich ciągów zmiennych losowych, dla których kolejny wyraz zależał wyłącznie od poprzedniego (tzw. własność zaniku pamięci). Kolejny wykład ma na celu przestawienie nowej struktury zależności dla ciągów zmiennych losowych, jaką określają martyngały.

Pierwsza część wykładu poświęcona jest przypomnieniu wiadomości na temat warunkowej wartości oczekiwanej. Zawiera ona definicję oraz kilka podstawowych twierdzeń. W kolejnej części wprowadzone zostaje pojęcie martyngału. Przedstawiono tu również kilka przykładów. Ostatnia, trzecia część wykładu dotyczy twierdzeń o zbieżności martyngałów.

2 Warunkowa wartość oczekiwana

2.1 Przypadek jednowymiarowy

Rozważania dotyczące warunkowej wartości oczekiwanej ograniczone zostaną do zmiennych losowych dyskretnych.

Definicja 1. *Warunkową wartością oczekiwaną zmiennej losowej Y pod warunkiem $X = x$ nazywamy liczbę $\psi(x)$ określoną następująco:* $\psi(x) = E(Y|X = x) = \sum_y yP(Y = y|X = x)$

Zauważmy, że $\psi(X)$ jest pewną funkcją zmiennej losowej X. Zatem $\psi(X)$ jest również zmienną losową. Oznaczamy ją $\psi(X) = E(Y|X)$. Wartość oczekiwana tak określonej zmiennej losowej jest równa $E(Y)$.

Twierdzenie 1. $E(\psi(X)) = E(Y)$

Dowód.

$$E(\psi(X)) = \sum_x \psi(x)p_X(x) = \sum_x \sum_y yp_{Y|X}(y|x)p_X(x) =$$
$$= \sum_x \sum_y yp_{X,Y}(x, y) = \sum_y y \sum_x p_{X,Y}(x, y) = \sum_y yp_Y(y) = E(Y).$$

□

Powyższe twierdzenie można uogólnić:

Twierdzenie 2. $E(\psi(X)g(X)) = E(Yg(X))$

Dowód tego faktu przebiega analogicznie.

2.2 Przypadek wielowymiarowy

Zajmiemy się teraz przypadkiem, gdy zmienna losowa $\mathbf{X}$ jest postaci $\mathbf{X} = (X_1, X_2, \ldots, X_n)$.

Lemat 1. *Zachodzą następujące własności:*
a) $E(Y_1 + Y_2|\mathbf{X}) = E(Y_1|\mathbf{X}) + E(Y_2|\mathbf{X})$,
b) $E(Yg(\mathbf{X})) = g(\mathbf{X})E(Y|\mathbf{X})$,
c) Jeśli h *jest funkcją różnowartościową, to* $E(Y|h(\mathbf{X})) = E(Y|\mathbf{X})$.

Lemat 2. (Własność wieżowa.)

$$E[E(Y|\mathbf{X_1}, \mathbf{X_2})|\mathbf{X_1}] = E(Y|\mathbf{X_1}).$$

Zdefiniujemy teraz warunkową wartość oczekiwaną zmiennej losowej Y względem zdarzenia A.

Definicja 2. *Niech* $(\Omega, \mathcal{F}, \mathbb{P})$ *będzie przestrzenią probabilistyczną, na której określone są zmienne losowe* X *oraz* Y. *Warunkową wartością oczekiwaną zmiennej losowej* Y *względem zdarzenia* $A \in \mathcal{F}$ *nazywamy liczbę*

$$E(Y|A) = \sum_y y\mathbb{P}(Y = y|A),$$

gdzie $\mathbb{P}(Y = y|A) = \mathbb{P}(Y = y|\{X(\omega) : \omega \in A\})$.

Mamy

$$E(Y|I_A) = \begin{cases} E(Y|A) \text{ gdy } \omega \in A \\ E(Y|A^c) \text{ gdy } \omega \notin A. \end{cases}$$

Zatem $E(Y|I_A)$ jest zmienną losową dwupunktową.
Ponadto zachodzi zależność:

$$E(I_B|A) = \mathbb{P}(B|A).$$

Lemat 3. *Niech* A *będzie zdarzeniem oraz niech* $A = \bigcup_{i=1}^n B_i$, *gdzie* $B_i \in \mathcal{F}$ *są zdarzeniami parami rozłącznymi. Wówczas*

$$E(Y|A)\mathbb{P}(A) = \sum_{i=1}^n E(Y|B_i)\mathbb{P}(B_i).$$

W szczególności, gdy $A = \Omega$ dostajemy uogólnienie wzoru na prawdopodobieństwo całkowite:

$$E(Y) = \sum_{i=1}^n E(Y|B_i)\mathbb{P}(B_i).$$

Podstawiając w powyższym wzorze $Y = I_C$ otrzymujemy znany wzór:

$$\mathbb{P}(C) = E(I_C) = \sum_{i=1}^n E(I_C|B_i)\mathbb{P}(B_i) = \sum_{i=1}^n \mathbb{P}(C|B_i)\mathbb{P}(B_i).$$

Dowód. Z części b) lematu (1) dostajemy:

$$E(YI_A) = E(Y|A)\mathbb{P}(A).$$

Mamy zatem

$$E(YI_A) = E(Y\sum_i I_{B_i}) = \sum_i E(YI_{B_i}) = \sum_i E(Y|B_i)\mathbb{P}(B_i).$$

□

Definicja 3. *Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie przestrzenią probabilistyczną, $\mathcal{G} \subseteq \mathcal{F}$ pod σ-ciałem σ-ciała $\mathcal{F}$. Niech ponadto $Y : \Omega \to \mathbb{R}$ będzie zmienną losową taką, że $EY^2 < \infty$. $\mathcal{G}$ mierzalną zmienną losową Z nazywamy warunkową wartością oczekiwaną względem $\mathcal{G}$, jeżeli*

$$E((Y - Z)I_G) = 0$$

dla każdego $G \in \mathcal{G}$. Oznaczamy ją $Z = E(Y|\mathcal{G})$.

3 Martyngały, definicja i przykłady

W tej części wykładu wprowadzona zostanie definicja martyngału oraz podanych będzie kilka przykładów.

Definicja 4. *Ciąg (S_n) zmiennych losowych (skończony lub nie) jest martyngałem względem ciągu (X_n), jeżeli*
a) $E|S_n| < \infty$,
b) $E(S_{n+1}|X_1, X_2, \dots, X_n) = S_n$.

Martyngał określa zatem tzw. grę sprawiedliwą w takim sensie, że średnia wygrana w chwili $n+1$, gdy znany jest przebieg gry do chwili n, jest równa S_n, czyli łącznej wygranej w chwili n.

Często definiuje się $S_n = X_n$ lub określa się S_n jako pewną funkcją X_n, tj. $S_n = \phi(X_n)$ (z definicji natomiast mamy $S_n = \phi(X_1, X_2, \dots, X_n)$.

Przykład 1. *Niech $X_1, X_2, \dots$ będą zmiennymi losowymi niezależnymi takimi, że $E(X_i) = 0$ oraz $E|X_i| < \infty$. Zdefiniujmy*

$$S_n = X_1 + X_2 + \cdots + X_n.$$

Zachodzi zależność

$$E(S_{n+1}|X_1, X_2, \dots, X_n) = E(S_n|X_1, X_2, \dots, X_n) + E(X_{n+1}|X_1, X_2, \dots, X_n) = S_n + E(X_{n+1}) = S_n.$$

(S_n) jest zatem martyngałem względem ciągu (X_n).

Przykład 2. *Do warunków z poprzedniego przykładu dodajmy $Var(X_i) < \infty$. Niech $T_n = S_n^2$. Mamy*

$$E(T_{n+1}|X_1, X_2, \dots, X_n) = T_n + 2E(X_{n+1})E(S_n|X_1, X_2, \dots, X_n) + E(X_{n+1}^2) \geq T_n.$$

T_n nie jest zatem martyngałem względem ciągu (X_n). Gdy spełniony jest taki rodzaj zależności, mówimy, że T_n jest supermartyngałem względem ciągu (X_n).

Przykład 3. *Rozważmy prosty spacer losowy, dla którego* $\mathbb{P}(X_n = 1) = p$*,* $\mathbb{P}(X_n = -1) = q$*. Niech*

$$S_0 = 0 \text{ oraz } S_n = \sum_{i=1}^{n} X_i.$$

Sprawdźmy, czy S_n *jest martyngałem względem* X_n*. Mamy* $|S_n| \leq n$*. Stąd* $E|S_n| < \infty$*. Ponadto*

$$E(S_{n+1}|X_1, X_2, \ldots, X_n) = S_n + p - q.$$

Zatem S_n *nie jest martyngałem. Zdefiniujmy ciąg* (Y_n) *następująco:*

$$Y_n = S_n - E(S_n) = S_n - n(p - q).$$

Wtedy $E|Y_n| < \infty$ *oraz*

$$E(Y_{n+1}|X_1, X_2, \ldots, X_n) = S_n + p - q - (n+1)(p - q) = S_n - n(p - q) = Y_n,$$

czyli (Y_n) *jest martyngałem względem ciągu* (X_n)*.*

Przykład 4. *Rozważmy pewną grę. Niech* S_0 *oznacza kapitał początkowy,* S_n *kapitał po* n *grach. Gra jest sprawiedliwa w potocznym sensie, gdy*

$$E(S_{n+1}|S_0, S_1, \ldots, S_n) = S_n,$$

czyli ciąg (S_n) *jest martyngałem względem samego siebie. Załóżmy, że gracz stosuje strategię podwajania stawki po każdej przegranej. Gra natomiast do czasu uzyskania pierwszego sukcesu. Z prawdopodobieństwem równym* 1 *strategia ta przynosi sukces, tzn. gracz zyskuje dokładnie* 1*. Zastanowimy się teraz jaki powinien być kapitał początkowy gracza, aby odniósł on sukces z prawdopodobieństwem* 1*. Niech* L *będzie zmienną losową określającą łączną przegraną do momentu pierwszej wygranej,* N *natomiast niech oznacza liczbę odbytych gier.* N *jest zmienną losową o rozkładzie geometrycznym z parametrem* $1/2$ *(Wtedy* $E(n) = 2$*.) Mamy:*

$$E(L) = E(E(L|N)) = \sum_{n=2}^{\infty} \left(\frac{1}{2}\right)^{n-1} \cdot \frac{1}{2}(1 + 2 + \cdots + 2^{n-2}) = \infty.$$

Zatem, aby wygrać z prawdopodobieństwem równym jeden należy przyjść z nieskończenie wielkim kapitałem początkowym. Strategia podwajania stawki określa historycznie pierwszy martyngał.

Przykład 5. *Określimy teraz dwa martyngały na bazie procesu gałązkowego, gdzie* $Z_0 = 1$ *oznacza pierwszego przodka oraz* Z_n *dla* $n \geq 1$ *są liczebnościami kolejnych pokoleń.*
a) Mamy

$$E(Z_{n+1}|Z_n = z_n) = z_n \cdot \mu,$$

gdzie $\mu = E(Z_0)$*. Stąd*

$$E(Z_{n+1}|Z_1, Z_2, \ldots, Z_n) = E(Z_{n+1}|Z_n) = \mu Z_n.$$

Ponadto $E(Z_n) = \mu^n$*. Zdefiniujmy*

$$W_n = \frac{Z_n}{\mu^n}.$$

Wówczas
$$E(W_{n+1}|Z_1, Z_2, \ldots, Z_n) = \frac{\mu Z_n}{\mu^n} = W_n$$
i (W_n) *jest martyngałem względem* (Z_n).
b) Niech η *będzie prawdopodobieństwem wyginięcia procesu gałązkowego. Wtedy ciąg* (V_n) *określony następująco:*
$$V_n = \eta^{Z_n}$$
jest martyngałem względem (Z_n).

4 Twierdzenia o zbieżności martyngałów

Na początku tej części wykładu przypomnimy definicje kilku typów zbieżności zmiennych losowych.

Definicja 5. *Mówimy, że* X_n *dąży do* X *z prawdopodobieństwem jeden, jeżeli*
$$\mathbb{P}(\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1.$$
Zapisujemy $X_n \xrightarrow{a.s.} X$.

Definicja 6. X_n *zbiega do* X *według* r*-tego momentu, gdy* $E(X_n)^r < \infty$ *oraz*
$$\lim_{n\to\infty} E(|X_n - X|^r) = 0.$$
Stosujemy oznaczenie $X_n \xrightarrow{r} X$.

Fakt 1. *Jeżeli* $r \geq s$, *to*
$$X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{a.s.} X.$$

Twierdzenie 3. (O zbieżności martyngałów.)
Niech (S_n) *będzie martyngałem oraz* $E(S_n^2) < \infty$. *Istnieje wówczas zmienna losowa* S *taka, że*
$$S_n \xrightarrow{a.s.} S$$
oraz
$$S_n \xrightarrow{2} S.$$

Wniosek 1. (Mocne prawo wielkich liczb)
Niech $X_1, X_2, \ldots, X_n$ *będą zmiennymi losowymi niezależnymi o takich samych rozkładach. Niech* $S_n = X_1 + X_2 + \cdots + X_n$. *Wtedy*
$$\exists \mu \in \mathbb{R} : \frac{S_n}{n} \xrightarrow{a.s.} \mu \Leftrightarrow E|X_1| < \infty.$$
Wówczas $\mu = E(X_1)$.

Dowód. Załóżmy bez straty ogólności, że $\mu = 0$. Zdefiniujmy zmienną losową

$$S'_n = \sum_{i=1}^{n} \frac{X_i}{i}.$$

Wtedy

$$S'_{n+1} = S'_n + \frac{X_{n+1}}{n+1}.$$

(S'_n) jest martyngałem względem (X_n). Z twierdzenia (3) wiemy, że istnieje zmienna losowa S taka, że $S'_n \xrightarrow{a.s.} S$. Stąd dostajemy, że $\sum_{i=1}^{\infty} \frac{X_i}{i} < \infty$ z prawdopodobieństwem jeden. Z lematu Kroneckera otrzymujemy zatem

$$\sum_{i=1}^{n} \frac{X_i}{i} = \frac{S_n}{n} \xrightarrow{a.s.} 0.$$

$\square$

Lemat 4. Kroneckera.
Jeżeli (b_n) jest ciągiem monotonicznie rosnącym do nieskończoności oraz $\sum_{i=1}^{\infty} \frac{a_i}{b_i} < \infty$, to

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} a_i}{b_n} = 0.$$

W twierdzeniu powyżej stosujemy lemat podstawiając: $b_n := n$, $a_i := X_i$.

Twierdzenie 4. (Nierówność Dooba-Kołmogorowa)
Jeżeli (S_n) jest martyngałem względem (X_n), to

$$\mathbb{P}(\max_{1 \le i \le n} |S_i| \ge \varepsilon) \le \frac{E(S_n^2)}{\varepsilon^2}.$$