Rachunek prawdopodobieństwa 2R

Kycia

Spis rozmaitości treściowalnych

Wykład 06.10.23 : Warunkowa wartość oczekiwana

1.1 Prawdopodobieństwo warunkowe

Tak jak zwykle do tej pory, pracować będziemy na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$.

Przypomnijmy definicję **prawdopodobieństwa warunkowego** z Rachunku Prawdopodobieństwa 1 (i z liceum). Dla zdarzenia $A \in \mathcal{F}$ takiego, że $\mathbb{P}[A] \in (0,1)$ definiujemy prawdopodobieństwo warunkowe jako

$$\mathbb{P}\left[B \mid A\right] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]}.$$

Wartość ta informuje nas o zajściu B wtedy, gdy jesteśmy pewni, że A zaszło. Ale co, jeśli nasza wiedza dotycząca A jest mniej pewna? To znaczy, *co jeśli* $\mathbb{P}[A] = 0$? Dość naturalne wydaje się rozważenie zdarzenia przeciwnego i zsumowania obu prawdopodobieństw:

$$\mathbb{1}_A \mathbb{P}\left[B \mid A\right] + \mathbb{1}_{A^c} \mathbb{P}\left[B \mid A^c\right].$$

Zauważmy od razu, że wyrażenie $\mathbb{1}_A \mathbb{P}\left[B \mid A\right]$ jest zmienną losową.

W przypadku, gdy mamy dwa zbiory, $A_1, A_2 \in \mathcal{F}$, i chcemy zbadać $\mathbb{P}\left[B \mid A_1 \cap A_2\right]$ możemy powyższe rozumowanie rozszerzyć na wszystkie możliwe kombinacje A_1, A_2 i ich dopełnień:

$$\mathbb{1}_{A_1 \cap A_2} \mathbb{P}\left[B \mid A_1 \cap A_2\right] + \mathbb{1}_{A_1 \cap A_2^c} \mathbb{P}\left[B \mid A_1 \cap A_2^c\right] + \mathbb{1}_{A_1^c \cap A_2} \mathbb{P}\left[B \mid A_1^c \cap A_2\right] + \mathbb{1}_{A_1^c \cap A_2^c} \mathbb{P}\left[B \mid A_1^c \cap A_2^c\right].$$

Działanie jak wyżej daje pełną informacje o każdym zdarzeniu z ciała generowanego przez zdarzenia A_1 i A_2. Nazywamy je **rozbiciem** względem σ-ciała generowanego przez A_1 i A_2.

Analogicznie możemy zdefiniować $\mathbb{E}\left[X \mid A\right]$ dla całkowalnej zmiennej losowej X (tzn. $\mathbb{E}\left[|X|\right] < \infty$):

$$\mathbb{E}\left[X \mid A\right] = \int_\Omega X(\omega) \mathbb{P}\left[d\omega \mid A\right] = \frac{1}{\mathbb{P}[A]} \mathbb{E}[X \mathbb{1}_A],$$

gdzie całka wyżej tłumaczy się na całkę po X względem miary $\mathbb{P}\left[B \mid A\right]$.

Uzasadnimy, dlaczego wzór wyżej jest zasadną definicją prawdopodobieństwa warunkowego przy ograniczonej wiedzy o zdarzeniu A.

1.2 Konstrukcja warunkowej wartości oczekiwanej

Zanim zdefiniujemy **warunkową wartość oczekiwaną [wwo]** zmiennej losowej X, zaczniemy od przyjrzenia się bliżej motywacji i konstrukcji stojącej za tym pojęciem.

Niech Z będzie całkowalną zmienną losową przyjmującą przeliczalnie wiele wartości. Zdefiniujmy funkcję

$$h(z) = \begin{cases} \mathbb{E}\left[X \mid Z = z\right] & \mathbb{P}[Z = z] > 0 \\ 0 & \text{wpp} \end{cases}$$

oraz zmienną losową $Y = h(Z)$. Weźmy dowolny $C \in \mathrm{Bor}(\mathbb{R})$ i zbadajmy $\mathbb{E}\left[Y \mathbb{1}_{\{Z \in C\}}\right]$. Zaczniemy od skorzystania z faktu, że Z przyjmuje przeliczalnie wiele wartości, więc możemy zapisać sumę po nich wszystkich

$$\begin{aligned}
\mathbb{E}\left[Y\mathbb{1}_{\{Z\in C\}}\right] &= \sum_{z\in C} h(z)\mathbb{P}[Z=z] = \\
&\overset{\star}{=} \sum_{z\in C} \mathbb{E}\left[X\mathbb{1}_{\{Z=z\}}\right] \frac{1}{\mathbb{P}[Z=z]}\mathbb{P}[Z=z] = \\
&= \sum_{z\in C} \mathbb{E}\left[X\mathbb{1}_{\{Z=z\}}\right] = \\
&= \mathbb{E}\left[\sum_{z\in C} X\mathbb{1}_{\{Z=z\}}\right] = \\
&= \mathbb{E}\left[X\mathbb{1}_{\{Z\in C\}}\right]
\end{aligned}$$

Równość $\star$ wynika ze sposobu w jaki zdefiniowaliśmy $\mathbb{E}\left[X \mid A\right]$ w poprzednim podrozdziale.

Zauważmy, że dowolne zdarzenie $F \in \sigma(Z)$ jest postaci $F = \{z \in C\}$ dla pewnego $C \in \mathrm{Bor}(\mathbb{R})$. Wyprowadziliśmy więc równość:

$$\mathbb{E}[Y\mathbb{1}_F] = \mathbb{E}[X\mathbb{1}_F] \qquad F \in \sigma(Z).$$

Pozostaje zapytać, co z tej zależności wynika?

Dla $F = \Omega$ dostajemy

$$\mathbb{E}\left[h(Z)\right] = \mathbb{E}[Y] = \mathbb{E}[X].$$

Dygresja.
W tym miejscu kuszące byłoby rozpisanie $Y = h(Z)$ wprost z definicji, tzn. $h(Z) = \mathbb{E}\left[X \mid Z = Z\right]$, ale jest to całkowitą brednią. W definicji funkcji h podanej na samym początku przykładu z jest teoretycznym punkcikiem, natomiast przy definiowaniu $Y = h(Z)$ ów Z jest już obserwowaną przez nas, konkretną zmienną losową. W takim razie, bardziej poprawny byłby zapis

$$h(Z(\omega)) = \mathbb{E}\left[X \mid \{\omega' \ : \ Z(\omega') = Z(\omega)\}\right].$$

Przykład(y) 1.1

1. Ze zbioru $\{1, 2, ..., 10\}$ losujemy w sposób jednostajny liczbę i oznaczamy ją jako N. W drugim losowaniu, również w sposób jednostajny, wybieramy liczbę ze zbioru $\{1, ..., N\}$ i nazywamy ją M. Chcemy znaleźć średnią wartość liczby M. Oczywiście, nie jest trudno zrobić to metodami poznanymi na poprzednich przygodach probabilistycznych, jednak w tym przypadku użyjemy konstrukcji wyżej.

   Funkcja h będzie wyglądać następująco:

   $$h(n) = \mathbb{E}\left[M \mid N = n\right] = \sum_{1\le i\le n} \frac{i}{n} = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$$

   czyli $h(N) = \frac{N+1}{2}$.

   Stosując notację jak wyżej, mamy

   $$\begin{gathered} Z = N \\ X = M \end{gathered}$$

czyli podstawiając do wzoru:

$$\begin{aligned}
\mathbb{E}\left[M\right] = \mathbb{E}\left[h(N)\right] = \mathbb{E}\left[\frac{N+1}{2}\right] &= \\
= \frac{1}{2}\left(\mathbb{E}\left[N\right]+1\right) = \frac{1}{2}\left(\sum_{1\leq i\leq 10}\frac{i}{10}+1\right) &= \\
= \frac{1}{2}\left(\frac{11}{2}+1\right)\frac{13}{4}
\end{aligned}$$

Rozbicie jak wyżej można w elegancki sposób zamienić w bardziej abstrakcyjną definicję warunkowej wartości oczekiwanej.

Definicja 1.1.
Niech $\mathcal{G} \subseteq \mathcal{F}$ będzie σ-ciałem, a X całkowalną zmienną losową.

Zmienną losową Y nazywamy **warunkową wartością oczekiwaną** [wwo] X pod warunkiem $\mathcal{G}$, jeśli następujące warunki są spełnione:

(W1) Y jest $\mathcal{G}$-mierzalne

(W2) $(\forall\ G \in \mathcal{G})\ \mathbb{E}\left[X\mathbb{1}_G\right] = \mathbb{E}\left[Y\mathbb{1}_G\right]$

Nasuwają się teraz pytania o poprawność Y zdefiniowanego jak wyżej. Czy zawsze istnieje i czy jest on jedyny?

Przykład(y) 1.2

1. Niech $\mathcal{G} = \sigma(Z)$, gdzie Z jest zmienną losową przyjmującą przeliczalnie wiele wartości. Wówczas $Y = h(Z)$ dla $h(z) = \mathbb{E}\left[X \mid Z = z\right]$ jest wwo X względem $\mathcal{G}$.

Twierdzenie 1.1 : poprawność wwo.
Dla σ-ciała $\mathcal{G} \subseteq \mathcal{F}$ i całkowalnej zmiennej losowej X **istnieje jedyna zmienna losowa** Y będąca wwo X względem $\mathcal{G}$. Będziemy ją oznaczać

$$\mathbb{E}\left[X \mid \mathcal{G}\right] = Y.$$

Jeśli Y, Y' są wwo X względem $\mathcal{G}$, to $Y = Y'$ prawie wszędzie.

Dowód
Dowód na następnym wykładzie.

Uwaga 1.2.

O wwo X pod warunkiem $\mathcal{G}$ należy myśleć jako o przybliżeniu X na podstawie informacji zawartych w $\mathcal{G}$ (więcej na wykładzie 3).

Przykład(y) 1.3

1. Jeśli X i $\mathcal{G}$ są niezależne, to znaczy dla każdego $B \in Bor(\mathbb{R})$ i dla każdego $G \in \mathcal{G}$ zachodzi

$$\mathbb{P}[X \in B, G] = \mathbb{P}[X \in B]\,\mathbb{P}[G],$$

to wtedy $\mathbb{E}\left[X \mid \mathcal{G}\right] = \mathbb{E}[X] = Y$.

Warunek (W1) jest oczywiście spełniony, bo Y jest funkcją stałą, więc jego przeciwobraz to całość lub $\emptyset$ (czyli jest $\mathcal{G}$-mierzalny). Warunek (W2) sprawdzamy dla dowolnego $G \in \mathcal{G}$:

$$\mathbb{E}[X\mathbb{1}_G] = \mathbb{E}[X]\,\mathbb{E}[\mathbb{1}_G] = \mathbb{E}[\mathbb{E}[X]\,\mathbb{1}_G] = \mathbb{E}[Y\mathbb{1}_G].$$

2. Rozważmy pokrycie Ω rozłącznymi zbiorami $\{A_n\}_{n\in\mathbb{N}}$, gdzie $A_i \in \mathcal{F}$ dla każdego i. Niech $\mathcal{G} = \sigma(A_i\ :\ i \in \mathbb{N})$ będzie σ-ciałem rozpinanym przez to pokrycie. Wówczas

$$\mathbb{E}\left[X \mid \mathcal{G}\right] = \sum_{i\in\mathbb{N}} \mathbb{1}_{A_i}\mathbb{E}\left[X \mid A_i\right]$$

Spełnianie pierwszego warunku jest oczywiste, bo mamy doczynienia z funkcją prostą. Warunek (W2) wystarczy sprawdzić dla atomów, czyli $G = A_j$, bo wszystkie zmienne losowe $\mathcal{G}$-mierzalne są stałe na A_i.

$$\begin{aligned}\mathbb{E}\left[\left[\sum \mathbb{1}_{A_i}\mathbb{E}\left[X \mid A_i\right]\right]\mathbb{1}_{A_j}\right] &= \mathbb{E}\left[\mathbb{E}\left[X \mid A_j\right]\mathbb{1}_{A_j}\right] = \\ &= \mathbb{E}\left[\mathbb{1}_{A_j}\frac{\mathbb{E}\left[X\mathbb{1}_{A_j}\right]}{\mathbb{P}\left[A_j\right]}\right] = \\ &= \mathbb{E}\left[\mathbb{1}_{A_j}\right]\frac{\mathbb{E}\left[X\mathbb{1}_{A_j}\right]}{\mathbb{P}\left[A_j\right]} = \mathbb{E}\left[X\mathbb{1}_{A_j}\right],\end{aligned}$$

gdyż $\mathbb{E}\left[\mathbb{1}_{A_j}\right] = \mathbb{P}\left[A_j\right]$.

3. Jeśli w przykładzie wyżej weźmiemy $A_1 = A$, $A_2 = A^c$ i $A_3 = A_4 = \ldots = \emptyset$ oraz $\mathcal{G} = \sigma(A)$, to dostajemy to samo co na samym początku tego wykładu:

$$\mathbb{E}\left[X \mid \mathcal{G}\right] - \mathbb{1}_A\mathbb{E}\left[X \mid A\right] + \mathbb{1}_{A^c}\mathbb{E}\left[X \mid A^c\right].$$

1.3 Prawdopodobieństwo warunkowe

Definicja 1.2 : prawdopodobieństwo warunkowe.
Dla σ-ciała $\mathcal{G} \subseteq \mathcal{F}$ definiujemy **prawdopodobieństwo warunkowe** pod warunkiem $\mathcal{G}$ jako

$$\mathbb{P}\left[A \mid \mathcal{G}\right] = \mathbb{E}\left[\mathbb{1}_A \mid \mathcal{G}\right]$$

Prawdopodobieństwo $\mathbb{P}\left[A \mid \mathcal{G}\right]$ jest zmienną losową taką, że:

- $\mathbb{P}\left[A \mid \mathcal{G}\right]$ jest $\mathcal{G}$-mierzalna (ze względu na wwo w definicji)

☕ $\mathbb{E}\left[\mathbb{P}\left[A \mid \mathcal{G}\right] \mathbb{1}_G\right] = \mathbb{E}\left[\mathbb{1}_A \mathbb{1}_G\right] = \mathbb{P}\left[A \cap G\right]$

Przykład(y) 1.4

1. Niech E_1, E_2 będą niezależnymi zmiennymi losowymi z rozkładem $\exp(1)$. Chcemy się zastanowić jak wygląda prawdopodobieństwo

$$\mathbb{P}\left[E_1 + E_2 > t \mid \sigma(E_1)\right]$$

dla $t > 0$. Ponieważ liczymy to prawdopodobieństwo względem $\sigma(E_1)$, to tak naprawdę wszystkie informacje o E_1 mamy w ręku, gdyż tę zmienną obserwujemy. Czyli E_1 możemy w takim przypadku potraktować jako zwykłą stałą i zgadnąć, że

$$\mathbb{P}\left[E_1 + E_2 > t \mid \sigma(E_1)\right] = e^{-(t-E_1)}.$$

Dla pewności, przerachujemy cały ten przykład wprost z definicji, żeby przekonać się że strzał był poprawny.

Niech $G \in \sigma(E_1)$, wtedy zgodnie z wcześniejszą obserwacją istnieje pewne $C \in \mathrm{Bor}(\mathbb{R})$ takie, że G jest postaci $G = \{E_1 \in C\}$. Policzymy $\mathbb{E}\left[\mathbb{P}\left[\{E_1 + E_2 > t\} \mid \sigma(E_1)\right]\right]$ gdyż jak wyżej zauważyliśmy, $\mathbb{P}\left[A \mid \mathcal{G}\right]$ jest zmienną losową. Mamy więc

$$\begin{aligned}
\mathbb{E}\left[\mathbb{P}\left[E_1 + E_2 > t \mid \sigma(E_1)\right] \mathbb{1}_G\right] &\overset{\star}{=} \mathbb{P}\left[\{E_1 + E_2 > t\} \cap G\right] = \\
&= \mathbb{P}\left[\{E_1 + E_2 > t\} \cap \{E_1 \in C\}\right] = \\
&= \iint\limits_{\substack{C \times \mathbb{R}_+ \\ x+y>t}} e^{-x} e^{-y} \, dx dy = \\
&= \int_C e^{-x} \underbrace{\left[\int_{x+y>t} e^{-y} dy\right]}_{\star\star} dx = \\
&= \int_C e^{-x} e^{-(t-x)_+} dx = \mathbb{E}\left[e^{-(t-E_1)_+} \mathbb{1}_{\{E_1 \in C\}}\right] = \\
&= \mathbb{E}\left[e^{-(t-E_1)_+} \mathbb{1}_G\right]
\end{aligned}$$

Równość $\star$ wynika z uwagi pod definicją prawdopodobieństwa warunkowego. Całka $\star\star$ jest równa 1 gdy $x > t$ (gdyż wtedy dla każdego y mamy $x + y > t$), natomiast dla $x \leq t$ wynosi ona $e^{-(t-x)}$.

Wykład 09.10.23 : Własności WWO

Na tym wykładzie zajmiemy się dowodzeniem własności wwo, w tym pokażemy jej istnienie i jedyność.

2.1 Poprawność: istnienie i jedyność

Lemat 2.1 : WWO jest całkowalna.
To znaczy, że mając całkowalną zmienną losową X oraz σ-ciało $\mathcal{G} \subseteq \mathcal{F}$, to zachodzi $\mathbb{E}\left[\left|\mathbb{E}\left[X \mid \mathcal{G}\right]\right|\right] < \infty$.

Dowód
Rozważmy zbiór

$$A = \{\omega \ : \ \mathbb{E}\left[X \mid \mathcal{G}\right](\omega) > 0\} = \{\omega \ : \ \mathbb{E}\left[X \mid \mathcal{G}\right] \in (0, \infty)\} = \left[\mathbb{E}\left[X \mid \mathcal{G}\right]\right]^{-1}((0, \infty))$$

jako przeciwobraz zbioru $(0, \infty) \in \mathrm{Bor}(\mathbb{R})$ przez funkcję $\mathcal{G}$-mierzalną $\mathbb{E}\left[X \mid \mathcal{G}\right]$ wiemy, że $A \in \mathcal{G}$. Ponieważ $\mathbb{E}\left[X \mid \mathcal{G}\right]$ jest wwo X pod warunkiem $\mathcal{G}$, to musi warunek (W2):

$$0 \leq \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_A\right] = \mathbb{E}\left[X\mathbb{1}_A\right] \leq \mathbb{E}\left[|X|\mathbb{1}_A\right] < \infty$$

bo X jest całkowalna.

Analogicznie postępujemy dla zbioru A^c:

$$0 \leq \mathbb{E}\left[-\mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_A\right] = \mathbb{E}\left[-X\mathbb{1}_{A^c}\right] \leq \mathbb{E}\left[|X|\mathbb{1}_{A^c}\right] < \infty.$$

Zauważmy, że

$$\left|\mathbb{E}\left[X \mid \mathcal{G}\right]\right| = \mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_A - \mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_{A^c}$$

Dodając obie te nierówności (i korzystając z liniowości wartości oczekiwanej) uzyskujemy

$$0 \leq \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_A\right] - \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_{A^c}\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_A - \mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_{A^c}\right] = \mathbb{E}\left[\left|\mathbb{E}\left[X \mid \mathcal{G}\right]\right|\right] < \infty$$

Lemat 2.2 : jedyność p.w..
Niech $\mathcal{G} \subseteq F$ będzie σ-ciałem. Jeśli Y i Y' są obie wersjami $\mathbb{E}\left[X \mid \mathcal{G}\right]$, to $Y = Y'$ p.w..

Dowód
Ustalmy $\varepsilon > 0$ i rozważmy zdarzenie

$$A_\varepsilon = \{Y - Y' > \varepsilon\} \in \mathcal{G}$$

które jest $\mathcal{G}$-mierzalne, bo Y i Y' takie są.

$$\begin{aligned}
\varepsilon\mathbb{P}\left[A_\varepsilon\right] + \mathbb{E}\left[Y'\mathbb{1}_{A_\varepsilon}\right] &= \mathbb{E}\left[\varepsilon\mathbb{1}_{A_\varepsilon}\right] + \mathbb{E}\left[Y'\mathbb{1}_{A_\varepsilon}\right] = \\
&= \mathbb{E}\left[(\varepsilon + Y')\mathbb{1}_{A_\varepsilon}\right] \leq \\
&\overset{\star}{\leq} \mathbb{E}\left[Y\mathbb{1}_{A_\varepsilon}\right] \overset{(W2)}{=} \mathbb{E}\left[X\mathbb{1}_{A_\varepsilon}\right] = \\
&= \mathbb{E}\left[Y'\mathbb{1}_{A_\varepsilon}\right]
\end{aligned}$$

gdzie $\star$ wynika z tego, że na zbiorze A_ε $Y > Y' + \varepsilon$.

Dostajemy więc, że

$$\varepsilon\mathbb{P}\left[A_\varepsilon\right] + \mathbb{E}\left[Y'\mathbb{1}_{A_\varepsilon}\right] \leq \mathbb{E}\left[Y'\mathbb{1}_{A_\varepsilon}\right]$$

co po przeniesieniu $\mathbb{E}$ na jedną stronę daje

$$\varepsilon\mathbb{P}\left[A_\varepsilon\right] \leq 0$$

a ponieważ $\varepsilon > 0$, to musi być $\mathbb{P}\left[A_\varepsilon\right] = 0$.

Wówczas

$$\mathbb{P}\left[Y > Y'\right] = \underbrace{\mathbb{P}\left[(\exists\ n)\ Y \geq Y' + \frac{1}{n}\right]}_{\mathbb{P}\left[A_{\frac{1}{n}}\right]} = \mathbb{P}\left[\bigcup A_{\frac{1}{n}}\right] = \lim \mathbb{P}\left[A_{\frac{1}{n}}\right] = 0$$

ponieważ $A_{\frac{1}{n}} \subseteq A_{\frac{1}{n+1}}$.

Zamieniając miejscami Y i Y' w dowodzie dostaniemy $\mathbb{P}\left[Y' > Y\right] = 0$, czyli obie możliwości są miary zero.

Twierdzenie 2.3 : o istnieniu WWO.
Niech $\mathcal{G} \subseteq \mathcal{F}$ będzie σ-ciałem, a X jest całkowalną zmienną losową. Istnieje zmienna losowa Y spełniająca oba postulaty wwo X pod warunkiem $\mathcal{G}$.

Jest to Twierdzenie 1.1 z poprzedniego wykładu.

Zanim jednak przejdziemy do dowodu 2.3, przypomnijmy *twierdzenie Radona-Nikodyma* z teorii miary:

Dygresja : twierdzenie Radona-Nikodyma.
Niech μ i ν będą σ-miarami na przestrzeni $(\Omega, \mathcal{G})$ takimi, że ν jest *absolutnie ciągła* względem μ $[\nu \ll \mu]$, tzn $\mu(A) = 0 \Rightarrow \nu(A) = 0$. Wówczas istnieje $\mathcal{G}$-mierzalna funkcja $f : \Omega \to \mathbb{R}$ taka, że

$$\nu(A) = \int_A f(x)\mu(dx)$$

Funkcję f jak wyżej często oznaczamy $f = \frac{d\nu}{d\mu}$ i nazywamy *pochodną Radona-Nikodyma*.

Dowód
Wracając do dowodu twierdzenia 2.3. Najpierw pokażemy prostszy przykład, gdy $X \geq 0$, a potem uogólnimy go do dowolnego X.

Załóżmy, że $X \geq 0$ p.w. Wtedy możemy rozważyć miary $\mu = \mathbb{P} \restriction \mathcal{G}$ oraz $\nu(A) = \mathbb{E}[X\mathbb{1}_A]$. Od razu widać, że w takim ułożeniu $\nu \ll \mu$, więc na mocy twierdzenia Radona-Nikodyma istnieje f $\mathcal{G}$-mierzalna taka, że

$$\mathbb{E}\left[f\mathbb{1}_A\right] = \int_A f(\omega)\mu(d\omega) = \nu(A) - \mathbb{E}[X\mathbb{1}_A].$$

Funkcja f spełnia (W1) z definicji wwo, bo jest $\mathcal{G}$-mierzalna, a (W2) jest potwierdzone przez rachunek wyżej. Czyli f jest wwo X pod warunkiem $\mathcal{G}$.

Niech teraz X będzie dowolną zmienną losową. Możemy ją rozbić jako

$$X = X^+ - X^-,$$

gdzie $X^+ = \max(0, X) \geq 0$ oraz $X^- = -\min(0, X) \geq 0$. Do obu tych zmiennych możemy zastosować pierwszą część dowodu, by dostać zmienne $\mathbb{E}\left[X^+ \mid \mathcal{G}\right]$ oraz $\mathbb{E}\left[X^- \mid \mathcal{G}\right]$. Wystarczy zauważyć, że dzięki liniowości $\mathbb{E}$ możemy w prosty sposób pokazać

$$\mathbb{E}\left[X \mid \mathcal{G}\right] = \mathbb{E}\left[X^+ \mid \mathcal{G}\right] - \mathbb{E}\left[X^- \mid \mathcal{G}\right]$$

2.2 Własności wwo

Twierdzenie 2.4 : o arytmetyce wwo.
Niech $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{F}$ będą σ-ciałami, a X, X_1, X_2 całkowalnymi zmiennymi losowymi

1. $\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right]\right] = \mathbb{E}[X]$
2. Jeśli $X \geq 0$, to również $\mathbb{E}\left[X \mid \mathcal{G}\right] \geq 0$
3. $\mathbb{E}\left[aX_1 + bX_2 \mid \mathcal{G}\right] = a\mathbb{E}\left[X_1 \mid \mathcal{G}\right] + b\mathbb{E}\left[X_2 \mid \mathcal{G}\right]$
4. $\left|\mathbb{E}\left[X \mid \mathcal{G}\right]\right| \leq \mathbb{E}\left[|X| \mid \mathcal{G}\right]$
5. Jeśli $\mathcal{G}_1 \subseteq \mathcal{G}_2$, to wówczas

   $$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_1\right] \mid \mathcal{G}_2\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_2\right] \mid \mathcal{G}_1\right] = \mathbb{E}\left[X \mid \mathcal{G}_1\right]$$

   To znaczy, że mając informacje o X w dwóch zawartych w sobie ciałach, to mniejsze zawsze wygrywa.
6. Jeśli Y jest $\mathcal{G}$-mierzalna i XY jest całkowalna, to $\mathbb{E}\left[XY \mid \mathcal{G}\right] = Y\mathbb{E}\left[X \mid \mathcal{G}\right]$, czyli Y możemy traktować jako stałą.

Dowód

1. Wystarczy wstawić $G = \Omega$ w warunek (W2):

$$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right]\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_{\Omega}\right] \overset{(W2)}{=} \mathbb{E}\left[X\mathbb{1}_{\Omega}\right] = \mathbb{E}\left[X\right]$$

2. Wynika z dowodu twierdzenia o istnieniu, bo $\frac{d\nu}{d\mu} = \mathbb{E}\left[X \mid \mathcal{G}\right]$. Gdyby $A = \{\omega \ : \ \mathbb{E}\left[X \mid \mathcal{G}\right] < 0\}$, to wówczas

$$\mathbb{E}\left[X\mathbb{1}_A\right] = \nu(A) = \int_A \mathbb{E}\left[X \mid \mathcal{G}\right](\omega)\mu(d\omega) = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right] \mathbb{1}_A\right] < 0$$

ale przecież $X \geq 0 \Rightarrow \mathbb{E}\left[X\mathbb{1}_A\right] \geq 0$, więc $A = \emptyset$.

3. Można to zrobić na dwa sposoby: licząc wszystko pokolei, albo można sprawdzić, czy $Y = a\mathbb{E}\left[X_1 \mid \mathcal{G}\right] + b\mathbb{E}\left[X_2 \mid \mathcal{G}\right]$ spełnia warunki wwo tej samej zmiennej co $\mathbb{E}\left[aX_1 + bX_2 \mid \mathcal{G}\right]$. Wówczas obie te zmienne są równe prawie wszędzie.

   Warunek $\mathcal{G}$-mierzalności dla Y jest spełniony, bo Y jest kombinacją liniową dwóch funkcji $\mathcal{G}$-mierzalnych. Wystarczy więc sprawdzić warunek (W2). W tym celu ustalmy $A \in \mathcal{G}$.

$$\begin{aligned}\mathbb{E}\left[Y\mathbb{1}_A\right] &\overset{\star}{=} a\mathbb{E}\left[\mathbb{E}\left[X_1 \mid \mathcal{G}\right] \mathbb{1}_A\right] + b\mathbb{E}\left[\mathbb{E}\left[X_2 \mid \mathcal{G}\right] \mathbb{1}_A\right] = \\ &= a\mathbb{E}\left[X_1\mathbb{1}_A\right] + b\mathbb{E}\left[X_2\mathbb{1}_A\right] = \\ &= \mathbb{E}\left[(aX_1 + bX_2)\mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{E}\left[aX_1 + bX_2 \mid \mathcal{G}\right] \mathbb{1}_A\right]\end{aligned}$$

4. Wiemy, że $-|X| \leq X \leq |X|$. Korzystając z punktu 2 dostajemy

$$0 \leq X + |X| \Rightarrow 0 \leq \mathbb{E}\left[|X| \mid \mathcal{G}\right] + \mathbb{E}\left[X \mid \mathcal{G}\right] \Rightarrow -\mathbb{E}\left[|X| \mid \mathcal{G}\right] \leq \mathbb{E}\left[X \mid \mathcal{G}\right]$$

$$0 \leq |X| - X \Rightarrow 0 \leq \mathbb{E}\left[|X| \mid \mathcal{G}\right] - \mathbb{E}\left[X \mid \mathcal{G}\right] \Rightarrow \mathbb{E}\left[X \mid \mathcal{G}\right] \leq \mathbb{E}\left[|X| \mid \mathcal{G}\right]$$

   Po złożeniu tych dwóch nierówności:

$$-\mathbb{E}\left[|X| \mid \mathcal{G}\right] \leq \mathbb{E}\left[X \mid \mathcal{G}\right] \leq \mathbb{E}\left[|X| \mid \mathcal{G}\right]$$

   wiemy, że $-|\mathbb{E}\left[X \mid \mathcal{G}\right]| \leq \mathbb{E}\left[X \mid \mathcal{G}\right] \leq |\mathbb{E}\left[X \mid \mathcal{G}\right]|$, więc musi być

$$|\mathbb{E}\left[X \mid \mathcal{G}\right]| \leq \mathbb{E}\left[|X| \mid \mathcal{G}\right].$$

5. Zaczniemy od sprawdzenia, że $\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_2\right] \mid \mathcal{G}_1\right] = \mathbb{E}\left[X \mid \mathcal{G}_1\right]$. Wybierzmy $A \in \mathcal{G}_1 \subseteq \mathcal{G}_2$:

$$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_1\right] \mathbb{1}_A\right] = \mathbb{E}\left[X\mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_2\right] \mathbb{1}_A\right]$$

   co potwierdza warunek (W2). $\mathcal{G}_1$-mierzalność $\mathbb{E}\left[X \mid \mathcal{G}_2\right]$ jest oczywista, gdyż $\mathcal{G}_1 \subseteq \mathcal{G}_2$, $\mathbb{E}\left[X \mid \mathcal{G}_2\right]$ jest $\mathcal{G}_2$-mierzalne, a po obcięciu do $\mathcal{G}_1$ dostajemy funkcję $\mathcal{G}_1$-mierzalną.

   Pozostaje nam sprawdzić czym jest $\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_1\right] \mid \mathcal{G}_2\right]$. Roboczo nazwiemy $Y = \mathbb{E}\left[X \mid \mathcal{G}_1\right]$. Jest to funkcja $\mathcal{G}_1$-mierzalna, ale dzięki $\mathcal{G}_1 \subseteq \mathcal{G}_2$ mamy też $\mathcal{G}_2$-mierzalność. W takim razie (tak jak w jednym z przykładów z pierwszego wykładu) $\mathbb{E}\left[Y \mid \mathcal{G}_2\right] = Y$. Pisząc bez używania litery Y dostajemy

$$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_1\right] \mid \mathcal{G}_2\right] = \mathbb{E}\left[X \mid \mathcal{G}_1\right]$$

6. Ćwiczenie, a poniżej moja próba.

   Jeśli Y jest $\mathcal{G}$-mierzalne, to $Y\mathbb{E}\left[X \mid \mathcal{G}\right]$ też takie jest jako iloczyn dwóch funkcji $\mathcal{G}$-mierzalnych. Pozostaje sprawdzić warunek (W2).

Zacznijmy od $Y = \sum a_i \mathbb{1}_{A_i}$ dla $A_i \in \mathcal{G}$, czyli od funkcji prostej. Wybierając $A \in \mathcal{G}$ możemy ograniczyć się do zbiorów A_i, gdyż są one rozłączne i na dowolnym innym zbiorze $Y = 0$. Mamy więc

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[XY \mid \mathcal{G}\right]\mathbb{1}_{A_i}\right] &\overset{(W2)}{=} \mathbb{E}\left[XY\mathbb{1}_{A_i}\right] = \mathbb{E}\left[a_i X \mathbb{1}_{A_i}\right] = a_i \mathbb{E}\left[X\mathbb{1}_{A_i}\right] = \\ &= a_i \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}\right]\mathbb{1}_{A_i}\right] = \mathbb{E}\left[(a_i \mathbb{E}\left[X \mid \mathcal{G}\right])\mathbb{1}_{A_i}\right] = \\ &= \mathbb{E}\left[(Y\mathbb{E}\left[X \mid \mathcal{G}\right])\mathbb{1}_{A_i}\right]\end{aligned}$$

Czyli $\mathbb{E}\left[XY \mid \mathcal{G}\right] = Y\mathbb{E}\left[X \mid \mathcal{G}\right]$ dla przypadku gdy Y jest funkcją prostą.

Jeśli teraz Y jest dowolną nieujemną funkcją mierzalną, to istnieje ciąg funkcji prostych

$$s_1 \leq s_2 \leq \ldots \leq s_n \leq \ldots \qquad \lim s_i = Y$$

Wówczas dla dowolnego $A \in \mathcal{G}$

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[XY\right]\mathbb{1}_A\right] = \mathbb{E}\left[XY\mathbb{1}_A\right] = \mathbb{E}\left[X \lim s_i \mathbb{1}_A\right] &\overset{\star}{=} \lim \mathbb{E}\left[X s_i \mathbb{1}_A\right] = \\ &\overset{\star\star}{=} \lim \mathbb{E}\left[s_i \mathbb{E}\left[X \mid \mathcal{G}\right]\mathbb{1}_A\right] = \mathbb{E}\left[\lim s_i \mathbb{E}\left[X \mid \mathcal{G}\right]\mathbb{1}_A\right] = \\ &= \mathbb{E}\left[Y\mathbb{E}\left[X \mid \mathcal{G}\right]\mathbb{1}_A\right]\end{aligned}$$

$\star$ można zrobić na mocy twierdzenia o monotoniczności ciągu s_i dla zwykłej $\mathbb{E}$, natomiast $\star\star$ stosuje poprzedni przypadek Y.

Pozostaje przypadek, gdy Y jest dowolną $\mathcal{G}$-mierzalną zmienną losową. Wówczas możemy rozbić $Y = Y^+ - Y^-$ i skorzystać z liniowości wwo:

$$\mathbb{E}\left[XY \mid \mathcal{G}\right] = \mathbb{E}\left[XY^+ \mid \mathcal{G}\right] - \mathbb{E}\left[XY^- \mid \mathcal{G}\right] = Y^+\mathbb{E}\left[X \mid \mathcal{G}\right] - Y^-\mathbb{E}\left[X \mid \mathcal{G}\right] = Y\mathbb{E}\left[X \mid \mathcal{G}\right]$$

Twierdzenie 2.5 : o zbieżności i ciągłości.

Niech $\mathcal{G} \subseteq \mathcal{F}$ będzie σ-ciałem, a $X, X_1, X_2, \ldots$ będzie ciągiem całkowalnych zmiennych losowych. Wówczas

1. Jeśli $0 \leq X_1 \leq X_2 \leq \ldots$ oraz $X_n \nearrow X$, to $\mathbb{E}\left[X_n \mid \mathcal{G}\right] \nearrow \mathbb{E}\left[X \mid \mathcal{G}\right]$ p.w. (twierdzenie o zbieżności monotonicznej)
2. Jeśli $X \geq 0$, to $\mathbb{E}\left[\liminf_n X_n \mid \mathcal{G}\right] \leq \liminf_n \mathbb{E}\left[X_n \mid \mathcal{G}\right]$ (lemat Fatou).
3. Jeśli $|X_n| \leq Y$ oraz Y jest całkowalny i $X_n \to X$ p.w., to $\mathbb{E}\left[X_n \mid \mathcal{G}\right] \to \mathbb{E}\left[X \mid \mathcal{G}\right]$ (twierdzenie o zbieżności ograniczonej)

Dowód

1. Zauważamy, że ciąg $\mathbb{E}\left[X_n \mid \mathcal{G}\right]$ jest niemalejący i ograniczony przez $\mathbb{E}\left[X \mid \mathcal{G}\right]$ (na mocy punktu 2 z poprzedniego twierdzenia).

Niech $Y = \lim \mathbb{E}\left[X_n \mid \mathcal{G}\right]$. Wystarczy, że pokażemy $Y = \mathbb{E}\left[X \mid \mathcal{G}\right]$ p.w., czyli sprawdzimy warunki (W1) i (W2). Oczywiście, warunek (W1) wynika z faktu, że granica ciągu funkcji $\mathcal{G}$-mierzalnych jest nadal $\mathcal{G}$-mierzalna. Dla sprawdzenia warunku (W2) wybierzmy $A \in \mathcal{G}$

$$\begin{aligned} \mathbb{E}\left[Y\mathbb{1}_A\right] = \mathbb{E}\left[\lim \mathbb{E}\left[X_n \mid \mathcal{G}\right]\mathbb{1}_A\right] &\overset{\star}{=} \lim \mathbb{E}\left[\mathbb{E}\left[X_n \mid \mathcal{G}\right]\mathbb{1}_A\right] = \\ &= \lim \mathbb{E}\left[X_n\mathbb{1}_A\right] = \mathbb{E}\left[\lim X_n\mathbb{1}_A\right] = \mathbb{E}\left[X\mathbb{1}_A\right] \end{aligned}$$

czyli $Y = \mathbb{E}\left[X \mid \mathcal{G}\right]$ p.w.

2. Zaczniemy od dwóch obserwacji:

   - Dla ciągu $\{a_n\}$ $\liminf a_n$ to najmniejszy z jego punktów skupienia, równoważnie:

   $$\liminf_n a_n = \lim_{k\to\infty} \inf_{n\geq k} a_n$$

   - Dla dowolnej przeliczalnej rodziny zmiennych losowych $\{Z_n\}_{n\in T}$ i dla dowolnego $t \in T$ mamy

   $$\inf_{s\in T} Z_s \leq Z_t$$
   $$\mathbb{E}\left[\inf_{s\in T} Z_s\right] \leq \mathbb{E}\left[Z_t\right]$$
   $$\mathbb{E}\left[\inf_{s\in T} Z_s\right] \leq \inf_{t\in T} \mathbb{E}\left[Z_t\right]$$

   (co jest tak naprawdę wersją lematu Fatou dla $\mathbb{E}$ z RP1R).

   Stosując obserwację w przejściach $\star$, obserwację w przejsciu $\star\star$ oraz ptk 1. ($\inf_{n>k} X_n \leq \inf_{n>k+1} X_n$) w $\star\star\star$, dostajemy

   $$\begin{aligned} \mathbb{E}\left[\liminf_n X_n \mid \mathcal{G}\right] \overset{\star}{=} \mathbb{E}\left[\lim_k \inf_{n>k} X_n \mid \mathcal{G}\right] &\overset{\star\star\star}{=} \lim_k \mathbb{E}\left[\inf_{n>k} X_n \mid \mathcal{G}\right] \leq \\ &\overset{\star\star}{\leq} \lim_k \inf_{n>k} \mathbb{E}\left[X_n \mid \mathcal{G}\right] \overset{\star}{=} \liminf_n \mathbb{E}\left[X_n \mid \mathcal{G}\right] \end{aligned}$$

3. Rozważmy zmienne $X_n' = Y + X_n$. Ponieważ $|X_n| \leq Y$, to $Y + X_n \geq 0$.

   $$\begin{aligned} \mathbb{E}\left[Y + \liminf X_n \mid \mathcal{G}\right] = \mathbb{E}\left[\liminf(X_n + Y) \mid \mathcal{G}\right] &\overset{3.}{\leq} \liminf \mathbb{E}\left[Y + X_n \mid \mathcal{G}\right] = \\ &= \mathbb{E}\left[Y\right] + \liminf \mathbb{E}\left[X_n \mid \mathcal{G}\right] \end{aligned}$$

   To daje nam, że $\mathbb{E}\left[\liminf X_n \mid \mathcal{G}\right] \leq \liminf \mathbb{E}\left[X_n \mid \mathcal{G}\right]$.

   Postępując analogicznie dla $X_n" = Y - X_n$ (które dalej jest ≥ 0) dostaniemy $\mathbb{E}\left[\limsup X_n \mid \mathcal{G}\right] \geq \limsup \mathbb{E}\left[X_n \mid \mathcal{G}\right]$:

   $$\begin{aligned} \mathbb{E}\left[Y - \limsup X_n \mid \mathcal{G}\right] &= \mathbb{E}\left[\limsup(Y - X_n) \mid \mathcal{G}\right] = \\ &= -\mathbb{E}\left[\liminf(X_n - Y) \mid \mathcal{G}\right] \overset{3.}{\geq} -\liminf \mathbb{E}\left[X_n - Y \mid \mathcal{G}\right] = \\ &= \limsup \mathbb{E}\left[Y - X_n \mid \mathcal{G}\right] \end{aligned}$$

   Ale wiemy, że $\liminf X_n = X$ oraz $\limsup X_n = X$, czyli

   $$\liminf \mathbb{E}\left[X_n \mid \mathcal{G}\right] \geq \mathbb{E}\left[\liminf X_n \mid \mathcal{G}\right] = \mathbb{E}\left[X \mid \mathcal{G}\right] = \mathbb{E}\left[\limsup X_n \mid \mathcal{G}\right] \geq \limsup \mathbb{E}\left[X_n \mid \mathcal{G}\right]$$

ale przecież $\liminf \mathbb{E}\left[X_n \mid \mathcal{G}\right] \leq \limsup \mathbb{E}\left[X_n \mid \mathcal{G}\right]$, czyli musi być

$$\liminf \mathbb{E}\left[X_n \mid \mathcal{G}\right] = \mathbb{E}\left[X \mid \mathcal{G}\right] = \limsup \mathbb{E}\left[X_n \mid \mathcal{G}\right]$$

i ponieważ $\liminf = \limsup = \lim$ to mamy

$$\lim \mathbb{E}\left[X_n \mid \mathcal{G}\right] = \mathbb{E}\left[X \mid \mathcal{G}\right].$$

Twierdzenie 2.6.
Niech $\mathcal{G} \subseteq \mathcal{F}$ będzie σ-ciałem. Załóżmy, że

- X jest $\mathcal{G}$-mierzalna
- Y jest niezależna od $\mathcal{G}$
- funkcja $\psi : \mathbb{R}^2 \to \mathbb{R}$ jest mierzalna taka, że

$$\mathbb{E}\left[|\psi(X, Y)|\right] < \infty.$$

Wówczas

$$\mathbb{E}\left[\psi(X, Y) \mid \mathcal{G}\right] = \Psi(X),$$

gdzie funkcja $\Psi : \mathbb{R} \to \mathbb{R}$ jest zdefiniowana jako $\Psi(x) = \mathbb{E}\left[\psi(x, Y)\right]$.

Dowód

Tak jak w dowodzie ostatniego punktu twierdzenia 2.4 zaczniemy od funkcji φ prostych i stopniowo przejdziemy do dowolnych funkcji mierzalnych.

Zaczniemy od funkcji φ postaci $\varphi(x, y) = \mathbb{1}_{A\times B}(x, y) = \mathbb{1}_A(x) \cdot \mathbb{1}_B(y)$ dla pewnych $A, B \in \mathrm{Bor}(\mathbb{R})$.

Po pierwsze zauważmy, że jeżeli X jest $\mathcal{G}$-mierzalne, to $\mathbb{1}_A(X)$ też takie jest. Analogicznie, jeśli Y jest niezależna od $\mathcal{G}$, to $\mathbb{1}_B(Y)$ też jest niezależne i wtedy $\mathbb{E}\left[\mathbb{1}_B(Y) \mid \mathcal{G}\right] = \mathbb{E}\left[\mathbb{1}_B(Y)\right]$. Korzystając z 2.4 w przejściu $\star$, dostajemy

$$\begin{aligned}\mathbb{E}\left[\varphi(X, Y) \mid \mathcal{G}\right] &= \mathbb{E}\left[\mathbb{1}_A(X)\mathbb{1}_B(Y) \mid \mathcal{G}\right] = \\ &\overset{\star}{=} \mathbb{1}_A(X)\mathbb{E}\left[\mathbb{1}_B(Y) \mid \mathcal{G}\right] = \\ &= \mathbb{1}_A(X)\mathbb{E}\left[\mathbb{1}_B(Y)\right]\end{aligned}$$

Z drugiej strony mamy

$$\begin{aligned}\Psi(x) = \mathbb{E}\left[\varphi(x, Y)\right] &= \mathbb{E}\left[\mathbb{1}_A(x)\mathbb{1}_B(Y)\right] = \\ &= \mathbb{1}_A(x)\mathbb{E}\left[\mathbb{1}_B(Y)\right]\end{aligned}$$

czyli $\Psi(X) = \mathbb{1}_A(X)\mathbb{E}\left[\mathbb{1}_B(Y)\right]$ tak jak chcieliśmy.

Chcemy przejść teraz do funkcji postaci $\varphi(x, y) = \mathbb{1}_C(x, y)$ dla $C \in \mathrm{Bor}(\mathbb{R}^2)$. Skorzystamy przy tym z lematu o $\pi - \lambda$ układach:

Dygresja : lemat o $\pi - \lambda$ układach.
Niech P będzie π-układem (tzn. $A, B \in P \Rightarrow A \cap B \in P$) oraz niech $P \subseteq L$ będzie λ-układem ($\Omega \in L, A \subseteq B \in L \Rightarrow B \setminus A \in L$ i $A_1 \subseteq A_2 \subseteq ... \in L \Rightarrow \bigcup A_i \in L$).

Wówczas L jest σ-ciałem i w szczególności zawiera σ-ciało generowane przez π-układ P.

Rozważmy zbiór

$$D = \{C \in \mathrm{Bor}\,\mathbb{R}^2 \ : \ \mathbb{E}\left[\mathbb{1}_C(X, Y) \mid \mathcal{G}\right] = \Psi_C(X)\}.$$

Oczywiście, zbiór wszystkich "kwadratów" $A \times B$ dla $A, B \subseteq \mathbb{R}$ jest π-układem zbiorów z $\mathbb{R}^2$ i zgodnie z tym co już pokazaliśmy, zawiera się on w D. Chcemy więc pokazać, że D jest λ-układem.

1. $\Omega \in D$

   Jest to prawdą, bo w tym przypadku $\Omega = \mathbb{R} \times \mathbb{R}$, czyli podlega pod przypadek udowodniony wyżej.

2. $A \subseteq B \in D \Rightarrow B \setminus A \in D$

   Niech $A \subseteq B \in D$, wówczas

   $$\mathbb{1}_{B\setminus A}(X, Y) = \mathbb{1}_B(X, Y) - \mathbb{1}_A(X, Y)$$

   czyli wówczas

   $$\begin{aligned}\mathbb{E}\left[\mathbb{1}_{B\setminus A}(X, Y) \mid \mathcal{G}\right] &= \mathbb{E}\left[\mathbb{1}_B(X, Y) - \mathbb{1}_A(X, Y) \mid \mathcal{G}\right] = \\ &= \mathbb{E}\left[\mathbb{1}_B(X, Y) \mid \mathcal{G}\right] - \mathbb{E}\left[\mathbb{1}_A(X, Y) \mid \mathcal{G}\right] = \\ &= \Psi_B(X) - \Psi_A(Y) = \mathbb{E}\left[\mathbb{1}_B(X, Y)\right] - \mathbb{E}\left[\mathbb{1}_A(X, Y)\right] = \\ &= \mathbb{E}\left[\mathbb{1}_{B\setminus A}(X, Y)\right] = \Psi(X)\end{aligned}$$

   i tym samym dostajemy $B \setminus A \in D$

3. $A_1 \subseteq A_2 \subseteq ... \Rightarrow \bigcup A_i \in D$

   Wystarczy zauważyć, że przy wstępującym ciągu zbiorów A_i mamy $\mathbb{1}_{A_n} \geq \mathbb{1}_{A_{n-1}}$ oraz $\mathbb{1}_{\bigcup A_i} = \lim \mathbb{1}_{A_i}$, a następnie zastosować twierdzenie o zbieżności monotonicznej:

   $$\begin{aligned}\mathbb{E}\left[\bigcup \mathbb{1}_{A_i}(X, Y) \mid \mathcal{G}\right] &= \mathbb{E}\left[\lim \mathbb{1}_{A_i}(X, Y) \mid \mathcal{G}\right] = \\ &= \lim \mathbb{E}\left[\mathbb{1}_{A_i}(X, Y) \mid \mathcal{G}\right] = \\ &= \lim \Psi_{A_i}(X) = \lim \mathbb{E}\left[\mathbb{1}_{A_i}(X, Y)\right] = \\ &= \mathbb{E}\left[\lim \mathbb{1}_{A_i}(X, Y)\right] = \\ &= \mathbb{E}\left[\mathbb{1}_{\bigcup A_i}(X, Y)\right]\end{aligned}$$

W ten sposób pokazaliśmy już, że twierdzenie jest prawdziwe dla funkcji prostych $\varphi : \mathbb{R}^2 \to \mathbb{R}$. Przejdziemy teraz do przypadku, gdy φ jest nieujemną funkcją mierzalną, czyli istnieje ciąg funkcji prostych $s_1 \leq s_2 \leq ... \leq s_n \leq ... \leq \varphi$ taki, że $\varphi = \lim s_i$.

$$\begin{aligned}\mathbb{E}\left[\varphi(X, Y) \mid \mathcal{G}\right] &= \mathbb{E}\left[\lim s_i(X, Y) \mid \mathcal{G}\right] = \\ &\overset{\heartsuit}{=} \lim \mathbb{E}\left[s_i(X, Y) \mid \mathcal{G}\right] = \\ &= \lim \mathbb{E}\left[s_i(X, Y)\right] = \\ &= \mathbb{E}\left[\lim s_i(X, Y)\right] = \mathbb{E}\left[\varphi(X, Y)\right]\end{aligned}$$

W przejściu ♡ skorzystaliśmy ponownie z twierdzeniu o zbieżności monotonicznej.

Pozostaje jedynie przypadek, gdy φ jest dowolną funkcją mierzalną. Wtedy możemy zapisać $\varphi = \varphi^+ - \varphi^-$ dla φ^+ oraz φ^- nieujemnych. Wtedy, korzystając z wcześniej już pokazanych form funkcji mierzalnych dostajemy

$$\begin{aligned}\mathbb{E}\left[\varphi(X, Y) \mid \mathcal{G}\right] &= \mathbb{E}\left[\varphi^+(X, Y) - \varphi^-(X, Y) \mid \mathcal{G}\right] = \\ &= \mathbb{E}\left[\varphi^+(X, Y) \mid \mathcal{G}\right] - \mathbb{E}\left[\varphi^-(X, Y) \mid \mathcal{G}\right] = \\ &= \mathbb{E}\left[\varphi^+(X, Y)\right] - \mathbb{E}\left[\varphi^-(X, Y)\right] = \\ &= \mathbb{E}\left[\varphi^+(X, Y) - \varphi^-(X, Y)\right] = \mathbb{E}\left[\varphi(X, Y)\right]\end{aligned}$$

🍻

2.3 Zadania

Zadanie 1.
Niech $\mathcal{G} \subseteq \mathcal{F}$ będzie σ-ciałem. Rozważmy $\mathcal{G}$-mierzalną zmienną losową X oraz niezależną od $\mathcal{G}$ zmienną losową Y. Załóżmy, że $\psi : \mathbb{R}^2 \to \mathbb{R}$ jest taką funkcją mierzalną, że $\mathbb{E}\left[|\psi(X, Y)|\right] < \infty$. Pokaż, że

$$\mathbb{E}\left[\psi(X, Y) \mid \mathcal{G}\right] = \Psi(X) \quad \Psi(x) = \mathbb{E}\left[\psi(x, Y)\right]$$

Rozwiązanie.
Patrz dowód twierdzenia 2.6.

Zadanie 2.
Niech (X, Y) będzie dwuwymiarowym wektorem losowych o rozkładzie jednostajnym na kwadracie o wierzchołkach $(-1, 0), (0, -1), (1, 0), (0, 1)$. Oblicz $\mathbb{P}\left[X > \frac{1}{2} \mid \mathcal{Y}\right]$.

Rozwiązanie.
Zaczniemy od znalezienia gęstości rozkładu Y.

Oczywiście, gęstość rozkładu wektora $(X, Y) = \frac{1}{2}$ gdyż kwadrat ma pole 2. W takim razie, gęstość zmiennej Y to

$$\begin{aligned}f_Y(y) &= \begin{cases} \int_{y-1}^{1-y} \frac{1}{2} dx & y \geq 0 \\ \int_{-y-1}^{1+y} \frac{1}{2} dx & y < 0 \end{cases} = \\ &= \begin{cases} 1 - y & y \geq 0 \\ 1 + y & y < 0 \end{cases} = 1 - |y|\end{aligned}$$

Skorzystamy z zadania 4 z listy 1, gdzie pokazaliśmy, że

$$\mathbb{E}\left[h(X) \mid Y\right] = \int_{\mathbb{R}} h(x) f_{X|Y}(x, Y) dx$$

gdzie

$$f_{X|Y}(x,y) = \begin{cases} \frac{f(x,y)}{f_Y(y)} & f_Y(y) \neq 0 \\ 0 & \text{wpp.} \end{cases}$$

W tym zadaniu chcemy wyliczyć

$$\mathbb{P}\left[2X > 1 \mid Y\right] = \mathbb{E}\left[\mathbb{1}_{\{2X>1\}} \mid Y\right] = \int_{1/2}^{1} \mathbb{1}_{\square}(x,Y)\frac{1}{2-2|Y|}dx =$$
$$= \int_{1/2}^{1-|Y|} \frac{\mathbb{1}_{[-1/2,1/2]}(Y)}{2-2|Y|}dx = \mathbb{1}_{[-1/2,1/2]}(Y)\frac{1/2-|Y|}{2-2|Y|}$$

Zadanie 3.
Niech $\{X_n\}$ będzie ciągiem niezależnych zmiennych losowych o takim samym rozkładzie z wartością oczekiwaną m. Niech N będzie dyskretną zmienną losową o wartościach w $\mathbb{N}$ niezależną od ciągu $\{X_n\}$ z wartością oczekiwaną M. Zdefiniujmy $S_n = \sum_{k=1}^{n} x_k$. Znajdź

$$\mathbb{E}\left[S_N \mid N\right] \quad \text{oraz} \quad \mathbb{E}\left[S_N\right]$$

Rozwiązanie.
Możemy od razu zacząć od tezy, że

$$\mathbb{E}\left[S_N \mid N\right] = N \cdot m$$

ale spróbujemy rozwiązać to w bardziej metodyczny sposób.

Niech $G \in \sigma(N)$, czyli $G = \{N \in C\}$ dla $C \in Bor(\mathbb{R})$.

$$\begin{aligned}
\mathbb{E}\left[\mathbb{E}\left[S_N \mid N\right]\mathbb{1}_G\right] &= \mathbb{E}\left[S_n\mathbb{1}_G\right] = \\
&= \mathbb{E}\left[\sum_{k=1}^{N} X_k\mathbb{1}_{\{N\in C\}}\right] = \\
&= \sum_{n\in C}\mathbb{E}\left[\sum_{k=1}^{n} X_k\right]\mathbb{P}\left[N=n\right] = \\
&= \sum_{n\in C}\sum_{k=1}^{n}\mathbb{P}\left[N=n\right]\mathbb{E}\left[\sum_{k=1}^{n} X_k\right] = \\
&= m\sum_{n\in C} n\cdot\mathbb{P}\left[N=n\right] = \\
&= m\cdot\mathbb{E}\left[N\mathbb{1}_G\right] = \mathbb{E}\left[N\cdot m\mathbb{1}_G\right]
\end{aligned}$$

Czyli warunek (W2) jest spełniony przez Nm, a warunek $\sigma(N)$-mierzlaności jest spełniony przez fakt, że N jest $\sigma(N)$-mierzalne.

Korzystając z 1. własności wwo (2.4) wiemy, że

$$\mathbb{E}\left[S_N\right] = \mathbb{E}\left[\mathbb{E}\left[S_N \mid N\right]\right] = \mathbb{E}\left[N\cdot m\right] = m\cdot M$$

Zadanie 4.
Zmienne losowe X i Y są niezależne i mają rozkład Exp(1).

(a) Oblicz $\mathbb{E}\left[X + Y \mid X\right]$

(b) Oblicz $\mathbb{E}\left[X \mid X + Y\right]$

Rozwiązanie.

(a) Zacznijmy od szybkiego przypomnienia, że jeśli X i Y są niezależne, to Y jest niezależne od $\sigma(X)$. Dla $A, B \in \mathrm{Bor}(\mathbb{R})$ i $G = \{X \in B\} \in \sigma(X)$ mamy

$$\mathbb{P}\left[Y \in A \mid G\right] = \mathbb{P}\left[Y \in A \mid X \in B\right] = \\ = \mathbb{P}\left[Y \in A\right]\mathbb{P}\left[X \in B\right] = \mathbb{P}\left[Y \in A\right]\mathbb{P}\left[G\right]$$

Czyli wracając do treści zadania

$$\mathbb{E}\left[X + Y \mid X\right] = \mathbb{E}\left[X \mid X\right] + \mathbb{E}\left[Y \mid X\right] = X + \mathbb{E}\left[Y\right] = X + 1$$

gdyż X jest mierzalne względem $\sigma(X)$, więc $\mathbb{E}\left[X \mid X\right] = X$, a z drugiej strony ponieważ Y jest niezależne od $\sigma(X)$, to $\mathbb{E}\left[Y \mid X\right] = \mathbb{E}\left[Y\right]$.

(b) Zaczniemy od obserwacji, że

$$\mathbb{E}\left[X \mid X + Y\right] = \mathbb{E}\left[Y \mid X + Y\right]$$

ponieważ

$$\mathbb{E}\left[\mathbb{E}\left[X \mid X + Y\right]\mathbb{1}_{\{X+Y\in C\}}\right] = \mathbb{E}\left[X\mathbb{1}_{\{X+Y\in C\}}\right] = \\ = \mathbb{E}\left[Y\mathbb{1}_{\{X+Y\in C\}}\right] = \\ = \mathbb{E}\left[\mathbb{E}\left[Y \mid X + Y\right]\mathbb{1}_{\{X+Y\in C\}}\right]$$

W takim razie

$$\mathbb{E}\left[X \mid X + Y\right] = \frac{1}{2}(\mathbb{E}\left[X \mid X + Y\right] + \mathbb{E}\left[Y \mid X + Y\right]) = \frac{1}{2}\mathbb{E}\left[X + Y \mid X + Y\right] = \frac{1}{2}(X + Y)$$

Zadanie 5.
Pokaż, że jeśli X i Y są zmiennymi losowymi takimi, że X oraz XY są całkowalne oraz Y jest zmienną losową mierzalną względem $\mathcal{G}$, to

$$\mathbb{E}\left[XY \mid \mathcal{G}\right] = Y\mathbb{E}\left[X \mid \mathcal{G}\right]$$

Rozwiązanie.
Patrz dowód twierdzenia 2.4.

Zadanie 6.
Niech X będzie całkowalną zmienną losową. Niech $\mathcal{C} \subseteq \mathcal{G}$ będzie π-układem generującym

σ-ciało $\sigma(\mathcal{C}) = \mathcal{G} \subseteq \mathcal{F}$.

Rozwiązanie.
Zadanie sprowadza się do skorzystania z lematu o $\pi - \lambda$ układach i sprawdzeniu czy zbiór

$$D = \{A \ : \ \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Z\mathbb{1}_A]\}$$

zawierający π-układ $\mathcal{C}$ jest λ-układem. Wówczas D samo w sobie będzie σ-układem, w szczególności zawierającym ciało $\mathcal{G}$.

1. $\Omega \in D$ bo aby $\mathcal{C}$ generowało σ-ciało, to musi zawierać Ω.
2. $A \subseteq B \in D \Rightarrow B \setminus A \in D$

   Niech $A \subseteq B \in D$, wówczas $\mathbb{1}_{B\setminus A} = \mathbb{1}_B - \mathbb{1}_A$. Daje to:

$$\begin{aligned}\mathbb{E}\left[X\mathbb{1}_{B\setminus A}\right] &= \mathbb{E}\left[X(\mathbb{1}_B - \mathbb{1}_A)\right] = \\ &= \mathbb{E}[X\mathbb{1}_B] - \mathbb{E}[X\mathbb{1}_A] = \\ &= \mathbb{E}[Z\mathbb{1}_B] - \mathbb{E}[Z\mathbb{1}_A] = \\ &= \mathbb{E}\left[Z(\mathbb{1}_B - \mathbb{1}_A)\right] = \\ &= \mathbb{E}\left[Z\mathbb{1}_{B\setminus A}\right]\end{aligned}$$

   czyli z $B \setminus A$ zaspokaja warunek należenia do D.

3. $A_1 \subseteq A_2 \subseteq ... \in D \Rightarrow \bigcup A_i \in D$

   Dla pokazania tego warunku będziemy korzystać z twierdzenia o zbieżności monotonicznej. Zauważmy, że dla każdego $n \in \mathbb{N}$ zachodzi $X\mathbb{1}_{\bigcup A_i} \geq X\mathbb{1}_{A_n}$ oraz $X\mathbb{1}_{A_1} \leq X\mathbb{1}_{A_2} \leq ....$

$$\begin{aligned}\mathbb{E}\left[X\mathbb{1}_{\bigcup A_i}\right] &= \mathbb{E}\left[X \lim \mathbb{1}_{A_i}\right] = \\ &= \lim \mathbb{E}\left[X\mathbb{1}_{A_i}\right] = \\ &= \lim \mathbb{E}\left[Z\mathbb{1}_{A_i}\right] = \\ &+ \mathbb{E}\left[Z \lim \mathbb{1}_{A_i}\right] = \mathbb{E}\left[Z\mathbb{1}_{\bigcup A_i}\right]\end{aligned}$$

Czyli tak długo jak Z jest $\mathcal{G}$-mierzalne, to jest ono wwo X pod warunkiem że $\mathcal{G}$.

Zadanie 7.
Niech $\mathcal{G}, \mathcal{D} \subseteq \mathcal{F}$ będą niezależnymi σ-ciałami. Niech X będzie całkowalną zmienną losową.

(a) Załóżmy, że X jest niezależna od σ-ciala $\mathcal{D}$. Czy prawdą jest, że

$$\mathbb{E}\left[X \mid \sigma(\mathcal{G} \cup \mathcal{D})\right] = \mathbb{E}\left[X \mid \mathcal{G}\right]? \quad (\star)$$

(b) Pokaż, że jeżeli $\mathcal{D}$ jest niezależne od $\sigma(\sigma(X) \cup \mathcal{G})$, to $(\star)$ zachodzi.

Wykład 23.10.23 : Interpretacje geometryczne WWO

Rozważmy zmienną losową X taką, że $\mathbb{E}\left[X^2\right] < \infty$. Interesuje nas zagadnienie regresji, mianowicie obserwując inną zmienną losową Z chcemy móc X aproksymować. Przez przybliżanie X rozumiemy przybliżanie średniokwadratowe.

Szukamy więc funkcji mierzalnej $h_0 : \mathbb{R} \to \mathbb{R}$ takiej, żeby

$$\mathbb{E}\left[(X - h_0(Z))^2\right] = \inf_{h:\mathbb{R}\to\mathbb{R}} \mathbb{E}\left[(X - h(Z))^2\right]$$

Fakt 3.1.
Dla każdej zmiennej losowej Y mierzalnej względem $\sigma(Z)$ (co w skrócie będziemy notować $Y \in \sigma(Z)$) istnieje h takie, że $Y = h(Z)$.

Dowód
Zadanie, moja próba poniżej.

Zaczynamy od Y będącego funkcją prostą i przechodzimy do coraz to bardziej skomplikowanych postaci funkcji mierzalnych.

Jeżeli $Y = \mathbb{1}_A$, to ponieważ Y jest $\sigma(Z)$-mierzalne, mamy $A \in \sigma(Z)$. To znaczy, że istnieje $B \in \mathrm{Bor}(\mathbb{R})$ taki, że $Z(A) = B$, czyli $A = Z^{-1}[B]$ i wówczas

$$Y = \mathbb{1}_A = \mathbb{1}_{Z^{-1}[B]} = \mathbb{1}_B(Z)$$

Zrobimy tutaj jeszcze krok $Y = \sum a_i \mathbb{1}_{A_i}$. Dla każdego i wiemy, że $\mathbb{1}_{A_i} = h_i(Z)$, gdyż są to funkcje $\sigma(Z)$-mierzalne. W takim razie:

$$Y = \sum a_i \mathbb{1}_{A_i} = \sum a_i \cdot h_i(Z) = \left[\sum a_i \cdot h_i\right](Z)$$

a więc szukane $h = \sum a_i \cdot h_i$.

Teraz załóżmy, że istnieje ciąg funkcji prostych $s_1 \leq s_2 \leq \ldots \leq Y$ taki, że $Y = \lim s_i$. Wówczas pokazaliśmy już, że każda funkcja $s_i = h_i(Z)$ dla pewnego borelowskiego $h_i : \mathbb{R} \to \mathbb{R}$. Ciąg s_i jest zbieżny, więc również ciąg h_i musi zbiegać do pewnego h. Wówczas dla dowolnego $\omega \in \Omega$

$$Y(\omega) = \lim s_i(\omega) = \lim h_i(Z(\omega)) = h(Z(\omega))$$

czyli $Y = h(Z)$ dla $h = \lim h_i$.

Dla formalności należy rozważyć jeszcze $Y = Y^+ - Y^-$, gdzie Y^+ oraz Y^- podlegają pod poprzedni podpunkt.

$$\mathbb{E}\left[(X - h(Z))^2\right] = \inf_{Y\in\sigma(Z)} \mathbb{E}\left[(X - Y)^2\right]$$

mając pewną wiedzę o przestrzeniach Hilberta jest do tego dość naturalne podejście: rzut ortogonalny.

Dla σ-ciała $\mathcal{G} \subseteq \mathcal{F}$ będziemy rozważać

$$L^2(\mathcal{G}) = \{Y \in \mathcal{G} \ : \ \mathbb{E}\left[Y^2\right] < \infty\}$$

wówczas $L^2(\mathcal{G})$ jest przestrzenią Euklidesową z iloczynem skalarnym $\langle X, Y\rangle = \mathbb{E}[XY]$, który z kolei zadaje normę

$$\|X\| = \sqrt{\mathbb{E}\left[X^2\right]}.$$

Używając tego języka będziemy wiedzieli jak szukać Y z początku wykładu, ale najpierw fakt pomocniczy do twierdzenia które nadejdzie lada moment.

Fakt 3.2 : warunkowa wersja nierówności Cauchy'ego-Schwartza.
Dla zmiennych X, Y takich, że $\mathbb{E}\left[X^2\right], \mathbb{E}\left[Y^2\right] < \infty$ i σ-ciała $\mathcal{G} \subseteq \mathcal{F}$ zachodzi

$$\mathbb{E}\left[|XY| \mid \mathcal{G}\right] \leq \mathbb{E}\left[X^2 \mid \mathcal{G}\right]^{\frac{1}{2}} \mathbb{E}\left[Y^2 \mid \mathcal{G}\right]^{\frac{1}{2}}$$

Dowód
Zadanie, tutaj moje podejście.

Zauważmy na początku, że

$$XY = \underbrace{\frac{(\mathbb{E}\left[Y^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/4}}{(\mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/4}} X}_{X_n} \cdot \underbrace{\frac{(\mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/4}}{(\mathbb{E}\left[Y^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/4}} Y}_{Y_n}$$

przy czym korzystamy z $\frac{1}{n}$, żeby na pewno nie dzielić przez 0 gdy np. $\mathbb{E}\left[X^2 \mid \mathcal{G}\right] = 0$.

Dalej zauważmy, że ponieważ $(X_n - Y_n)^2 \geq 0$, to również

$$0 \leq \mathbb{E}\left[(X_n - Y_n)^2 \mid \mathcal{G}\right] = \mathbb{E}\left[X_n^2 + Y_n^2 - 2X_nY_n \mid \mathcal{G}\right]$$

czyli korzystając z liniowości i przenosząc $\mathbb{E}\left[X_nY_n \mid \mathcal{G}\right]$ na lewą stronę nierówności dostajemy

$$\begin{aligned}\mathbb{E}\left[XY \mid \mathcal{G}\right] &= \mathbb{E}\left[X_nY_n \mid \mathcal{G}\right] \leq \frac{1}{2}\mathbb{E}\left[X_n^2 \mid \mathcal{G}\right] + \frac{1}{2}\mathbb{E}\left[Y_n^2 \mid \mathcal{G}\right] = \\ &= \frac{1}{2}\mathbb{E}\left[\frac{(\mathbb{E}\left[Y^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}}{(\mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}} X^2 \mid \mathcal{G}\right] + \frac{1}{2}\mathbb{E}\left[\frac{(\mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}}{(\mathbb{E}\left[Y^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}} Y^2 \mid \mathcal{G}\right] = (\heartsuit)\end{aligned}$$

a ponieważ $\frac{(\mathbb{E}\left[Y^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}}{(\mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}}$ jest $\mathcal{G}$-mierzalne, to możemy wyciągnąć je przed nawias:

$$\begin{aligned}(\heartsuit) &= \frac{1}{2} \cdot \frac{(\mathbb{E}\left[Y^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}}{(\mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}} \mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{2} \cdot \frac{(\mathbb{E}\left[X^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}}{(\mathbb{E}\left[Y^2 \mid \mathcal{G}\right] + \frac{1}{n})^{1/2}} \mathbb{E}\left[Y^2 \mid \mathcal{G}\right] \to \\ &\xrightarrow{n\to\infty} (\mathbb{E}\left[Y^2 \mid \mathcal{G}\right])^{1/2} (\mathbb{E}\left[X^2 \mid \mathcal{G}\right])^{1/2}\end{aligned}$$

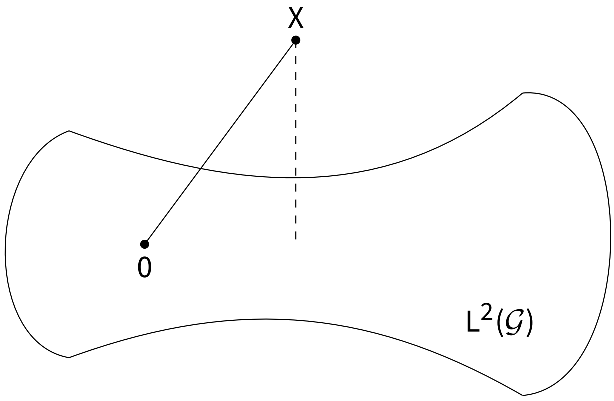
Z tego nierówności w fakcie 3.2 wynika, że dla $Y = 1$ mamy

$$\mathbb{E}\left[|X| \mid \mathcal{G}\right]^2 \leq \mathbb{E}\left[X^2 \mid \mathcal{G}\right] \Rightarrow \mathbb{E}\left[\mathbb{E}\left[X \mid Y\right]^2\right] < \infty$$

Twierdzenie 3.3 : wwo jest rzutem ortogonalnym na $L^2(\mathcal{G})$.
Niech X będzie zmienną losową taką, że $\mathbb{E}\left[X^2\right] < \infty$, a $\mathcal{G} \subseteq \mathcal{F}$ jest σ-ciałem. Wówczas

$$\mathbb{E}\left[X \mid \mathcal{G}\right] \in L^2(\mathcal{G})$$

jest *rzutem ortogonalnym* X na $L^2(\mathcal{G})$.



Dokładniej, $\mathbb{E}\left[X \mid \mathcal{G}\right]$ daje minimum zbioru $\{\mathbb{E}\left[(X - Y)^2\right] \;:\; Y \in L^2(\mathcal{G})\}$. Z faktu 3.1 dla $\mathcal{G} = \sigma(Z)$, $\mathbb{E}\left[X \mid \mathcal{G}\right] = h_0(Z)$ dla pewnego h_0.

Dowód
Dla $Y \in L^2(\mathcal{G})$ mamy

$$\begin{aligned}\mathbb{E}\left[(X-Y)^2\right] &= \mathbb{E}\left[\left((X - \mathbb{E}\left[X \mid \mathcal{G}\right]) - \overbrace{(Y - \mathbb{E}\left[X \mid \mathcal{G}\right])}^{=Y'}\right)^2\right] = \\ &= \mathbb{E}\left[(X - \mathbb{E}\left[X \mid \mathcal{G}\right])^2\right] - 2\mathbb{E}\left[(X - \mathbb{E}\left[X \mid \mathcal{G}\right])Y'\right] + \mathbb{E}\left[(Y')^2\right] = \\ &\overset{\star}{=} \mathbb{E}\left[(X - \mathbb{E}\left[X \mid \mathcal{G}\right])^2\right] + \mathbb{E}\left[(Y')^2\right]\end{aligned}$$

Zauważmy, że

$$\mathbb{E}\left[Y'X \mid \mathcal{G}\right] = Y'\mathbb{E}\left[X \mid \mathcal{G}\right]$$
$$\mathbb{E}\left[\mathbb{E}\left[Y'X \mid \mathcal{G}\right]\right] = \mathbb{E}\left[Y'X\right] = \mathbb{E}\left[Y'\mathbb{E}\left[X \mid \mathcal{G}\right]\right]$$

Przykład(y) 3.1

1. Niewiele mający z tym co przed chwilą było. Niech $\Omega = [0,1]$, $\mathcal{F} = \mathrm{Bor}([0,1])$ i $\mathbb{P} = \lambda$. Chcemy rozważyć $t \in (0,1)$ oraz $\mathcal{G} = \sigma(\mathrm{Bor}([0,t])$.

   Dla całkowalnej zmiennej losowej X szukamy $\mathbb{E}\left[X \mid \mathcal{G}\right]$.

   Dowolny $G \in \mathcal{G}$ ma postać $G = A \cup B$, gdzie $A \in \mathrm{Bor}([0,t])$ i $B \in \{(t,1], \emptyset\}$. W takim razie, jeśli $Y \in \mathcal{G}$, to Y jest stała na $(t,1]$. Czyli żeby $Y = \mathbb{E}\left[X \mid \mathcal{G}\right]$ to zapewne będzie postaci:

   $$Y(\omega) = X(\omega)\mathbb{1}_{[0,t]}(\omega) + c\mathbb{1}_{(t,1]}(\omega)$$

   gdzie c jest pewną stałą.

   Musimy sprawdzić, czy (i kiedy) $\mathbb{E}[X\mathbb{1}_G] = \mathbb{E}[Y\mathbb{1}_G]$. Widać od razu, że dla $G = A \cup B$ jak wyżej, mamy

   $$X\mathbb{1}_A = Y\mathbb{1}_A,$$

   czyli $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$. Zostaje nam uzgodnić część B kiedy jest on niepusty:

   $$\mathbb{E}[X\mathbb{1}_B] = \int_t^1 X(s)ds$$

   $$\mathbb{E}[Y\mathbb{1}_B] = c(1-t),$$

   czyli c musi być średnią X na przedziale $(t,1]$:

   $$c = \frac{1}{1-t}\int_t^1 X(s)ds.$$

3.1 Regularne rozkłady warunkowe

Dla zmiennej losowej Y i całkowalnej zmiennej losowe X napis

$$\mathbb{E}\left[X \mid Y\right] := \mathbb{E}\left[X \mid \sigma(Y)\right]$$

będzie wwo X względem σ-ciała generowanego przez Y.

Zadanie dla dociekliwych:

Jeśli $\mathbb{P}[Y = y] > 0$ dla $y \in \mathbb{R}$, to biorąc $\omega \in \{Y = y\}$ dostajemy:

$$\mathbb{E}\left[X \mid Y\right](\omega) = \mathbb{E}\left[X \mid Y = y\right] = \frac{1}{\mathbb{P}[Y = y]}\mathbb{E}\left[X\mathbb{1}_{\{Y=y\}}\right]$$

Definicja 3.1 : wwo X pod warunkiem $\{Y = y\}$.
Po pierwsze zauważamy, że istnieje funkcja $h : \mathbb{R} \to \mathbb{R}$ taka, że $\mathbb{E}\left[X \mid Y\right] = h(Y)$. $(\star)$

Niech X i Y będą dowolnymi zmiennymi losowymi, przy czym od X wymagamy całkowalności. Dla $y \in \mathbb{R}$ warunkową wartość oczekiwaną X pod warunkiem $\{Y = y\}$ definiujemy przez

$$\mathbb{E}\left[X \mid Y = y\right] = h(y)$$

gdzie h spełnia warunek $(\star)$.

Podobnie definiujemy prawdopodobieństwo zbioru A pod warunkiem $\{Y = y\}$:

$$\mathbb{P}\left[A \mid Y = y\right] = \mathbb{E}\left[\mathbb{1}_A \mid Y = y\right]$$

Przykład(y) 3.2

1. Jeżeli X i Z są niezależne, to chcemy zapytać o

$$\mathbb{P}\left[X + Z \in B \mid X = x\right] \overset{?}{=} \mathbb{P}\left[Z + x \in B\right] = \mu_Z(B - x)$$

   Wysławiając tę wartość w terminach całego X:

$$\mathbb{P}\left[X + Z \in B \mid X\right] = \mathbb{E}\left[\mathbb{1}_{X+Z\in B} \mid X\right] \overset{\star}{=} \mathbb{E}\left[\varphi(X, Z) \mid X\right] = \Phi(X)$$

   w $\star$ definiujemy: $\varphi(x, z) = \mathbb{1}_{x+z\in B}$. Dla ustalonego x mamy więc:

$$h(x) = \mathbb{E}\left[\varphi(x, Z)\right] = \mathbb{E}\left[\mathbb{1}_{x+Z\in B}\right] = \mathbb{P}\left[Z + x \in B\right]$$

2. Niech wektor losowy (X, Y) ma gęstość łączną f(x, y). Wówczas

$$\mathbb{P}\left[X \in B \mid Y = y\right] = \int_B f_{X|Y}(x, y)dx,$$

   gdzie

$$f_{X|Y}(x, y) = \frac{f(x, y)}{\int f(t, y)dt}.$$

3. Rozważmy $\mathbb{P}\left[E_1 \in \bullet \mid E_1 + E_2 = y\right]$ dla E_1, E_2 niezależnych o rozkładzie Exp(1).

   Przyłożymy do tego przypadku wzór z przykładu 2. Wektor losowy $(E_1, E_1 + E_2)$ ma rozkład losowy o łącznej gęstości $f(x, y) = e^{-x}e^{-(y-x)}\mathbb{1}_{y\geq x\geq 0}$.

$$\int f(s, y)ds = \int_0^y e^{-y}ds = y,$$

   czyli

$$f_{E_1|E_1+E_2}(x, y) = \frac{1}{y}\mathbb{1}_{y\geq x\geq 0}$$

   co daje rozkład jednostajny:

$$\mathbb{P}\left[E_1 \in \bullet \mid E_1 + E_2 = y\right] = U[0, y](\bullet)$$

Można zadać sobie pytanie, czy

$$\mathbb{P}\left[A \mid Y = y\right] = \mathbb{E}\left[\mathbb{1}_A \mid Y = y\right]$$

zawsze zadaje miarę probabilistyczną? Okazuje się, że tak faktycznie jest.

Definicja 3.2 : regularny rozkład warunkowy.
Niech X będzie zmienną losową, a $\mathcal{G} \subseteq \mathcal{F}$ będzie σ-ciałem. Funkcja

$$\kappa_{X,\mathcal{G}} : \Omega \times \mathrm{Bor}(\mathbb{R}) \to [0, 1]$$

nazywa się **regularnym rozkładem warunkowym** [rrw] X pod warunkiem $\mathcal{G}$, jeżeli

(R1) Dla każdego $B \in \mathrm{Bor}(\mathbb{R})$ zmienna losowa $\kappa_{X,\mathcal{G}}(\bullet, B)$ jest $\mathcal{G}$-mierzalna

(R2) Dla każdej $\omega \in \Omega$ $\kappa_{X,\mathcal{G}}(\omega, \bullet)$ jest rozkładem prawdopodobieństwa na prostej.

(R3) Dla każdego $B \in \mathrm{Bor}(\mathbb{R})$ i $\mathbb{P}$-p.w. $\omega \in \Omega$

$$\mathbb{P}\left[X \in B \mid \mathcal{G}\right](\omega) = \kappa_{X,\mathcal{G}}(\omega, B)$$

Twierdzenie 3.4 : rrw istnieje.
Dla każdego X i dla każdego $\mathcal{G}$ istnieje rrw X pod warunkiem $\mathcal{G}$

Dowód
W notatkach

Fakt 3.5.
Dla funkcji $f : \mathbb{R} \to \mathbb{R}$ i zmiennej losowej X takiej, że $\mathbb{E}\left[|f(X)|\right] < \infty$ mamy

$$\mathbb{E}\left[f(X) \mid \mathcal{G}\right](\omega) = \int_{\mathbb{R}} f(x)\kappa_{X,\mathcal{G}}(\omega, dx)$$

Pisząc to mówię "weź f(x) i weź miarę κ w punkcie ω i scałkuj κ".

Dowód
Ćwiczenia.

Będziemy przechodzić od najprostszych możliwych funkcji f do coraz to bardziej skomplikowanych konstrukcji.

W pierwszych kroku niech $f(x) = \mathbb{1}_B(x)$ dla $B \in \mathrm{Bor}(\mathbb{R})$. Wówczas

$$\begin{aligned}\mathbb{E}\left[f(X) \mid \mathcal{G}\right](\omega) &= \mathbb{E}\left[\mathbb{1}_B(X) \mid \mathcal{G}\right](\omega) = \mathbb{E}\left[\mathbb{1}_{\{X\in B\}} \mid \mathcal{G}\right](\omega) = \\ &= \mathbb{P}\left[X \in B \mid \mathcal{G}\right](\omega) = \kappa_{X,\mathcal{G}}(\omega, B) = \\ &= \int_B \kappa_{X,\mathcal{G}}(\omega, dx) = \int_{\mathbb{R}} \mathbb{1}_B(x)\kappa_{X,\mathcal{G}}(\omega, dx)\end{aligned}$$

Dalej, niech $f(x) = \sum a_i \mathbb{1}_{A_i}(x)$. Wtedy mamy

$$\begin{aligned}
\mathbb{E}\left[f(X) \mid \mathcal{G}\right](\omega) &= \mathbb{E}\left[\sum a_i \mathbb{1}_{A_i}(X) \mid \mathcal{G}\right](\omega) = \sum a_i \mathbb{E}\left[\mathbb{1}_{A_i}(X) \mid \mathcal{G}\right](\omega) = \\
&= \sum a_i \int_{\mathbb{R}} \mathbb{1}_{A_i}(x) \kappa_{X,\mathcal{G}}(\omega, dx) = \int_{\mathbb{R}} \sum a_i \mathbb{1}_{A_i}(x) \kappa_{X,\mathcal{G}}(\omega, dx) = \\
&= \int_{\mathbb{R}} f(x) \kappa_{X,\mathcal{G}}(\omega, dx)
\end{aligned}$$

Teraz niech $f(x) = \lim s_i(x)$ dla $0 \leq s_1 \leq s_2 \leq \ldots \leq f$ dla prostych funkcji s_i jak z poprzednich podpunktów. Zauważmy, że mamy tutaj predyspozycje do skorzystania z twierdzenia o zbieżności monotonicznej.

$$\begin{aligned}
\mathbb{E}\left[f(X) \mid \mathcal{G}\right](\omega) &= \mathbb{E}\left[\lim s_i(X) \mid \mathcal{G}\right](\omega) = \lim \mathbb{E}\left[s_i(X) \mid \mathcal{G}\right](\omega) = \\
&= \lim \int_{\mathbb{R}} s_i(x) \kappa_{X,\mathcal{G}}(\omega, dx) = \int_{\mathbb{R}} \lim s_i(x) \kappa_{X,\mathcal{G}}(\omega, dx) = \\
&= \int_{\mathbb{R}} f(x) \kappa_{X,\mathcal{G}}(\omega, dx)
\end{aligned}$$

Ostatni krok w dowodzie to $f = f^+ - f^-$ i korzysta się tutaj już tylko z poprzednich podpunktów oraz liniowości wwo:

$$\begin{aligned}
\mathbb{E}\left[f(X) \mid \mathcal{G}\right](\omega) &= \mathbb{E}\left[f^+(X) - f^-(X) \mid \mathcal{G}\right](\omega) = \mathbb{E}\left[f^+(X) \mid \mathcal{G}\right](\omega) - \mathbb{E}\left[f^-(X) \mid \mathcal{G}\right](\omega) = \\
&= \int_{\mathbb{R}} f^+(x) \kappa_{X,\mathcal{G}}(\omega, dx) - \int_{\mathbb{R}} f^-(x) \kappa_{X,\mathcal{G}}(\omega, dx) = \\
&= \int_{\mathbb{R}} (f^+(x) - f^-(x)) \kappa_{X,\mathcal{G}}(\omega, dx) = \int_{\mathbb{R}} f(x) \kappa_{X,\mathcal{G}}(\omega, dx)
\end{aligned}$$

Jeżeli $G = \sigma(Z)$, to pojęcie rrw troszkę się upraszcza (z naciskiem na trochę):

Definicja 3.3.
Dla zmiennych losowych X, Y funkcję $\kappa_{X,Y} : \mathbb{R} \times \mathrm{Bor}(\mathbb{R}) \to [0, 1]$ nazywamy rrw X pod warunkiem Y, jeżeli:

(P1) Dla każdego $B \in \mathrm{Bor}(\mathbb{R})$ funkcja $\kappa_{X,Y}(\bullet, B)$ jest borelowska

(P2) Dla każdego $y \in \mathbb{R}$, $\kappa_{X,Y}(y, \bullet)$ jest rozkładem prawdopodobieństwa na prostej.

(P3) $\mathbb{P}\left[X \in B \mid Y = y\right] = \kappa_{X,Y}(y, B)$

3.2 Zadania

Zadanie 1.
Niech Y i Z będą dowolnymi zmiennymi losowymi. Pokaż, że jeżeli zmienna Y jest $\sigma(Z)$-mierzalna,

to istnieje borelowska funkcja $h : \mathbb{R} \to \mathbb{R}$ taka, że $Y = h(Z)$.

Rozwiązanie.
Treść dowodu faktu 3.1.

Zadanie 2.
Pokaż, że dla zmiennych X i Y takich, że $\mathbb{E}\left[X^2\right], \mathbb{E}\left[Y^2\right] < \infty$ i σ-ciała $\mathcal{G} \subseteq \mathcal{F}$ zachodzi

$$|\mathbb{E}\left[XY \mid \mathcal{G}\right]| \leq [\mathbb{E}\left[X^2 \mid \mathcal{G}\right]]^{1/2}[\mathbb{E}\left[Y^2 \mid \mathcal{G}\right]]^{1/2}$$

Rozwiązanie.
Patrz dowód twierdzenia 3.2.

Zadanie 3.
Niech $\kappa_{X,\mathcal{G}}$ będzie regularnym rozkładem warunkowym X pod warunkiem σ-ciała $\mathcal{G} \subseteq \mathcal{F}$. Pokaż, że dla każdej funkcji $f : \mathbb{R} \to \mathbb{R}$ takiej, że $\mathbb{E}\left[|f(X)|\right] < \infty$ zachodzi

$$\mathbb{E}\left[f(X) \mid \mathcal{G}\right](\omega) = \int_{\mathbb{R}} f(x)\kappa_{X,\mathcal{G}}(\omega, dx)$$

Rozwiązanie.
Kolejne rozwiązanie jako dowód faktu 3.5.

Zadanie 4.
(Nierówność Jensena) Dana jest funkcja wypukła $\varphi : \mathbb{R} \to \varphi$, przestrzeń probabilistyczna $(\Omega, \mathcal{F}, \mathbb{P})$ oraz $\mathcal{G}$ pod-σ-ciało $\mathcal{F}$. Załóżmy, że zmienne losowe X i $\varphi(X)$ są całkowalne. Pokaż, że

$$\varphi(\mathbb{E}\left[X \mid \mathcal{G}\right]) \leq \mathbb{E}\left[\varphi(X) \mid \mathcal{G}\right]$$

Rozwiązanie.
Korzystając z faktu 3.5 możemy powiedzieć, że

$$\begin{aligned}\mathbb{E}\left[\varphi(X) \mid \mathcal{G}\right](\omega) &= \int_{\mathbb{R}} \varphi(x)\kappa_{X,\mathcal{G}}(\omega, dx) \geq \\ &\geq \varphi\left[\int_{\mathbb{R}} \kappa_{X,\mathcal{G}}(\omega, dx)\right] = \varphi(\mathbb{E}\left[X \mid \mathcal{G}\right])\end{aligned}$$

nierówność wynika z twierdzenia Jensena dla całek (które mówi, że $\int \varphi \circ f \, d\mu \geq \varphi\left(\int f \, d\mu\right)$) a ostatnie przejście to ponowne zastosowanie faktu 3.5, tym razem dla $\mathbb{E}\left[id(X) \mid \mathcal{G}\right]$.

Zadanie 5.
Załóżmy, że wektor losowy (X, Y) ma dwuwymiarowy rozkład normalny.

(a) Znajdź $a \in \mathbb{R}$ takie, że zmienne $X - aY$ i Y są niezależne.

(b) Pokaż, że

$$\mathbb{E}\left[X \mid Y\right](\omega) = \mu_X + \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}(Y(\omega) - \mu_Y),$$

gdzie $\mu_X = \mathbb{E}[X]$ oraz $\mu_Y = \mathbb{E}[Y]$.

(c) Dla $y \in \mathbb{R}$ znajdź rozkład X pod warunkiem $Y = y$.

Rozwiązanie.

(a) Z Rachunku Prawdopodbieństwa 1R wiemy, że jeśli wektor losowy ma rozkład normalny, a jego poszczególne elementy są nieskorelowane, to są one również niezależne. Patrzymy więc na kowariancję

$$\mathrm{Cov}(X - aY, Y) = \mathrm{Cov}(X, Y) - a\mathrm{Cov}(Y, Y) = \mathrm{Cov}(X, Y) - a\mathrm{Var}(Y)$$

i przyrównujemy ją do 0

$$0 = \mathrm{Cov}(X, Y) - a\mathrm{Var}(Y) \Rightarrow a = \frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}(Y)}$$

(b) Zauważmy, że korzystając z poprzedniego punktu możemy przepisać równość jako

$$\mathbb{E}\left[X \mid Y\right](\omega) = \mathbb{E}[X] - a\mathbb{E}[Y] + aY(\omega) = \mathbb{E}[X - aY] + a\mathbb{E}\left[Y \mid Y\right]$$

gdzie $Y = \mathbb{E}\left[Y \mid Y\right]$, bo Y jest $\sigma(Y)$-mierzalne.

Kolejne przekształcenia dają

$$\mathbb{E}\left[X \mid Y\right] - a\mathbb{E}\left[Y \mid Y\right] = \mathbb{E}[X - aY]$$

co jest prawdą, gdyż po skorzystaniu z liniowości wwo po lewej stronie mamy

$$\mathbb{E}\left[X - aY \mid Y\right] = \mathbb{E}[X - aY]$$

a ponieważ $X - aY$ dobraliśmy tak, żeby było niezależne od Y, to jest ono również niezależne od $\sigma(Y)$. Czyli wwo jest równe wartości oczekiwanej $X - aY$.

(c) Rozkład X pod warunkiem $Y = y$ to $\kappa_{X,Y}(y, B) = \mathbb{P}\left[X \in B \mid Y = y\right]$. Wiemy, że

$$\mathbb{P}\left[X \in B \mid Y = y\right] = \frac{\mathbb{P}[X \in B \text{ i } Y = y]}{\mathbb{P}[Y = y]}$$

czyli można wydedukować, że szukamy

$$\mathbb{P}\left[X \mid Y = y\right] = f_{X|Y=y}(x, y) = \frac{f(x, y)}{f_Y(y)}$$

co jest zbyt dużą liczbą brzydkich obliczeń żeby nawet mi się chciało to dokładnie spisywać. Wystarczy podstawić pod gęstość (X, Y) na górze i do gęstości Y na dole.

Wykład 30.10.23 : Martyngały

Mają coś wspólnego z jazdą konną (podobno).

Przykład(y) 4.1

1. Rozważmy dowolną grę i dla uproszczenia niech polega ona na rzucaniu monetą, na której orzeł wypada z $\mathbb{P} = p \in (0,1)$. Obstawiamy według zasady double or nothing:

   - jeśli wypada orzeł, to podwajamy nasz kapitał
   - jeżeli wypada reszka, to tracimy wszystko

   Czy taka gra jest sprawiedliwa?

   Rozważmy ciąg niezależnych zmiennych losowych $\{\xi_k\}_{k\in\mathbb{N}}$ o tym samym rozkładzie
   $$\mathbb{P}\left[\xi_k = 2\right] = 1 - \mathbb{P}\left[\xi_k = 0\right] = p.$$
   Wówczas ciąg
   $$X_n = \xi_n \cdot \xi_{n-1} \cdot \ldots \cdot \xi_1 \cdot X_0$$
   reprezentuje stan konta po n-tym rzucie, gdzie X_0 to jakaś stała.

   Rozważmy σ-ciało $\mathcal{F}_n = \sigma(\xi_n, \ldots, \xi_1, X_0)$ generowane przez pierwszych n rzutów i stan początkowy. Zadajmy sobie teraz pytanie, ile wynosi
   $$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right]?$$
   Mamy zależność rekurencyjną $X_{n+1} = \xi_{n+1}X_n$, stąd możemy powiedzieć, że
   $$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[\xi_{n+1}X_n \mid \mathcal{F}_n\right]$$
   samo X_n jest w $\mathcal{F}_n$, więc możemy je wyciągnąć przed $\mathbb{E}$. Dodatkowo, ξ_{n+1} jest niezależne od $\mathcal{F}_n$, więc
   $$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = X_n\mathbb{E}\left[\xi_{n+1} \mid \mathcal{F}_n\right] = X_n\mathbb{E}\left[\xi_{n+1}\right] = (2p) \cdot X_n$$

   - Jeżeli $p > \frac{1}{2}$, to wówczas
     $$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] \geq X_n$$
     i wtedy taka gra jest korzystna, bo z coraz to kolejnym rzutem oczekiwania rosną.
   - Jeżeli $p < \frac{1}{2}$, to wówczas
     $$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] \leq X_n$$
     i gra jest korzystna, ale dla kasyna a nie gracza.
   - Jeżeli $p = \frac{1}{2}$, to wówczas
     $$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = X_n$$
     i w takim przypadku powiemy, że gra jest sprawiedliwa.

Ten ostatni, uczciwy przypadek to jest jeden ze sposobów, na które możemy myśleć o martyngałach.

Definicja 4.1 : o martyngałach słów kilka.

- Wstępującą rodzinę σ-ciał
  $$\mathbb{F} = \{\mathcal{F}_n\}_{n\in\mathbb{N}},$$

$\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$, nazywamy **filtracją**

- Powiemy, że ciąg zmiennych losowych $\{X_n\}_{n\in\mathbb{N}}$ jest **$\mathbb{F}$-adaptowalny**, jeżeli $X_n \in \mathcal{F}_n$.
- Adaptowalny i całkowalny ciąg $\{X_n\}$ nazywamy *nadmartyngałem*, jeśli

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] \leq X_n$$

- Adaptowalny i całkowalny ciąg $\{X_n\}$ nazywamy *podmartyngałem*, jeśli

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] \geq X_n$$

- Z kolei ciąg $\{X_n\}$ jest **martyngałem**, jeśli jest jednocześnie nadmartyngałem i podmartyngałem, czyli zachodzi równość

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = X_n$$

Przykład(y) 4.2

1. Niech $\{\eta_k\}_{k\in\mathbb{N}}$ będą niezależne takie, że $\mathbb{E}\left[\eta_k\right] = 0$ dla każdego $k \in \mathbb{N}$. Wówczas jako filtrację możemy rozważyć
$$\mathcal{F}_n = \sigma(\eta_1, \ldots, \eta_n)$$
a jako nowy ciąg zmiennych losowych zdefiniujemy jako $M_0 = 0$ i
$$M_n = \sum_{k=1}^{n} \eta_k.$$
Tak zdefiniowany ciąg $\{M_n\}$ jest $\mathbb{F} = \{\mathcal{F}_n\}$-martyngałem:
$$\begin{aligned}\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[\eta_{n+1} + M_n \mid \mathcal{F}_n\right] = \\ &= \mathbb{E}\left[\eta_{n+1} \mid \mathcal{F}_n\right] + \mathbb{E}\left[M_n \mid \mathcal{F}_n\right] = \\ &= \mathbb{E}\left[\eta_{n+1}\right] + M_n = 0 + M_n = M_n\end{aligned}$$

2. Dla dowolnej filtracji $\mathbb{F} = \{\mathcal{F}_n\}$ i całkowalnej zmiennej losowej X rozważmy
$$M_n = \mathbb{E}\left[X \mid \mathcal{F}_n\right].$$
Wówczas
$$\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{F}_{n+1}\right] \mid \mathcal{F}_n\right] = \mathbb{E}\left[X \mid \mathcal{F}_n\right] = M_n$$

Uwaga 4.1.

Jeżeli $\{X_n\}$ jest martyngałem, to
$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = X_n$$
czyli mam dwie zmienne losowe które są sobie równe, czyli
$$\mathbb{E}\left[X_{n+1}\right] = \mathbb{E}\left[\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right]\right] = \mathbb{E}\left[X_n\right]$$

W szczególności, jeśli zastosujemy indukcję, to dostaniemy, że dla dowolnego $n \in \mathbb{N}$

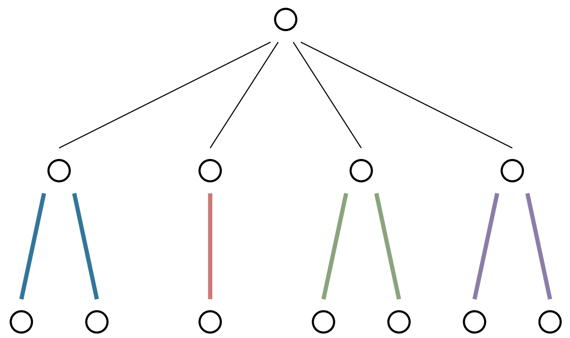
$$\mathbb{E}[X_n] = \mathbb{E}[X_0]$$

Przykład(y) 4.3

1. **Proces Galtona-Watsona**

   Rozważmy populację, w której osobniki rozmnażają się bezpłciowo, niezależnie od siebie. Można myśleć o tym jako o obserwacji populacji pantofelków z pomiarami w jakiś określonych odstępach czasu.

   Myślimy o tym jako o drzewie, w którym liczba krawędzi z danego wierzchołka oznacza liczbę potomstwa, a ilość wierzchołków na danej głębokości oznacza ilość pantofelków po n-tym pokoleniu.



   Niech μ będzie dowolnym rozkładem prawdopodobieństwa na $\mathbb{N} = \{0, 1, \ldots\}$. Rozważmy zmienne losowe losowe indeksowane parami liczb naturalnych $\{Y_{n,k}\}_{n,k\in\mathbb{N}}$. Kładziemy

$$Z_1 = 1$$

$$Z_{n+1} = \sum_{k=1}^{Z_n} Y_{n+1,k}$$

   Z_1 to liczba pantofelków na samym początku, $Z_2 = Y_{2,1}$ to liczba dzieci pierwszego pantofelka, z kolei

$$Z_3 = Y_{3,1} + Y_{3,2} + Y_{3,3} + Y_{3,4},$$

   co odpowiada kolorom na rysunku. To znaczy, że $Y_{n+1,k}$ to liczba potomstwa w generacji $n+1$ zrodzona z k-tego pantofelka w generacji n.

   Filtracją będzie dla nas ciąg o elementach $\mathcal{F}_n = \sigma(Y_{j,k} \ : \ k \in \mathbb{N}, j \leq n)$. Chcemy zapytać się o wartość oczekiwaną Z_{n+1}

$$\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[\sum_{k=1}^{Z_n} Y_{n+1,k} \mid \mathcal{F}_n\right] = h(Z_n)$$

   gdzie

$$h(z) = \mathbb{E}\left[\sum_{k=1}^{z} Y_{n+1,k}\right] = z \cdot \underbrace{\mathbb{E}\left[Y_{n,k}\right]}_{m} = m \cdot z,$$

   bo wszystkie $Y_{n,k}$ mają taką samą średnią. Oznacza to, że

$$\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] = m \cdot Z_n.$$

Jeżeli $m < 1$, to dostajemy w ten sposób nadmartyngał, jeśli $m > 1$ to mamy podmartyngał, a w krytycznym przypadku $m = 1$, to $\{Z_n\}$ jest martyngałem.

Jeśli pomnożymy

$$\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] = mZ_n$$

oboma stronami przez m^{-n-1}, to dostajemy

$$\mathbb{E}\left[m^{-n-1}Z_{n+1} \mid \mathcal{F}_n\right] = m^{-n}Z_n$$

i wtedy $W_n = m^{-n}Z_n$ jest zawsze martyngałem, bo

$$\mathbb{E}\left[W_{n+1} \mid \mathcal{F}_n\right] = W_n$$

4.1 Transformata martyngałowa

Stan konta gracza wynosi X_n po n-tej sprawiedliwej grze. Przychodzi drugi gracz i obstawia on wyniki w grze tego pierwszego. Wypłata drugiego gracza wynosi $B_n \cdot (X_n - X_{n-1})$, tzn. za każdy przychód pierwszego gracza dostaje jakąś część tej wygranej.

Dla ciągu funkcji $B_n \in \mathcal{F}_{n+1} = \sigma(X_0, ..., X_{n-1})$. Żeby było łatwiej, niech drugi gracz zaczyna z tym samym kapitałem co pierwszy. Stan konta drugiego gracza po n-tej grze wynosi

$$W_n = \sum_{k=1}^{n} B_k \cdot (X_k - X_{k-1}) + X_0.$$

Tak zdefiniowany ciąg $\{Q_n\}$ jest również martyngałem, bo

$$\begin{aligned}\mathbb{E}\left[W_{n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[B_{n+1} \cdot (X_{n+1} - X_n) \mid \mathcal{F}_n\right] + \mathbb{E}\left[W_n \mid \mathcal{F}_n\right] = \\ &= B_{n+1}\mathbb{E}\left[X_{n+1} - X_n \mid \mathcal{F}_n\right] + W_n = \\ &= B_{n+1}(\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] - X_n) + W_n = W_n\end{aligned}$$

bo X_n sam w sobie był martyngałem, więc $X_n = \mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right]$.

Definicja 4.2.
Niech $\mathbb{F}$ będzie filtracją. Zmienną losową $T : \Omega \to \mathbb{N} \cup \{+\infty\}$ nazywamy **$\mathbb{F}$-czasem zatrzymania**, jeżeli zdarzenie $\{T = n\}$ jest mierzalne względem $\mathcal{F}_n$ dla każdego $n \in \mathbb{N}$.

Przykład(y) 4.4

1. Rzucamy 10-krotnie monetą. Zdefiniujmy zmienną losową

$$X_n = \begin{cases} 1 & \text{orzeł w n-tym} \\ 0 & \text{wpp} \end{cases}$$

   Filtracją niech będzie $\mathcal{F}_n = \sigma(X_1, ..., X_n)$. Jeśli T będzie momentem wypadnięcia pierwszego orła, a S - wypadnięcia ostatniego orła, to T jest czasem zatrzymania, bo

$$\{T = n\} = \{X_1 = 0, X_2 = 0, ..., X_n = 1\} \in \mathcal{F}_n$$

   a S nim nie jest, bo wymaga informacji wybiegającej w przyszłość:

$$\{S = n\} = \{X_n = 1, X_{n+1} = 0, ...\} \notin \mathcal{F}_n$$

2. Rozważmy $\mathbb{F}$-adaptowalny ciąg zmiennych losowych $\{X_n\}$. Dla $B \in \mathrm{Bor}(\mathbb{R})$ kładziemy

$$T(B) = \inf\{n \ : \ X_n \in B\}.$$

Tak zdefiniowane T jest czasem zatrzymania:

$$\{T = n\} = \{X_1 \notin B, ..., X_{n-1} \notin B, X_n \in B\} \in \mathcal{F}_n$$

3. Jeżeli $T = n_0$ dla pewnego $n_0 \in \mathbb{N}$, to taka stała funkcja nadal jest czasem zatrzymania, bo

$$\{T = n\} = \begin{cases} \emptyset & n \neq n_0 \\ \Omega & n = n_0 \end{cases}$$

4.2 Zadania

Zadanie 1.
Załóżmy, że $\{X_n\}_{n\in\mathbb{N}}$ jest ciągiem niezależnych zmiennych losowych o takim samym rozkładzie, średniej 0 i skończonej wariancji. Rozważmy filtrację $\mathbb{F} = \{\mathcal{F}_n\}$ zadaną przez $\mathcal{F}_n = \sigma(X_0, X_1, ..., X_n)$. Udowodnij, że ciąg

$$Z_n = X_0X_1 + X_1X_2 + ... + X_{n-1}X_n, \quad Z_0 = 0$$

jest $\mathbb{F}$-martyngałem.

Rozwiązanie.
Tak naprawdę to sprawdzenie że Z_n są całkowalne powinno być na samym początku, bo inaczej to nie ma sensu.

Chcemy pokazać, że

$$\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] = Z_n$$

dla dowolnego $n \in \mathbb{N}$.

$$\begin{aligned}\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[Z_n + X_nX_{n+1} \mid \mathcal{F}_n\right] = \\ &= \mathbb{E}\left[Z_n \mid \mathcal{F}\right] + \mathbb{E}\left[X_nX_{n+1} \mid \mathcal{F}_n\right] = \\ &= Z_n + X_n\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right]\end{aligned}$$

ponieważ X_n jest $\mathcal{F}_n$-mierzalne oraz

$$\mathbb{E}\left[|X_nX_{n+1}|\right] \leq \mathbb{E}\left[X_n^2\right]^{1/2} \mathbb{E}\left[X_{n+1}^2\right]^{1/2} < \infty$$

gdzie nierówność wynika z nierówności Cauchy'ego-Schwarza, a $\mathbb{E}\left[X_n^2\right] = \mathrm{Var}(X_n) < \infty$.

Zauważmy teraz, że X_{n+1} jest niezależne od $\mathcal{F}_n$, gdyż X_n jest niezależne od każdej ze zmiennych $X_1, ..., X_n$. W takim razie, $\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[X_{n+1}\right] = 0$, a więc ostatecznie dostajemy

$$\mathbb{E}\left[Z_{n+1} \mid \mathcal{F}_n\right] = Z_n + X_n\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = Z_n + X_n \cdot 0 = Z_n$$

Czyli Z_n faktycznie jest martyngałem.

Zadanie 2.
Ustalmy $\theta \in \mathbb{R}$. Niech $X_1, X_2, \ldots$ będzie ciągiem niezależnych zmiennych losowych o takim samym rozkładzie takich, że

$$\mathbb{E}\left[e^{\theta X_1}\right] < \infty.$$

Pokaż, że

$$M_n = \mathbb{E}\left[e^{\theta X_1}\right]^{-n} \prod_{j=1}^{n} e^{\theta X_j}$$

jest $\mathbb{F}$-martyngałem dla filtracji $\mathbb{F} = \{\mathcal{F}_n\}$ danej przez $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

Rozwiązanie.
Tutaj ta sama uwaga co do poprzedniego zadania.

Zacznijmy od obserwacji, że

$$M_{n+1} = \mathbb{E}\left[e^{\theta X_1}\right]^{-n-1} \prod_{j=1}^{n+1} e^{\theta X_j} = M_n \cdot \mathbb{E}\left[e^{\theta X_1}\right]^{-1} e^{\theta X_{n+1}}$$

w takim razie, wwo M_{n+1} to jest

$$\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[e^{\theta X_1}\right]^{-1} \cdot \mathbb{E}\left[M_n \cdot e^{\theta X_{n+1}} \mid \mathcal{F}_n\right]$$

Od razu widać, że M_n jest mierzalne względem $\mathcal{F}_n$, bo zależy tylko od zmiennych $X_1, \ldots, X_n$ które $\mathcal{F}_n$ generują. Chcemy teraz sprawdzić, czy $\mathbb{E}\left[|M_n \cdot e^{\theta X_{n+1}}|\right] < \infty$, wówczas możemy wyciągnąć M_n przed wwo.

$$\begin{aligned}
\mathbb{E}\left[|M_n \cdot e^{\theta X_{n+1}}|\right] &= \mathbb{E}\left[\left|\mathbb{E}\left[e^{\theta X_1}\right]^{-1} \cdot \prod_{j=1}^{n} e^{\theta X_j} \cdot e^{\theta X_{n+1}}\right|\right] = \\
&= |\mathbb{E}\left[e^{\theta X_1}\right]|^{-n} \cdot \mathbb{E}\left[\prod_{j=1}^{n+1} e^{\theta X_j}\right] = \\
&= |\mathbb{E}\left[e^{\theta X_1}\right]|^{-n} \cdot \prod_{j=1}^{n+1} \mathbb{E}\left[e^{\theta X_j}\right] = \\
&= \mathbb{E}\left[e^{\theta X_{n+1}}\right] = \mathbb{E}\left[e^{\theta X_1}\right] < \infty
\end{aligned}$$

ponieważ jeśli $\{X_n\}$ są niezależne, to e^{X_n} też są niezależne, a dla niezależnych X, Y zachodzi $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$. W takim razie dostajemy

$$\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[e^{\theta X_1}\right]^{-1} \mathbb{E}\left[M_n \cdot e^{\theta X_{n+1}} \mid \mathcal{F}_n\right] = M_n \mathbb{E}\left[e^{\theta X_1}\right]^{-1} \mathbb{E}\left[e^{\theta X_{n+1}} \mid \mathcal{F}_n\right]$$

ale ponieważ $\mathcal{F}_n$ nie zawiera ani grama informacji o X_{n+1}, to $e^{\theta X_{n+1}}$ jest niezależne od $\mathcal{F}_n$, więc

$$\mathbb{E}\left[e^{\theta X_{n+1}} \mid \mathcal{F}_n\right] = \mathbb{E}\left[e^{\theta X_{n+1}}\right] = \mathbb{E}\left[e^{\theta X_1}\right]$$

a to już daje to co chcieliśmy.

Zadanie 3.
Niech $\{Y_n\}$ będzie ciągiem niezależnych zmiennych losowych o średniej 0 i wariancji σ^2. Pokaż, że ciąg

$$X_n = \left(\sum_{k=1}^{n} Y_k\right)^2 - n\sigma^2$$

jest martyngałem.

Rozwiązanie.
Tutaj powinniśmy sprawdzić jeszcze $\mathbb{F}$-adaptowalność $\{X_n\}$ i jego całkowalność.

Zacznijmy od wyrażenia X_{n+1} przy użyciu X_n

$$\begin{aligned}X_{n+1} &= \left(\sum_{k=1}^{n+1} Y_k\right)^2 - (n+1)\sigma^2 = \left(\sum_{k=1}^{n} Y_k + Y_{n+1}\right)^2 - (n+1)\sigma^2 = \\ &= \left(\sum_{k=1}^{n} Y_k\right)^2 - n\sigma^2 + 2Y_{n+1}\left(\sum_{k=1}^{n} Y_k\right) + Y_{n+1}^2 - \sigma^2 = \\ &= X_n + 2\left(\sum_{k=1}^{n} Y_{n+1}Y_k\right) - \sigma^2 + Y_{n+1}^2\end{aligned}$$

Rozważmy teraz filtrację $\mathbb{F} = \{\mathcal{F}_n\}$ dla ciągu $\mathcal{F}_n = \sigma(Y_1, ..., Y_n)$. Wtedy mamy

$$\begin{aligned}\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[X_n + 2\left(\sum_{k=1}^{n} Y_{n+1}Y_k\right) - \sigma^2 + Y_{n+1}^2 \mid \mathcal{F}_n\right] = \\ &= \mathbb{E}\left[X_n \mid \mathcal{F}_n\right] + 2\sum \mathbb{E}\left[Y_{n+1}Y_n \mid \mathcal{F}_n\right] - \sigma^2 + \mathbb{E}\left[Y_{n+1}^2 \mid \mathcal{F}_n\right] = (\star)\end{aligned}$$

Zmienna X_n jest $\mathcal{F}_n$-mierzalna bo korzysta tylko z informacji zapisywanych przez $Y_1, ..., Y_n$. Zmienna Y_{n+1} jest niezależna od $\mathcal{F}_n$, bo zmienne $\{Y_n\}$ są niezależne. Zmienna $Y_{n+1}Y_k$ jest całkowalna dla dowolnego k, bo $\mathbb{E}\left[|Y_{n+1}Y_k|\right] = \mathbb{E}\left[|Y_{n+1}|\right]\mathbb{E}\left[|Y_k|\right] = 0$. W takim razie wracając do równości wyżej, można napisać

$$\begin{aligned}(\star) &= X_n + 2\sum Y_k\mathbb{E}\left[Y_{n+1} \mid \mathcal{F}_n\right] - \sigma^2 + \mathbb{E}\left[Y_{n+1}^2\right] = \\ &= X_n + 2\sum Y_k\mathbb{E}\left[Y_{n+1}\right] - \sigma^2 + \sigma^2 = X_n + 2\sum Y_k \cdot 0 = X_n\end{aligned}$$

W takim razie ciąg $\{X_n\}$ jest $\mathbb{F}$-martyngałem.

Zadanie 4.

Niech $\{X_n\}$ będzie $\mathbb{F}$-martyngałem. Pokaż, że dla każdych naturalnych m, n

$$\mathbb{E}\left[X_{m+n} \mid \mathcal{F}_n\right] = X_n.$$

Rozwiązanie.
Dla $m = 1$ działa z definicji martyngału.

Zakładamy teraz, że $\mathbb{E}\left[X_{m+n} \mid \mathcal{F}_n\right] = X_n$ i chcemy to samo dostać dla $m + 1$ (indukcja).

Z własnośći wwo 2.4 wiemy, że jeśli mamy dwa zawarte w sobie σ-ciała $\mathcal{G}_1 \subseteq \mathcal{G}_2$, to

$$\mathbb{E}\left[\mathbb{E}\left[X \mid \mathcal{G}_2\right] \mid \mathcal{G}_1\right] = \mathbb{E}\left[X \mid \mathcal{G}_1\right]$$

tutaj bierzemy $\mathcal{G}_1 = \mathcal{F}_n$ natomiast $\mathcal{G}_2 = \mathcal{F}_{n+m}$. Mamy więc

$$\mathbb{E}\left[\mathbb{E}\left[X_{m+n+1} \mid \mathcal{F}_{m+n}\right] \mid \mathcal{F}_n\right] = \mathbb{E}\left[X_{m+n+1} \mid \mathcal{F}_n\right]$$

Z faktu, że $\{X_n\}$ jest martyngałem to mamy

$$\mathbb{E}\left[X_{m+n+1} \mid \mathcal{F}_{n+m}\right] = X_{m+n}$$

czyli przechodząc już do sedna sprawy,

$$\begin{aligned}\mathbb{E}\left[X_{m+n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[\mathbb{E}\left[X_{m+n+1} \mid \mathcal{F}_{m+n}\right] \mid \mathcal{F}_n\right] = \\ &= \mathbb{E}\left[X_{m+n} \mid \mathcal{F}_n\right] = X_n\end{aligned}$$

bo ostatnie przejście wynika z założenia indukcyjnego.

Zadanie 5.
Niech $\{X_n\}$ będzie nadmartyngałem takim, że $\mathbb{E}[X_n] = \mathbb{E}[X_0] < \infty$ dla każdego $n \in \mathbb{N}$. Pokaż, że $\{X_n\}$ jest martyngałem.

Rozwiązanie.
Nadmartyngał spełnia

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] \leq X_n$$

Po pierwsze zauważmy, że

$$\mathbb{E}\left[X_n - \mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right]\right] = \mathbb{E}[X_n] - \mathbb{E}\left[\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right]\right] = \mathbb{E}[X_n] - \mathbb{E}[X_{n+1}] = 0,$$

czyli $\mathbb{E}\left[X_n - \mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right]\right] = 0$, a ponieważ z warunku na nadmartyngał $X_n - \mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] \geq 0$ prawie wszędzie, to mamy nieujemną funkcję której całka jest 0 $\Rightarrow$ ta funkcja jest prawie wszędzie równa zero.

Zadanie 6.
Pokaż, że jeżeli T_1, T_2 są czasami zatrzymania, to $\min\{T_1, T_2\}$ i $\max\{T_1, T_2\}$ również są czasami zatrzymania. Czy $T_1^2, T_1 + 1, T_1 + T_2, T_1 - 1, \min\{T_1, 2T_2\}$ są też czasami zatrzymania?

Rozwiązanie.

Przychodzimy do tego zadania z wiedzą, że zdarzenie $\{T_i = n\}$ jest w zbiorze $\mathcal{F}_n$ dla dowolnego n. Od razu zauważmy, że skoro $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq ...$, to jeśli $\{T_i = n\} \in \mathcal{F}_n$ dla dowolnego n, to $\mathcal{F}_n$ nadal trzyma informację o $\{T_i = k\}$ dla $k \leq n$, czyli możemy równoważnie powiedzieć, że $\{T_i \leq n\} \in \mathcal{F}_n$.

Zdarzenie $\{\max\{T_1, T_2\} \leq n\}$ jest równoważne zdarzeniu

$$\{\max\{T_1, T_2\} \leq n\} = \{T_1 \leq n \text{ i } T_2 \leq n\} = \{T_1 \leq n\} \cap \{T_2 \leq n\}$$

a ponieważ $\mathcal{F}_n$ jest σ-ciałem, to skoro dwa zbiory do niego należą, to należeć też musi ich przekrój.

Zdarzenie $\{\min\{T_1, T_2\} = n\}$ potencjalnie pyta o czas w nieskończonej przyszłości, bo możemy je zapisać jako

$$\min\{T_1, T_2\} = n \iff (T_1 = n \ \wedge T_2 \geq n) \ \vee \ (T_1 \geq n \ \wedge \ T_2 = n)$$

ale jeśli popatrzymy na zdarzenie

$$\{\min\{T_1, T_2\} \leq n\} = \{T_1 \leq n \text{ lub } T_2 \leq n\} = \{T_1 \leq n\} \cup \{T_2 \leq n\}$$

wiemy, że oba składniki tej sumy są w $\mathcal{F}_n$, więc i cała suma w $\mathcal{F}_n$ siedzi.

1. T_1^2 jest czasem zatrzymania, bo

   $$\{T_1^2 \leq n\} = \{T_1 \leq \sqrt{n} \leq n\}$$

   jest elementem $\mathcal{F}_n$.

2. $T_1 + 1$ jest czasem zatrzymania, bo

   $$\{T_1 + 1 = n\} = \{T_1 = n - 1\}$$

   co w oczywisty sposób daje nam $\{T_1 = n - 1\} \in \mathcal{F}_{n-1} \subseteq \mathcal{F}_n$

3. $T_1 + T_2$ jak najbardziej jest czasem zatrzymania, bo dla $i \neq j$ zbiory $\{T_1 = i \ \wedge \ T_2 = n - i\}$ oraz $\{T_1 = j \ \wedge \ T_2 = n - j\}$ są rozłączne, a więc

   $$\{T_1 + T_2 = n\} = \bigcup_{k \leq n} \{T_1 = k \text{ i } T_2 = n - k\}$$

   każdy element sumy należy do $\mathcal{F}_n$ jak przekrój dwóch jego elementów, więc i cała suma do niego należy.

4. $T_1 - 1$ nie koniecznie jest czasem zatrzymania, bo

   $$\{T_1 - 1 = n\} = \{T_1 = n + 1\}$$

   może należeć tylko do $\mathcal{F}_{n+1}$, ale nie trafiać w $\mathcal{F}_n$. Rozważmy choćby modelowanie gry w której wygrywamy i przegrywamy dolara z takim samym prawdopodobieństwem (wykład następny), ale dorzućmy możliwość nie obstawiania, czyli "kroku 0". Załóżmy, że startujemy z zerem dolarów. Innymi słowy, rozważamy skakanie w losowy sposób po $\mathbb{Z}$ z możliwością podskoczenia tylko pionowo w górę. Wtedy jeśli X_i będzie naszym aktualnym położeniem na $\mathbb{Z}$ po i krokach, a $\mathcal{F}_i = \sigma(X_1, ..., X_i)$, to $\mathcal{F}_1$ będzie posiadało tylko informację o zdarzeniu $\{X_1 = -1\}$, $\{X_1 = 0\}$ i $\{X_1 = 1\}$, a więc gdy $T_1 = \inf\{n : X_n = 1\}$, to zdarzenie

$$\{T_1 - 1 = 1\} = \{T_1 = 2\} = \{X_1 = 1\} \cup \{X_1 = 0 \text{ i } X_2 = 1\}$$

należy do $\mathcal{F}_2$, ale do $\mathcal{F}_1$ już należeć nie ma możliwości.

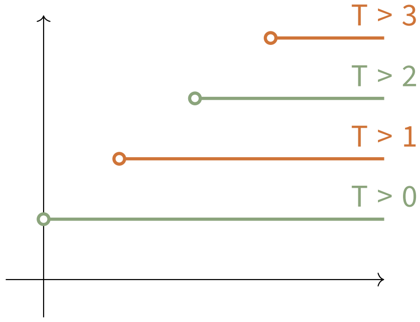
Zadanie 7.
Niech T będzie czasem zatrzymania. Załóżmy, że istnieje $\varepsilon \in (0, 1)$ oraz $N \in \mathbb{N}$ takie, że

$$\mathbb{P}\left[T \leq N + n \mid \mathcal{F}_n\right] > \varepsilon$$

dla każdego $n \in \mathbb{N}$. Pokaż, że $\mathbb{E}[T] < \infty$.

Rozwiązanie.
Zaczniemy od obserwacji, że

$$\mathbb{E}[T] = \sum_{n\geq 0} n\mathbb{P}[T = n] = \sum_{n\geq 0}\sum_{k=0}^{n-1} \mathbb{P}[T = n] = \sum_{n\geq 0} \mathbb{P}[T > n]$$



Będziemy chcieli oszacować całe to wyrażenie od góry przez zbieżny szereg. Dokładnie, będzie nam potrzebne oszacowanie go przez $\mathbb{P}[T > kN]$:

$$\mathbb{E}[T] = \sum_{n\geq 0} \mathbb{P}[T > n] = \sum_{k\geq 0}\sum_{n=kN}^{(k+1)N-1} \mathbb{P}[T > n] \leq \sum_{k\geq 0} N \cdot \mathbb{P}[T > kN] \quad (☕)$$

gdyż funkcja $\mathbb{P}[T > n]$ jest malejąca, a więc przyjmuje ekstremum na lewym krańcu. Dzielimy więc $\sum \mathbb{P}[T > n]$ na odcinki $[kN, (k+1)N)$ i szacujemy sumę na każdym z nich przez $N \cdot \mathbb{P}[T > kN]$.

Żeby jednak dokończyć szacowanie wyżej, potrzebujemy ograniczyć $\mathbb{P}[T > kN]$ od góry. Z definicji prawdopodobieństwa warunkowego wiemy, że

$$1 \geq \mathbb{P}\left[T \leq N + n \mid \mathcal{F}_n\right] = \mathbb{E}\left[\mathbb{1}_{\{T\leq N+n\}} \mid \mathcal{F}_n\right] > \varepsilon$$

Funkcja $\mathbb{1}_{\{T\leq N+n\}} = 1 - \mathbb{1}_{\{T>N+n\}}$, więc mamy

$$1 \geq \mathbb{E}\left[1 - \mathbb{1}_{\{T>N+n\}} \mid \mathcal{F}_n\right] = \mathbb{E}\left[1 \mid \mathcal{F}_n\right] - \mathbb{E}\left[\mathbb{1}_{\{T>N+n\}} \mid \mathcal{F}_n\right] = 1 - \mathbb{E}\left[\mathbb{1}_{\{T>N+n\}} \mid \mathcal{F}_n\right] > \varepsilon$$

Po przerzuceniu ε na lewą stronę, a wwo na prawą stronę, dostajemy

$$1 - \varepsilon > \mathbb{E}\left[\mathbb{1}_{\{T>N+n\}} \mid \mathcal{F}_n\right] \qquad (\star)$$

co po nałożeniu całki na obie strony implikuje, że

$$1 - \varepsilon > \mathbb{E}\left[\mathbb{1}_{\{T>N+n\}}\right] = \mathbb{P}\left[T > N + n\right]$$

Wróćmy do punktu $(\star)$ i rozważmy zdarzenie

$$G = [\{T \leq n\}]^c = \{T > n\} \in \mathcal{F}_n.$$

Zdarzenie $\{T \leq n\} \in \mathcal{F}_n$, bo T jest czasem zatrzymania, a jego dopełnienie należy do tego zbioru ze względu na fakt, że $\mathcal{F}_n$ jest σ-ciałem. Jeśli nałożymy na obie strony $(\star)$ całkowanie po G, to dostajemy

$$\begin{aligned}
&\mathbb{P}\left[T > n\right](1 - \varepsilon) = \\
&= (1 - \varepsilon)\mathbb{E}\left[\mathbb{1}_G\right] = \\
&= \mathbb{E}\left[(1 - \varepsilon)\mathbb{1}_G\right] > \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{T>N+n\}} \mid \mathcal{F}_n\right]\mathbb{1}_G\right] = \\
&\qquad = \mathbb{E}\left[\mathbb{1}_{\{T>N+n\}}\mathbb{1}_G\right] = \mathbb{E}\left[\mathbb{1}_{\{T>N+n>n\}}\right] = \\
&\qquad = \mathbb{P}\left[T > N + n\right]
\end{aligned}$$

Dostaliśmy więc

$$\mathbb{P}\left[T > n\right](1 - \varepsilon) > \mathbb{P}\left[T > N + n\right]$$

Niech teraz $n = k \cdot N$. Wtedy

$$\begin{aligned}
&\mathbb{P}\left[T > (k + 1)N\right] = \mathbb{P}\left[N + kN\right] < \\
&\qquad < \mathbb{P}\left[T > k \cdot N\right](1 - \varepsilon) = \\
&\qquad\quad = \mathbb{P}\left[T > n\right](1 - \varepsilon) = \mathbb{P}\left[T > N + (k - 1)N\right](1 - \varepsilon) < \\
&\qquad\qquad < (1 - \varepsilon)(1 - \varepsilon)\mathbb{P}\left[T > (k - 1)N\right].
\end{aligned}$$

a w szczególności gdy $n = 0$, to $\mathbb{P}\left[T > 1 \cdot N\right] < (1 - \varepsilon)^1 = 1 - \varepsilon$. Sugeruje to, że zachodzi wzór

$$(1 - \varepsilon)^k > \mathbb{P}\left[T > k \cdot N\right]$$

który udowadniamy indukcyjnie. Dla $k = 0$ sprawa jest oczywista, a dla kroku indukcyjnego $k \Rightarrow k + 1$ zaczynamy od góry i kontynuujemy jak w nierówności wyżej aż do czerwonego $\mathbb{P}\left[T > n\right]$.

Wróćmy teraz do (☕) i dokończmy szacowanie $\mathbb{E}\left[T\right]$

$$\mathbb{E}\left[T\right] \leq \sum_{k\geq 0} N\mathbb{P}\left[T > kN\right] \leq N\sum_{k\geq 0}(1 - \varepsilon)^k \leq N\sum e^{-k\varepsilon} < \infty$$

Wykład 06.11.23 : Twierdzenie Dooba o zatrzymaniu, czyli jak uprawiać hazard

Dla $T : \Omega \to \mathbb{N}$ i procesu $\{X_n\}$ definiujemy zmienną X_T wzorem

$$X_T(\omega) = X_{T(\omega)}(\omega)$$

Dla martyngału $\{X_n\}_{n\in\mathbb{N}}$ i czasu zatrzymania T rozważamy ciąg zmiennych $\{X_{n\wedge T}\}_{n\in\mathbb{N}}$

$$X_{n\wedge T}(\omega) = \begin{cases} X_n & n \leq T(\omega) \\ X_{T(\omega)}(\omega) & n \geq T(\omega) \end{cases}$$

Tutaj dla $X, y \in \mathbb{R}$ piszemy $x \wedge y$ aby przekazać, że interesuje nas $\min\{x, y\}$. To znaczy $x \wedge y = \min\{x, y\}$.

Czyli gramy w pewna uczciwą grę i mamy strategię wyjścia T, ale musimy np. zdążyć na obiad, więc chcemy wyjść po co najwyżej n rundach.

Twierdzenie 5.1 : Dooba o zatrzymaniu (uproszczone).
Niech $\{X_n\}$ będą odpowiednio martyngałem i czasu zatrzymania względem tej samej filtracji $\mathbb{F} = \{\mathcal{F}_n\}$. Wówczas proces (ciąg) $\{X_{n\wedge T}\}$ zdefiniowany wyżej jest martyngałem. W szczególności

$$\mathbb{E}\left[X_{n\wedge T}\right] = \mathbb{E}\left[X_0\right]$$

dla każdego n (średnia jest stała w czasie).

Dowód
Mamy

$$X_{n\wedge T} = \sum_{k=1}^{n\wedge T}(X_k - X_{k-1}) + X_0 = \sum_{k=1}^{n} \mathbb{1}_{T\geq k}(X_k - X_{k-1}) + X_9$$

gdzie

$$\mathbb{1}_{T\geq k} = 1 - \mathbb{1}_{T<k} = 1 - 1_{T\leq k-1} \in \mathcal{F}_{k-1}$$

i teza wynika z przykładu o transformacie martyngałowej.

Przykład(y) 5.1

1. Gracz rozpoczyna grę z kapitałem j\$. W każdym rozdaniu może z prawdopodobieństwem $\frac{1}{2}$ zyskać jednego dolara lub go stracić. Celem gracza jest wzbogacenie się o k dolarów. Jakie jest prawdopodobieństwo sukcesu $p_{k,j}$?

   Zaczynamy w punkcie j i chcemy dojść do punktu j + k, a boimy się punktu 0

Niech $\{\xi_k\}$ będą niezależne o tym samym rozkładzie $\mathbb{P}\left[\xi_k = \pm 1\right] = \frac{1}{2}$. Rozważmy

$$X_n = \sum_{k=1}^{n} \xi_n.$$

Żeby rozwiązać to zadanie to chcemy rozważyć funkcję

$$T = \inf\{n \in \mathbb{N} \ : \ X_n = -j \text{ lub } X_n = k\}$$

Teraz szukane przez nas prawdopodobieństwo to

$$p_{k,j} = \mathbb{P}\left[X_T = k\right]$$

Rozważamy filtrację $\mathbb{F} = \{\mathcal{F}_n\}$, $\mathcal{F}_n = \sigma(\xi_1, ..., \xi_n)$. Ciąg $\{X_n\}$ jest $\mathbb{F}$-adaptowalny, więc T jest $\mathbb{F}$-czasem zatrzymania.

Ciąg $\{X_n\}$ jest $\mathbb{F}$-martyngałem, co wynika z faktu, że ξ_{n+1} są niezależne od $\mathcal{F}_n$ i mają $\mathbb{E}$ równą 0:

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[\xi_{n+1} + X_n \mid \mathcal{F}_n\right] = \mathbb{E}\left[\xi_{n+1}\right] + \mathbb{E}\left[X_n \mid \mathcal{F}_n\right] = 0 + X_n$$

Z twierdzenia o zatrzymaniu wiemy więc, że

$$\mathbb{E}\left[X_{n \wedge T}\right] = \mathbb{E}\left[X_0\right] = 0$$

i tutaj szkopuł jest taki, że nas interesuje X_T a nie $X_{n\wedge T}$. Musimy więc przejść z n do nieskończoności.

W pierwszej kolejności chcemy się upewnić, że $\mathbb{P}\left[T < \infty\right] = 1$, bo

$$\mathbb{P}\left[T \geq n\right] \leq \mathbb{P}\left[|X_n| \leq k + j\right] = \mathbb{P}\left[\frac{|X_n|}{\sqrt{n}} \leq \frac{j+k}{\sqrt{n}}\right] \xrightarrow{CTG} 0$$

a ponieważ

$$\mathbb{P}\left[T = \infty\right] = \lim_{n\to\infty} \mathbb{P}\left[T \geq n\right] = 0.$$

W takim razie ciąg $X_{n\wedge T}$ zbiega prawie wszędzie do ciągu X_T. Mało tego, dla pewnego n się zacznie stabilizować. Pozostaje uzasadnić, że możemy wejść z granicą pod całkę, ale to wynika z faktu, że

$$|X_{n\wedge T}| \leq j + k,$$

więc mamy

$$0 = \mathbb{E}\left[X_0\right] = \lim \mathbb{E}\left[X_{n\wedge T}\right] = \mathbb{E}\left[\lim X_{n\wedge T}\right] = \mathbb{E}\left[X_T\right].$$

Rozpisując już na końcu

$$0 = \mathbb{E}\left[X_T\right] = k\mathbb{P}\left[X_T = k\right] - j\mathbb{P}\left[X_T = -j\right] = k \cdot p_{k,j} - j(1 - p_{k,j})$$

co pozwala nam wyliczyć

$$p_{k,j} = \frac{j}{k+j}.$$

W szczególności mamy

$$\mathbb{P}\left[\{X_n\} \text{ osiągnie } k\right] = \lim_{j\to\infty} \mathbb{P}\left[\{X_k\} \text{ osiągnie } k \text{ przed osiągnięciem -}j\right] =$$

$$= \lim_{j\to\infty} p_{k,j} = \lim_{j\to\infty} \frac{j}{j+k} = 1$$

2. Gracz rozpoczyna grę z kapitałem j\$. W każdym rozdaniu może z prawdopodobieństwem p zyskać jednego dolara lub stracić go z prawdopodobieństwem $(1-p)$. Celem gracza jest wzbogacenie się o k dolarów. Jakie jest prawdopodobieństwo sukcesu $p_{k,j}$ gdy $p > \frac{1}{2}$?

   Jest to niemalże takie samo zadanie jak wcześniej, z tym że tym razem nie mamy martyngału. Niemniej jednak modelować będziemy to w niemalże identyczny sposób.

   Niech $\{\eta_k\}$ będą iid takie, że $\mathbb{P}\left[\eta_k = 1\right] = p$ oraz $\mathbb{P}\left[\eta_k = -1\right] = 1-p$. Określamy $X_n = \sum_{k=1}^{n} \eta_k$. Mamy wówczas

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[\eta_{n+1}\right] + X_n = X_n + (2p-1) > X_n$$

   czy $\{X_n\}$ jest podmartyngałem. Określmy czas zatrzymania

$$T = \inf\{n \ : \ X_n = -j \text{ lub } X_n = k\}.$$

   Chcemy sobie zorganizować nowy martyngał postaci

$$M_n = f(X_n)$$

   dla pewnej funkcji $f : \mathbb{Z} \to \mathbb{R}$ takiej, że

$$\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] = M_n.$$

   Mamy

$$\begin{aligned}\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[f(X_n + \eta_{n+1} \mid \mathcal{F}_n\right] = \\ &= F(X_n),\end{aligned}$$

   gdzie $F(x) = \mathbb{E}\left[f(x + \eta_{n+1})\right] = pf(x+1) + (1-p)f(x-1)$ jest oznaczeniem pomocniczym przy "odcałkowaniu niezależnej η_{n+1}".

   Aby $\{M_n\}$ był martyngałem musi zachodzić

$$M_n = f(X_n) = pf(X_n+1) + (1-p)f(X_n-1) = F(X_n) = \mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right]$$

   f musi zatem spełniać rekurencję

$$f(x) = pf(x+1) + (1-p)f(x-1)$$

   Szukamy rozwiązania postaci $f(x) = \gamma^x$. Mamy więc

$$\gamma^x = p\gamma^{x+1} + (1-p)\gamma^{x-1}$$

$$\gamma = p\gamma^2 + (1-p)$$

   i istnieją dwa rozwiązania: $\gamma = 1$ oraz $\gamma = \frac{1-p}{p}$. Wówczas

$$M_n = f(X_n) = \left(\frac{1-p}{p}\right)^{X_n}$$

   jest martyngalem. Znowu $M_{n\wedge T} \leq \left(\frac{p}{1-p}\right)^{k+j}$, a z twierdzenia Dooba

$$\mathbb{E}\left[M_{n\wedge T}\right] = \mathbb{E}\left[M_0\right] = 1$$

i poprzez przejście graniczne

$$\mathbb{E}[M_T] = 1$$

Oznaczamy $\mathbb{P}[X_T = k] = r_{j,j}$ i mamy

$$1 = \mathbb{E}[M_T] = \gamma^{-j}(1 - r_{k,j}) + \gamma^k r_{k,j}$$

gdzie

$$r_{k,j} = \frac{1 - \gamma^{-j}}{\gamma^k - \gamma^{j-1}}$$

w szczególności

$$\mathbb{P}\left[\{X_n\} \text{ osiągnie } k\right] = \lim_{j\to\infty} r_{j,k} = 1$$

$$\mathbb{P}\left[\{X_n\} \text{ osiągnie } j\right] = \lim_{k\to\infty} (1 - r_{k,j}) = \lim_{k\to\infty} \ldots = \gamma^j$$

3. Rozważmy $X_n = \sum Y_k$, gdzie Y_k są iid takie, że $\mathbb{P}[Y_1 = \pm 1] = \frac{1}{2}$. To znaczy, że jeden gracz bierze udział w uczciwej grze i obstawiamy.

5.1 Zadania

Zadanie 1.
Uzasadnij, że jeżeli $\{X_n\}$ są niezależnymi całkowalnymi zmiennymi losowymi o tym samym rozkładzie, a T jest czasem zatrzymania względem filtracji $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, takim że $\mathbb{E}[T] < \infty$, to

$$\mathbb{E}[S_T] = \mathbb{E}[T] \cdot \mathbb{E}[X_1]$$

gdzie $S_n = X_1 + X_2 + \ldots + X_n$.

Rozwiązanie.
PLAN:

$$\text{martyngał } Y_n = S_n - n\mathbb{E}[X_1]$$
$$\downarrow$$
$$T \text{ jest prawie wszędzie skończony}$$
$$\downarrow$$
$$Y_{n\wedge T} \text{ od pewnego momentu jest stale równy } Y_T$$
$$\downarrow$$
$$\text{granica } \mathbb{E}[\lim Y_{n\wedge T}] = \lim \mathbb{E}[Y_{n\wedge T}] \text{ ma sens}$$
$$\downarrow$$
$$\mathbb{E}[Y_T] = \mathbb{E}[S_T - T\mathbb{E}[X_1]] = 0$$

Jesteśmy w temacie martyngałów, więc możemy chcemy tego użyć.

Niech $m = \mathbb{E}[X_1] = \mathbb{E}[X_n]$ dla każdego n. Dobrym początkiem będzie pokazanie, że ciąg $\{S_n - nm\}$ jest martyngałem. W tym celu potrzebujemy całkowalności $[S_n - nm]$, $\mathcal{F}_n$-mierzalności i równości wwo.

1. $[S_n - nm]$ jest całkowalne

$$\mathbb{E}\left[|S_n - nm|\right] = \mathbb{E}\left[\left|\sum X_k - m\right|\right] \leq \mathbb{E}\left[\sum |X_k - m|\right] = \sum \mathbb{E}\left[|X_k - m|\right] < \infty$$

2. $[S_n - nm]$ jest $\mathcal{F}_n$-mierzalne, bo jest skończoną sumą $\mathcal{F}_n$-mierzalnych funkcji (wraz z funkcją stałą).

3. $\mathbb{E}\left[S_{n+1} - (n+1)m \mid \mathcal{F}_n\right] = S_n - nm$

$$\begin{aligned}\mathbb{E}\left[S_{n+1} - (n+1)m \mid \mathcal{F}_n\right] &= \mathbb{E}\left[\sum^{n+1}(X_i - m) \mid \mathcal{F}_n\right] = \\ &= \mathbb{E}\left[X_{n+1} - m \mid \mathcal{F}_n\right] + \sum(X_i - m) = \\ &= \mathbb{E}\left[X_{n+1}\right] - m + S_n - nm = m - m + S_n - nm = S_n - nm\end{aligned}$$

Dla uładnienia zapisu niech $Y_k = n = S_n - n \cdot m$, wtedy $\{Y_n\}$ jest martyngałem względem filtracji jak w zadaniu. Z twierdzenia Dooba o zatrzymaniu wiemy, że

$$\mathbb{E}\left[Y_{n \wedge T}\right] = \mathbb{E}\left[Y_1\right] = S_1 - m = 0$$

Będziemy chcieli przejść z n do granicy, do czego potrzebujemy aby $\mathbb{P}\left[T \geq n\right] \to 0$, bo wówczas ciąg $X_{n \wedge T}$ zbiega prawie wszędzie do X_T. Wystarczy przypomnieć sobie ostatnie zadanie z poprzedniej listy, aby dostać

$$\mathbb{E}\left[T\right] = \sum_{k \geq 0} \mathbb{P}\left[T > k\right] < \infty$$

czyli w pewnym momencie wyrazu muszą być dowolnie blisko 0, czyli faktycznie $\mathbb{P}\left[T \geq n\right] \to 0$.

Przechodząc z n do granicy dostajemy

$$\mathbb{E}\left[Y_T\right] = \mathbb{E}\left[\lim Y_{n \wedge T}\right] = \lim \mathbb{E}\left[Y_T\right] = 0$$

ponieważ $Y_{n \wedge T}$ zbiega do Y_T, więc od pewnego momentu jest stały i granica ma sens.

Rozbijając więc Y_T na wzór podany wyżej, dostajemy

$$0 = \mathbb{E}\left[Y_T\right] = \mathbb{E}\left[S_T - Tm\right] = \mathbb{E}\left[S_T\right] - \mathbb{E}\left[T\mathbb{E}\left[X_1\right]\right]$$

czyli

$$\mathbb{E}\left[S_T\right] = \mathbb{E}\left[T\mathbb{E}\left[X_1\right]\right] = \mathbb{E}\left[X_1\right]\mathbb{E}\left[T\right]$$

tak jak chcieliśmy.

Można też zapisać najpierw $\mathbb{E}\left[|S_T|\right] = \sum_{t=1}^{\infty} \mathbb{E}\left[|S_T| \mathbb{1}_{\{T=t\}}\right]$. JEDNAK BŁĄD Z $\mathbb{P}\left[T \geq k\right] \to 0$ bo to niekoniecznie musi być prawdą - trzeba by używać $\mathbb{E}\left[S_{n \wedge T}\right]$

Zadanie 2.
Rzucamy kostką tak długo, aż pięciokrotnie wyrzucimy szóstkę. Znajdź średnią wartość sumy wyrzuconych oczek.

Rozwiązanie.
Zadania wygląda bardzo podobnie jak równość udowadniana wyżej. Chcemy tylko znaleźć martyngał i czas zatrzymania.

Niech X_i będzie liczbą oczek wyrzuconych w i-tym rzucie, a $S_n = \sum_{k=1}^{n} X_i$ będzie sumą oczek wyrzuconych w pierwszych n rzutach. Oczywiście, X_i mają ten sam rozkład jednostajny na zbiorze $\{1, ..., 6\}$ i są od siebie niezależne. Zdefiniujmy teraz funkcję

$$T = \inf\{n \ : \ (X_1, ..., X_n) \text{ posiada 5 szóstek}\}$$

która jest czasem zatrzymania względem filtracji $\mathbb{F} = \{\mathcal{F}_n\}$, $\mathcal{F}_n = \sigma(X_1, ..., X_n)$, bo jej definicja opiera się wyłącznie na informacjach o $X_1, ..., X_n$.

Korzystając więc z poprzedniego zadania, dostajemy

$$\mathbb{E}[S_T] = \mathbb{E}[T]\,\mathbb{E}[X_1] = \mathbb{E}[T] \cdot \frac{7}{2}$$

i jedynym problemem jest obliczenie

$$\mathbb{E}[T] = \sum_{n \geq 0} n\mathbb{P}[T = n].$$

Oczywiście, dla $\mathbb{P}[T = 1] = \mathbb{P}[T = 4] = 0$, a w pozostałych przypadkach jest to stosunek wszystkich ciągów długości $n - 1$ które posiadają dokładnie 4 szóstki do ilości wszystkich ciągów posiadających co najwyżej 4 szóstki.

Zadanie 3.
Niech $\{X_n\}$ będzie niesymetrycznym spacerem losowym na $\mathbb{Z}$ (tzn. $X_n = \sum_{k=1}^{n} \xi_k$, gdzie ξ_k są iid takie, że $\mathbb{P}\left[\xi_k = 1\right] = 1 - \mathbb{P}\left[\xi_k = -1\right] = p \neq \frac{1}{2}$) i niech $T = \min\{n \ : \ X_n = -j \text{ lub } X_n = k\}$ dla ustalonych $k, j > 0$.

(a) Pokaż, że $M_n = X_n + n(1 - 2p)$ jest martyngałem.

(b) Wykorzystując twierdzenie Dooba oblicz $\mathbb{E}[T]$.

Rozwiązanie.
Filtrem u mnie będzie $\mathcal{F}_n = \sigma(\xi_1, ...\xi_n)$.

(a) Wypadałoby pokazać, że M_n jest całkowalne, czyli

$$\begin{aligned}\mathbb{E}\left[|M_n|\right] &= \mathbb{E}\left[|X_n + n(1 - 2p)|\right] \leq \mathbb{E}\left[|X_n|\right] + n(1 - 2p) = \\ &= \mathbb{E}\left[\left|\sum \xi_k\right|\right] + n(1 - 2p) \leq \sum \mathbb{E}\left[|\xi_k|\right] + n(1 - 2p) < \infty\end{aligned}$$

Jest $\mathcal{F}_n$-mierzalne bo jest kombinacją funkcji $\mathcal{F}_n$-mierzalnej z funkcją stałą. Pozostaje warunek z wwo:

$$\begin{aligned}
\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right] &= \mathbb{E}\left[X_{n+1} + (n+1)(1-2p) \mid \mathcal{F}_n\right] = \\
&= \mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] + (n+1)(1-2p) = \\
&= \mathbb{E}\left[\sum^{n+1} \xi_k \mid \mathcal{F}_n\right] + (n+1)(1-2p) = \\
&= \sum^{n+1} \mathbb{E}\left[\xi_k \mid \mathcal{F}_n\right] + (n+1)(1-2p) = \\
&= X_n + n(1-2p) + \mathbb{E}\left[\xi_{n+1} \mid \mathcal{F}_n\right] + (1-2p) = \\
&= X_n + n(1-2p) = M_n
\end{aligned}$$

(b) To jest tak samo jak w zadaniu 1, tylko trzeba pokazać, że z prawdopodobieństwem 1 w kończenie wielu krokach dojdziemy do –j lub k. A tutaj można użyć Borel-Cantalliego :3

Zadanie 4.
Niech $\{M_n\}$ będzie nieujemnym martyngałem. Pokaż, że dla $m > n$, $\{M_n = 0\} \subseteq \{M_m = 0\}$ prawie wszędzie.

Rozwiązanie.
Rozważmy zbiór $A = \{M_n = 0\} \in \mathcal{F}_n$. Ponieważ $\{M_n\}$ jest martyngałem, to na poprzedniej liście pokazywaliśmy, że

$$\mathbb{E}\left[M_m \mid \mathcal{F}_n\right] = M_n.$$

W takim razie mamy

$$0 = \mathbb{E}\left[M_n \mathbb{1}_A\right] = \mathbb{E}\left[\mathbb{E}\left[M_m \mid \mathcal{F}_n\right] \mathbb{1}_A\right] = \mathbb{E}\left[M_m \mathbb{1}_A\right]$$

Ponieważ $M_m \geq 0$ oraz $\int_A M_m d\mathbb{P} = 0$, to $M_m = 0$ prawie wszędzie na zbiorze A. Czyli

$$A \subseteq \{M_m = 0\}$$

prawie wszędzie.

Zadanie 5.
Niech $\mathbb{F} = \{\mathcal{F}_n\}$ będzie filtracją.

(a) Pokaż, że dla każdych $m, n \in \mathbb{N}$, $m < n$ i zdarzenia $A \in \mathcal{F}_m$ zmienna losowa

$$\tau = m + (n-m)\mathbb{1}_A$$

jest $\mathbb{F}$-czasem zatrzymania.

(b) Niech $\{X_n\}$ będzie $\mathbb{F}$-adaptowalnym ciągiem całkowalnych zmiennych losowych takim, że

$$\mathbb{E}\left[X_\tau\right] = \mathbb{E}\left[X_0\right]$$

dla każdego skończonego czasu zatrzymania τ. Pokaż, że $\{X_n\}$ jest $\mathbb{F}$-martyngałem.

Rozwiązanie.

(a) Musimy sprawdzić, że zdarzenie $\{\tau = k\} \in \mathcal{F}_k$. Od razu widzimy, że τ ma jedynie dwie możliwe wartości:

$$\tau(\omega) = \begin{cases} m & \omega \notin A \\ n & \omega \in A \end{cases}$$

Zacznijmy od $\{\tau = m\}$. Ale tak jak wyżej napisaliśmy, to zachodzi tylko dla $\omega \notin A$, czyli $\{\tau = m\} = A^c \in \mathcal{F}_m$, bo $A \in \mathcal{F}_m$.

Z kolei $\{\tau = n\} = A$, a wiemy, że $A \in \mathcal{F}_m \subseteq \mathcal{F}_n$ bo $m < n$.

W takim razie τ zachowuje się jak bardzo dziwny czas zatrzymania.

(b) PLAN na ten podpunkt jest taki (pokazujemy $\mathbb{E}\left[X_{m+1} \mid \mathcal{F}_m\right] = X_m$):

$\tau' = m + (n - m)\mathbb{1}_{\emptyset}$ z założenia wynika, że $\mathbb{E}[X_0] = \mathbb{E}[X_m]$

$\downarrow$

nowy czas zatrzymania $\tau = m + (m + 1 - m)\mathbb{1}_A$ dla $A \in \mathbb{1}_m$

$\downarrow$

z wwo wyprowadzamy, że $\mathbb{E}[X_{m+1}\mathbb{1}_A] = \mathbb{E}[X_\tau \mathbb{1}_A]$

$\downarrow$

$\mathbb{E}[X_\tau] = \mathbb{E}[X_\tau \mathbb{1}_A] + \mathbb{E}\left[X_\tau \mathbb{1}_{A^c}\right]$

$\downarrow$

korzystamy z pierwszego kroku

Zaczniemy od wprowadzenia czasu zatrzymania

$$\tau' = m + (n - m)\mathbb{1}_{\emptyset}$$

dla dowolnego $n > m$ i użycia założenia z tego punktu, by otrzymać

$$\mathbb{E}[X_0] = \mathbb{E}\left[X_{\tau'}\right] = \mathbb{E}\left[X_{\tau'}\mathbb{1}_{\emptyset}\right] + \mathbb{E}\left[X_{\tau'}\mathbb{1}_{\emptyset^c}\right] = \mathbb{E}\left[X_m \mathbb{1}_{\Omega}\right] = \mathbb{E}[X_m]$$

gdyż $\mathbb{E}\left[X_{\tau'}\mathbb{1}_{\Omega}\right] = \int_{\Omega} X_{\tau'(\omega)}(\omega) = \int X_m$ - całkujemy tę część tylko po zbiorze gdzie $\tau' = m$.

Rozważamy teraz czas zatrzymania

$$\tau = m + (m + 1 - m)\mathbb{1}_A = m + (m + 1)\mathbb{1}_A$$

dla dowolnego $A \in \mathcal{F}_m$.

Z jednej strony wiemy, że

$$\mathbb{E}\left[\mathbb{E}\left[X_{m+1} \mid \mathcal{F}_m\right]\mathbb{1}_A\right] = \mathbb{E}[X_{m+1}\mathbb{1}_A] = \mathbb{E}[X_\tau \mathbb{1}_A]$$

A z drugiej

$$\mathbb{E}[X_0] = \mathbb{E}[X_\tau] = \mathbb{E}[X_\tau \mathbb{1}_A] + \mathbb{E}\left[X_\tau \mathbb{1}_{A^c}\right] = \mathbb{E}[X_{m+1}\mathbb{1}_A] + \mathbb{E}\left[X_m \mathbb{1}_{A^c}\right]$$

czyli korzystając z $\mathbb{E}[X_0] = \mathbb{E}[X_m]$ mamy

$$\mathbb{E}[X_{m+1}\mathbb{1}_A] = \mathbb{E}[X_0] - \mathbb{E}\left[X_m\mathbb{1}_{A^c}\right] = \mathbb{E}[X_m] - \mathbb{E}\left[X_m\mathbb{1}_{A^c}\right] = \mathbb{E}\left[X_m(1 - \mathbb{1}_{A^c}\right] = \mathbb{E}[X_m\mathbb{1}_A]$$

i to już wystarczy, bo X_m jest mierzalne względem $\mathcal{F}_m$ ponieważ jest $\mathbb{F}$-adaptowalne.

Zadanie 6.
Niech $\{M_n\}$ będzie nieujemnym martyngałem o wartościach całkowitym takim, że $M_0 = m \geq 1$, $M_n - M_{n-1} \leq 1$ oraz $M_n \to 0$ p.w.. Pokaż, że dla $k \geq m$,

$$\mathbb{P}\left[\sup_{n\in\mathbb{N}} M_n \geq k\right] = \frac{m}{k}$$

Rozwiązanie.
Zacznijmy od rozważenia co znaczy, że $\sup_{n\in\mathbb{N}} M_n \geq k$. Moim skromnym zdaniem jest to powiedzenie, że M_n dojdzie do k. Wiemy, że w każdym przejściu z M_n do M_{n+1} możemy skoczyć w górę o nie więcej niż 1, a w dół możemy skakać aż do 0.

PLAN na to zadanie jest taki:

czas zatrzymania $T = \inf\{n \ : \ M_n \geq k\}$

$\downarrow$

$\sup M_n \geq k \iff T < \infty$, więc $\mathbb{P}[\sup M_n \geq k] = \mathbb{P}[T < \infty]$

$\downarrow$

twierdzenie Dooba daje $m = \mathbb{E}[M_0] = \mathbb{E}[M_{n\wedge T}]$

$\downarrow$

$\lim_n M_{n\wedge T}$ będzie $M_T = k$ jeśli $T < \infty$ i M_n wpp.

$\downarrow$

$M_{n\wedge T} \in \{0, ..., k\}$

$\downarrow$

$\mathbb{P}[M_{n\wedge T} < k] \to 0$, bo jeśli $T = \infty$ to $M_n \to 0$, a jeśli $T < \infty$, to $M_{n\wedge T} \to M_T = k$ jest stałe od pewnego momentu

$\downarrow$

$\lim \mathbb{P}[M_{n\wedge T} = k] = \mathbb{P}[T < \infty]$

$\downarrow$

$m = \sum_{0\leq i\leq k} i\mathbb{P}[M_{n\wedge T}] \to k\mathbb{P}[T < \infty]$

Spróbujmy zobaczyć co się stanie, jeśli wplączemy w to zadanie czas zatrzymania

$$T = \inf\{n \ : \ M_n \geq k\}$$

to możemy zauważyć, że

$$\mathbb{P}\left[\sup_{n\in\mathbb{N}} M_n \geq k\right] = \mathbb{P}[T < \infty] = \mathbb{P}[M_{n\wedge T} = k]$$

ponieważ $T < \infty$ oznacza, że zbiór $\{n \ : \ M_n \geq k\}$ jest niepusty.

Z twierdzenia Dooba wiemy, że

$$m = \mathbb{E}[M_0] = \mathbb{E}[M_{n\wedge T}]$$

zauważmy, że jeżeli $T(\omega) < \infty$, to $M_{n\wedge T}(\omega) = M_{T(\omega)}(\omega) = k$, a jeśli $T(\omega) = \infty$, to $M_{n\wedge T}(\omega) = M_n(\omega)$.

Oznacza to, że jeśli $M_{n\wedge T} \geq k$, to $M_{n\wedge T} = k$ i jest od tego momentu funkcją stałą. W takim razie

$$m = \mathbb{E}\left[M_{n\wedge T}\right] = \sum_{j=0}^{k} j\mathbb{P}\left[M_{n\wedge T} = j\right] = k \cdot \mathbb{P}\left[M_{n\wedge T} = k\right] + \sum_{j=0}^{k-1} j\mathbb{P}\left[M_{n\wedge T} = j\right]$$

przechodząc teraz z n do granicy, albo $M_{n\wedge T} = k$ albo $M_{n\wedge T} = M_n$ i w obu przypadkach $\mathbb{P}\left[M_{n\wedge T} = j\right] = 0$ dla $j < k$ (bo $M_n \to 0$). W takim razie, po przejściu do granicy z n zostaje nam

$$m = k \cdot \mathbb{P}\left[\sup_{n\in\mathbb{N}} M_n \geq k\right]$$

i to jest to co chcieliśmy.

Zadanie 7.
Niech Y_k będą iid takie, że $\mathbb{P}\left[Y_k \in \{-1, 0, 1\}\right] = 1$ oraz $\mathbb{E}\left[Y_k\right] = 0$. Niech $S_0 = 0$, $S_n = Y_1 + \ldots + Y_n$. Dla $k \in \mathbb{N}$ rozważmy moment zatrzymania

$$T_{-k} = \inf\{n \ : \ S_n = -k\}.$$

Znajdź rozkład zmiennej losowej

$$\sup_{n\leq T_{-k}} S_n$$

Rozwiązanie.
Zmienna losowa

$$\sup_{n\leq T_{-k}} S_n$$

pyta, jak wysoko możemy dojść, jeśli zatrzymamy się przy pierwszym dojściu do $-k$.

Popatrzmy teraz na $\mathbb{P}\left[Y_k = 1\right]$. Oznaczając $\mathbb{P}\left[Y_k = 0\right] = p$ mamy

$$0 = \mathbb{E}\left[Y_k\right] = \mathbb{P}\left[Y_k = 1\right] - \mathbb{P}\left[Y_k = -1\right]$$

$$\mathbb{P}\left[Y_k = 1\right] = \mathbb{P}\left[Y_k = -1\right] = \frac{1-p}{2}$$

Wykład 13.11.23 : Czyli odrabiam wykład z notatek

Wykład 18.11.23 : Sobotnia mordęga

7.1 Nierówności maksymalne

Niech $X = \{X_n\}$ będzie $\mathbb{F} = \{\mathcal{F}_n\}$ podmartyngałem. Wówczas

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] \geq X_n$$

dla każdego $n \in \mathbb{N}$.

Twierdzenie 7.1 : o zatrzymaniu dla podmartyngałów.
Jeżeli X jest $\mathbb{F}$-podmartyngałem, a T $\mathbb{F}$-czasem zatrzymania, to dla każdego $n \in \mathbb{N}$

$$\mathbb{E}\left[X_n\right] \geq \mathbb{E}\left[X_{n\wedge T}\right]$$

Dowód
Mamy

$$\begin{aligned} X_n - X_{n\wedge T} &= \sum_{k=n\wedge T+1}^{n} (X_k - X_{k-1}) = \sum_{k=1}^{n} \mathbb{1}_{\{k\geq n\wedge T+1\}}(X_k - X_{k-1}) = \\ &= \sum_{k=1}^{n} \underbrace{\mathbb{1}_{\{k\geq T+1\}}}_{\in\mathcal{F}_{k-1}}(X_k - X_{k-1}) \end{aligned}$$

Co z własności martyngału da nam

$$\mathbb{E}\left[\mathbb{1}_{\{k\geq T+1\}}(X_k - X_{k-1}) \mid \mathcal{F}_{k-1}\right] = \mathbb{1}_{\{k\geq T+1\}}\mathbb{E}\left[X_k - X_{k-1} \mid \mathcal{F}_{k-1}\right] \geq 0$$

Czyli jak zsumujemy rzeczy ≥ 0, to również dostaniemy wartość ≥ 0, tzn.

$$\mathbb{E}\left[X_n - X_{n\wedge T}\right] \geq 0$$

i to już kończy dowód.

Twierdzenie 7.2 : nierówność maksymalna słabego typu.
Jeżeli X jest podmartyngałem, to dla każdego $n \in \mathbb{N}$ oraz $\lambda > 0$ mamy

$$\mathbb{P}\left[\max_{k=0,\dots,n} X_k \geq \lambda\right] \leq \mathbb{E}\left[X_n \mathbb{1}_{\{\max_{k=0,\dots,n} X_k \geq \lambda\}}\right] \leq \mathbb{E}\left[|X_n|\right]$$

(środkowa nierówność nie jest pamiętana przez zdrowych psychicznie ludzi, ale będzie potrzebna do dowodu następnego dowodu)

Dowód
Wybierzmy czas zatrzymania

$$\tau = \inf\{k \ : \ X_k \geq \lambda\}.$$

Z twierdzenia o zatrzymaniu 7.1 wiemy, że

$$\mathbb{E}[X_n] \geq \mathbb{E}[X_{n\wedge\tau}] = \mathbb{E}\left[\underbrace{X_{n\wedge\tau}}_{=X_\tau}\mathbb{1}_{\{\max X_k \geq \lambda\}}\right] + \mathbb{E}\left[\underbrace{X_{n\wedge\tau}}_{=X_n}\mathbb{1}_{\{\max X_k < \lambda\}}\right] = (\star)$$

nawiasy na dole są bo pierwsza całka jest po zbiorze gdzie już mamy max $\geq \lambda$, a druga jest po zbiorze gdzie to nigdy nie dojdzie do λ. Idąc dalej mamy

$$\mathbb{E}[X_n] \geq (\star) \geq \lambda\mathbb{P}\left[\max X_k \geq \lambda\right] + \mathbb{E}\left[X_n \mathbb{1}_{\max X_k < \lambda}\right]$$

Stad, jeśli przerzucimy $\mathbb{E}$ na jedną stronę, to dostajemy

$$\mathbb{E}\left[X_n \mathbb{1}_{\max X_k \geq \lambda}\right] \geq \lambda\mathbb{P}\left[\max X_k \geq \lambda\right]$$

🍻

Wniosek 7.3.
Jeżeli $M = \{M_n\}$ jest martyngałem takim, że dla pewnego $p \geq 1$

$$\mathbb{E}\left[|M_n|^p\right] < \infty$$

dla każdego $n \in \mathbb{N}$, to jeśli zastosuje się nierówność Jensena,

$$X_n = |M_n|^p$$

jest podmartyngałem.

Dowód

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = \mathbb{E}\left[|M_{n+1}|^p \mid \mathcal{F}_n\right] \geq \left|\mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n\right]\right|^p = |M_n|^p = X_n$$

🍻

Wówczas mamy

$$\lambda\mathbb{P}\left[\max |M_n|^p \geq \lambda\right] \leq \mathbb{E}\left[|M_n|^p \mathbb{1}_{\{\max |M_k|^p \geq \lambda\}}\right]$$

Czyli dla $s = \lambda^{1/p}$ dostajemy

$$\mathbb{P}\left[\max |M_k| \geq s\right] \leq \frac{1}{s^p}\mathbb{E}\left[|M_n|^p \mathbb{1}_{\{\max |M_k| \geq s\}}\right]$$

Przykład(y) 7.1

1. Nierówność maksymalna Kołmogorowa.

   Niech ξ_k będzie ciągiem niezależnych zmiennych z $\mathbb{E}\left[\xi_k\right] = 0$ i $\mathbb{E}\left[\xi_k^2\right] < \infty$, to oznaczając $S_n = \sum_{k \le n} \xi_k$ wiemy już, że S_n jest martyngałem. Wówczas

$$\mathbb{P}\left[\max|S_k| \ge \epsilon\right] \le \frac{1}{\epsilon^2}\mathbb{E}\left[S_n^2\right] = \frac{1}{\varepsilon^2}\sum_{k \le n}\mathbb{E}\left[\xi_j^2\right]$$

Twierdzenie 7.4 : nierówność maksymalna mocnego typu.

Niech $M = \{M_n\}$ będzie martyngałem takim, że $\mathbb{E}\left[|M_n|^p\right] < \infty$ dla pewnego $p > 1$ i wszystkich $n \in \mathbb{N}$, to wtedy

$$\mathbb{E}\left[\max|M_k|^p\right] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|M_n|^p\right]$$

Dowód

Oznaczmy

$$M_n^* = \max_{k \le n}|M_k|$$

Mamy dla $k > 0$

$$\mathbb{E}\left[(M_n^* \wedge k)^p\right] = \mathbb{E}\left[\int_0^{M_n^* \wedge k} ps^{p-1}ds\right] = \mathbb{E}\left[\int_0^k \mathbb{1}_{\{M_n^* \ge s\}} ps^{p-1}ds\right] =$$

$$= \int_0^k ps^{p-1}\mathbb{P}\left[M_n^* \ge s\right] ds \le \int_0^k ps^{p-1}\overbrace{\frac{1}{s}\mathbb{E}\left[|M_n|\mathbb{1}_{\{M_n^* \ge s\}}\right]}^{7.2} =$$

$$= \mathbb{E}\left[|M_n|\int_0^k ps^{p-2}\mathbb{1}_{\{M_n^* \ge s\}}ds\right] = \mathbb{E}\left[|M_n|\int_0^{k \wedge M_n^*} ps^{p-2}ds\right] =$$

$$= \frac{p}{p-1}\mathbb{E}\left[|M_n|(M_n^* \wedge k)^{p-1}\right] \le \frac{p}{p-1}\mathbb{E}\left[|M_n|^p\right]^{1/p}\mathbb{E}\left[|M_n^* \wedge k|^{p/(p-1)}\right]^{(p-1)/p}$$

Przerzucając potworka z $\frac{p}{p-1}$ na lewą stronę, dostajemy

$$\mathbb{E}\left[(M_n^* \wedge k)^p\right]^{1/p} \le \frac{p}{p-1}\mathbb{E}\left[|M_n|^p\right]^{1/p}$$

przechodząc z k do nieskończoności. Użycie k było potrzebne przy dzieleniu.

Twierdzenie 7.5.

Niech $p > 1$. Dla martyngału $\{M_n\}$ następujące warunki są równoważne:

1. $\sup \mathbb{E}\left[|M_n|^p\right] < \infty$
2. istnieje zmienna losowa M_∞ taka, że $\mathbb{E}\left[|M_\infty|^p\right] < \infty$ i $M_n \to M_\infty$ prawie wszędzie w L^p

3. Istnieje zmienna losowa M_∞ taka, że $\mathbb{E}\left[|M_\infty|^p\right] < \infty$ i $\mathbb{E}\left[M_\infty \mid \mathcal{F}_n\right] = M_n$.

Dygresja.
Zbieżność w L^p oznacza, że $\mathbb{E}\left[|M_n - M_\infty|^p\right] \to 0$ prawie wszędzie.

Dowód

$1 \Rightarrow 2$ Wiemy, że

$$\sup \mathbb{E}\left[M_n^*\right]^p \leq \sup \mathbb{E}\left[|M_n|^p\right] < \infty,$$

zatem istnieje M_∞ taka, że $M_n \to M_\infty$. Dla każdego $n \in \mathbb{N}$ mamy

$$\mathbb{E}\left[\max |M_k|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|M_n|^p\right] \leq \left(\frac{p}{p-1}\right)^p \sup \mathbb{E}\left[|M_k|^p\right]$$

przy $n \to \infty$ daje to

$$\mathbb{E}\left[\sup |M_k|^p\right] \leq \left(\frac{p}{p-1}\right)^p \sup \mathbb{E}\left[|M_k|^p\right]$$

Mamy

$$|M_n - M_\infty|^p \leq (|M_n| + |M_\infty|)^p \leq 2^p \sup |M_k|^p$$

z twierdzenia o zbieżności ograniczonej wiemy, że

$$\lim \mathbb{E}\left[|M_n - M_\infty|^p\right] = \mathbb{E}\left[\lim |M_n - M_\infty|^p\right] = 0$$

$2 \Rightarrow 3$ Chcemy pokazać, że

$$\mathbb{E}\left[M_\infty \mid \mathcal{F}_n\right] = M_n$$

czyli, że dla każdego $A \in \mathcal{F}_n$

$$\mathbb{E}\left[M_\infty \mathbb{1}_A\right] = \mathbb{E}\left[M_n \mathbb{1}_A\right]$$

Wiemy, że dla każdego $m \in \mathbb{N}$ zachodzi

$$\mathbb{E}\left[M_{n+m} \mid \mathcal{F}_n\right] = M_n$$

czyli

$$\mathbb{E}\left[M_{n+m} \mathbb{1}_A\right] = \mathbb{E}\left[M_n \mathbb{1}_A\right]$$

przy $m \to \infty$ dostajemy

$$\left[\mathbb{E}\left[(M_\infty - M_{n+m})\mathbb{1}_A\right]\right]^p \leq \mathbb{E}\left[|M_\infty - M_{n+m}|^p\right] to 0$$

Skoro więc

$$\mathbb{E}\left[M_{n+m} \mathbb{1}_A\right] \xrightarrow{m \to \infty} \mathbb{E}\left[M_\infty \mathbb{1}_A\right]$$

$3 \Rightarrow 1$ Z założenia i nierówności Jensena mamy, że

$$|M_n|^p = |\mathbb{E}\left[M_\infty \mid \mathcal{F}_n\right]|^p \leq \mathbb{E}\left[|M_\infty|^p \mid \mathcal{F}_n\right]$$

stąd

$$\mathbb{E}\left[|M_n|^p\right] \leq \mathbb{E}\left[|M_\infty|^p\right]$$

w szczególności $\sup \mathbb{E}\left[|M_n|^p\right] \leq \infty$.

Przykład(y) 7.2

1. Niech $f : [0,1] \to \mathbb{R}$ będzie funkcją w L^p, tzn.:

$$\int_0^1 |f(t)|^p dt < \infty$$

Dla $n \in \mathbb{N}$ niech $I_k^{(n)} = (k-1)2^{-n}, k2^{-n})$. Niech teraz $f_k^{(n)} = 2^n \int_{I_k^{(n)}} f(t)dt$ będzie średnią f na przedziale $I_k^{(n)}$.

Niech teraz $f^{(n)} : [0,1] \to \mathbb{R}$ będzie dana przez

$$f^{(n)}(t) = \sum_{k=1}^{2^n} f_k^{(n)} \mathbb{1}_{I_k^{(n)}}(t)$$

Wówczas funkcje $f^{(n)}$ przybliżają f w $L^p[0,1]$, tzn $f^{(n)} \to f$.

Działamy w przestrzeni z $\Omega = [0,1]$, $\mathbb{P} = \lambda \restriction [0,1]$. Rozważamy rodzinę $\mathcal{F}_n = \sigma(I_k^{(n)} \;:\; k = 1, ..., 2^n)$, w której mamy $I_{2k}^{(n+1)} \cup I_{2k-1}^{(n+1)} = I_k^{(n)}$, więc jest ona filtracją.

$$\mathbb{E}\left[f \mid \mathcal{F}_n\right] = \sum_{k=1}^{2^n} \mathbb{1}_{I_k^{(n)}} \mathbb{E}\left[f \mid I_k^{(n)}\right] = \sum_{k=1}^{2^n} \mathbb{1}_{I_k^{(n)}} f_k^{(n)}$$

Czyli jesteśmy w 3. punkcie równoważności z twierdzenia wyżej, czyli mamy od razu dane 2, gdzie $M_\infty = f$, $M_n = f^{(n)}$.

7.2 Zadania

Zadanie 1.
Niech $\{X_n\}$ będzie martyngałem. Uzasadnij wykładniczą wersję nierówności Dooba: dla każdego $h > 0$ i każdego $n \in \mathbb{N}$

$$\mathbb{P}\left[\max X_k > x\right] \leq e^{-hx} \mathbb{E}\left[e^{hX_n}\right]$$

Rozwiązanie.
Rozważmy czas zatrzymania

$$T = \inf\{k \;:\; X_k > x\}$$

dowód wygląda podobnie jak dowód twierdzenia 7.2:

$$\begin{aligned} \mathbb{E}\left[e^{hX_n}\right] &\geq \mathbb{E}\left[e^{hX_{n\wedge T}}\right] = \mathbb{E}\left[e^{hX_{n\wedge T}} \mathbb{1}_{\{\max X_k > x\}}\right] + \mathbb{E}\left[e^{hX_{n\wedge T}} \mathbb{1}_{\{\max X_k \leq x\}}\right] = \\ &= \mathbb{E}\left[e^{hX_T} \mathbb{1}_{\{\max X_k > x\}}\right] + \mathbb{E}\left[e^{hX_n} \mathbb{1}_{\{\max X_k \leq x\}}\right] \geq \mathbb{E}\left[e^{hx} \mathbb{1}_{\{\max X_k > x\}}\right] = \\ &= e^{hx} \mathbb{E}\left[\mathbb{1}_{\{\max X_k > x\}}\right] = e^{hx} \mathbb{P}\left[\max X_k > x\right] \end{aligned}$$

czyli po przerzuceniu e^{hx} na lewą stronę, dostajemy tezę.

Zadanie 2.
Niech $\{X_n\}$ będzie martyngałem całkowalnym z kwadratem ($\mathbb{E}\left[X_n^2\right] < \infty$ dla każdego $n \in \mathbb{N}$). Pokaż, że

$$\mathbb{E}\left[(X_n - X_m)^2 \mid \mathcal{F}_m\right] = \mathbb{E}\left[X_n^2 \mid \mathcal{F}_m\right] - X_m^2$$

Dodatkowe założenia: $n > m$.

Rozwiązanie.
Po pierwsze, wypadałoby pokazać, że $X_n X_m$ jest całkowalne

$$\mathbb{E}\left[|X_n X_m|\right] \leq \mathbb{E}\left[|X_n|^2\right]^{1/2} \mathbb{E}\left[|X_m|^2\right]^{1/2} < \infty$$

W takim razie, wiedząc, że X_m jest zawsze mierzalna względem $\mathcal{F}_m$ wiemy, że $\mathbb{E}\left[X_n X_m \mid \mathcal{F}_m\right] = X_m \mathbb{E}\left[X_n \mid \mathcal{F}_m\right]$ i teraz jeśli $n > m$ to i $\mathbb{E}\left[X_n \mid \mathcal{F}_m\right] = X_m$, co pokazaliśmy już dawno temu na liście. Przechodząc z tą wiedzą do pisania, mamy;

$$\begin{aligned}
\mathbb{E}\left[(X_n - X_m)^2 \mid \mathcal{F}_m\right] &= \mathbb{E}\left[X_n^2 - 2X_n X_m + X_m^2 \mid \mathcal{F}_m\right] = \\
&= \mathbb{E}\left[X_n^2 \mid \mathcal{F}_m\right] + \mathbb{E}\left[X_m^2 \mid \mathcal{F}_m\right] - 2X_m \mathbb{E}\left[X_n \mid \mathcal{F}_m\right] = \\
&= \mathbb{E}\left[X_n^2 \mid \mathcal{F}_m\right] + X_m^2 - 2X_m^2 = \mathbb{E}\left[X_n^2 \mid \mathcal{F}_m\right] - X_m^2
\end{aligned}$$

Zadanie 3.
Niech $\{Z_n\}$ będzie procesem Galtona-Watsona dla którego $\mathbb{E}[Z_1] = \mu$ oraz $\mathrm{Var}(Z_1) = \sigma^2 < \infty$. Rozważmy martyngał $M_n = \mu^{-n} Z_n$.

(a) Pokaż, że

$$\mathbb{E}\left[M_n^2\right] = 1 + \sigma^2 \sum_{k=2}^{n+1} \mu^{-k}$$

(b) Uzasadnij, że jeśli $\mu > 1$, to M_n jest zbieżny w L^2

(c) Uzasadnij, że jeśli $\mu < 1$, to M_n nie jest zbieżny w L^2.

Rozwiązanie.
Proces Galtona-Watsona pojawił się w rozdziale 4.3, gdy chcieliśmy obserwować pantofelki rozmnażające się bezpłciowo, niezależnie od siebie. Rozważaliśmy zmienne losowe $Y_{n,k}$ takie oraz ciąg

$$Z_1 = 1$$

$$Z_{n+1} = \sum_{k=1}^{Z_n} Y_{n+1,k}$$

gdzie Z_n to liczba nowych pantofelków w n-tej generacji, a $Y_{n,k}$ to liczba potomstwa w n-tej generacji zrodzona przez k-tego pantofelka w $n-1$ generacji.

(a) Wiem już, że

$$\mathbb{E}[Z_1] = \mathbb{E}\left[\sum_{k=1}^{Z_0} Y_{n,k}\right] = \mathbb{E}\left[Y_{1,1}\right] = \mu$$

oraz (korzystając z tego co pokazaliśmy na wykładzie):

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}\left[Y_{n,k}\right]\mathbb{E}[Z_n] = \mu\mathbb{E}[Z_n] = \mu^{n+1}.$$

Całość pokażemy za pomocą indukcji. Jeśli $n = 1$, to mamy

$$\mathbb{E}\left[M_1^2\right] = \mathbb{E}\left[\mu^{-2}Z_1^2\right] = \mu^{-2}\mathbb{E}\left[Z_1^2\right] = \mu^{-2}[\mathrm{Var}(Z_1) + \mathbb{E}[Z_1]^2] = \sigma^2\mu^{-2} + 1$$

tak jak chcieliśmy.

Zróbmy teraz krok indukcyjny, czyli $n \Rightarrow n + 1$. Będziemy korzystać z zadania 2, więc chcemy wyliczyć $M_{n+1} - M_n$

$$M_{n+1} - M_n = \mu^{-n-1}\sum_{k=1}^{Z_n} Y_{n,k} - \mu^{-n}Z_n = \mu^{-n-1}\sum_{k=1}^{Z_n}(Y_{n,k} - \mu)$$

Wstawiając do równości w zadaniu 2 (po scałkowaniu), dostajemy

$$\begin{aligned}\mathbb{E}\left[M_{n+1}^2\right] &= \mathbb{E}\left[(M_{n+1} - M_n)^2\right] + \mathbb{E}\left[M_n^2\right] = \\ &= \mathbb{E}\left[\left(\mu^{-n-1}\sum_{k=1}^{Z_n}(Y_{n,k} - \mu)\right)^2\right] + 1 + \sigma^2\sum_{k=2}^{n+1}\mu^{-k} = \\ &= \mu^{-2n-2}\mathbb{E}\left[\left(\sum_{k=1}^{Z_n}(Y_{n,k} - \mu)\right)^2\right] + 1 + \sigma^2\sum_{k=2}^{n+1}\mu^{-k} = (\star)\end{aligned}$$

Pozostaje pokazać, że $\mathbb{E}\left[\left(\sum_{k=1}^{Z_n}(Y_{n,k} - \mu)\right)^2\right] = \mu^n\sigma^2$.

Rozważmy funkcję

$$\begin{aligned}h(z) = \mathrm{Var}(\sum_{k=1}^{z} Y_{n+1,k}) &= \mathbb{E}\left[\left(\sum_{k=1}^{z} Y_{n+1,k} - z\mu\right)^2\right] = \\ &= \sum_{k=1}^{z}\mathrm{Var}(Y_{n+1,k}) = z\mathrm{Var}(Y_{1,1}) = z\mathrm{Var}(Z_1) = z\sigma^2\end{aligned}$$

i zauważmy, że wwo

$$\mathbb{E}\left[\left(\sum_{k=1}^{Z_n}(Y_{n+1,k} - \mu)\right)^2 \mid \mathcal{F}_n\right] = h(Z_n) = Z_n\sigma^2$$

czyli całkując obie strony otrzymujemy

$$\mathbb{E}\left[\left(\sum_{k=1}^{Z_n}(Y_{n+1,k}-\mu)\right)^2\right]=\mathbb{E}\left[Z_n\sigma^2\right]=\mu^n\sigma^2$$

i to jest tym co chcieliśmy, bo wracając do kroku indukcyjnego

$$(\star)=\mu^{-n-2}\mu^n\sigma^2+1+\sigma^2\sum_{k=2}^{n+1}\mu^{-k}=1+\sigma^2\sum_{k=2}^{n+2}\mu^{-k}$$

(b) Aby M_n zbiegało w L^2, musimy znaleźć funkcję X taką, że $\mathbb{E}\left[|M_n-X|^2\right]\to 0$. Zauważmy, że Z_n zawsze przyjmuje skończone wartości, czyli $M_n=\mu^{-n}Z_n\to 0$, bo $\mu^{-n}\to 0$. W takim razie mamy $X=0$ jest potencjalną granicą M_n. Wystarczy wyliczyć

$$\begin{aligned}\mathbb{E}\left[|M_n|^2\right]&=1+\sigma^2\sum_{k=2}^{n+1}\mu^{-k}=1+\sigma^2\sum_{k=0}^{n-1}\mu^{-k-2}=\\&=1+\sigma^2\mu^{-2}\sum_{k=0}^{n-1}\mu{-}k=1+\sigma^2\mu^{-2}\sum_{k=0}^{n-1}\mu^{-k}=\\&=1+\sigma^2\mu^{-2}\frac{1-\mu^{-n}}{1-\mu^{-1}}=1+\sigma^2\mu^{-1}\frac{1-\mu^{-n}}{\mu-1}\leq 1+\sigma^2\mu^{-1}\frac{1}{\mu-1}<\infty\end{aligned}$$

czyli stosuje się do tego twierdzenie 7.5, ponieważ spełniony jest warunek $\sup\mathbb{E}\left[|M_n|^p\right]<\infty\iff$ istnieje M_∞ że $M_n\to M_\infty$ w L^p dla $p=2$.

(c) Argument jest prawie taki sam jak wyżej, z tym, że

$$\mathbb{E}\left[|M_n|^2\right]=1+\sigma^2\mu^{-2}\frac{1-\mu^{-n}}{1-\mu^{-1}}=1+\sigma^2\mu^{-1}\frac{1-\mu^{-n}}{\mu-1}$$

jest cały czas rosnące, czyli $\sup|M_n|^p=\infty$ i wówczas musimy mieć fałsz w warunku 2, czyli istnieniu M_∞ takiego, że $M_n\to M_\infty$ w L^p.

Zadanie 4.
Niech $f:[0,1]\to\mathbb{R}$ będzie funkcją spełniającą warunek Lipschitza: istnieje $L>0$ takie, że

$$|f(x)-f(y)|\leq L|x-y|$$

dla wszystkich $x,y\in[0,1]$. Celem zadania jest pokazanie, że f jest całką z ograniczonej funkcji mierzalnej. Niech X będzie zmienną o rozkładzie jednostajnym na $[0,1)$. Dla każdego $n\in\mathbb{N}$ położmy

$$X_n=2^{-n}\lfloor 2^nX\rfloor,\qquad Z_n=2^n(f(X_n+2^{-n})-f(X_n))$$

oraz $\mathcal{F}_n=\sigma(X_0,X_1,\ldots,X_n)$.

(a) Pokaż, że $\sigma(X_0,X_1,\ldots,)=\sigma(X)$ i $\mathcal{F}_n=\sigma(X_n)$

(b) Dla ograniczonej funkcji mierzalnej $h : [0,1] \to \mathbb{R}$ znajdź $\mathbb{E}\left[h(X_{n+1}) \mid \mathcal{F}_n\right]$. Wywnioskuj, że $(X_n)_{n\in\mathbb{N}}$ jest martyngałem.

(c) Pokaż, że $Z_n \to g(X)$ dla pewnej ograniczonej, mierzalnej funkcji $g : [0,1] \to \mathbb{R}$.

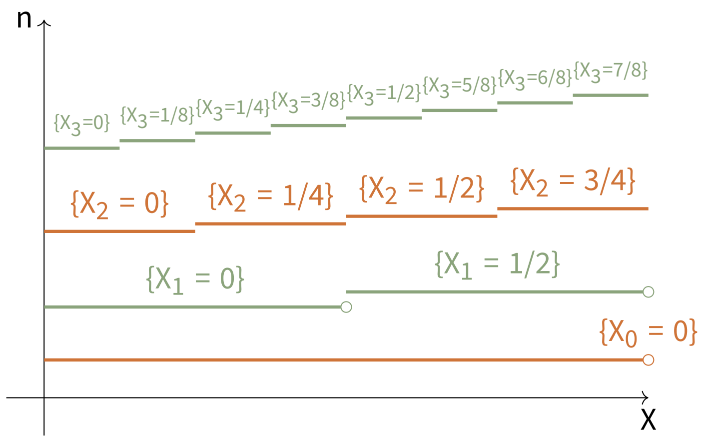
(d) Pokaż, że

$$Z_n = 2^n \int_{X_n}^{X_n+2^{-n}} g(u)du$$

(e) Wywnioskuj, że dla każdego $x \in [0,1]$

$$f(x) = f(0) + \int_0^x g(u)du$$

Rozwiązanie. (a) Zacznijmy od obserwacji, że każdy X_n dzieli X na odcinki $[k2^{-n}, (k+1)2^{-n})$, $k = 0, 1, \ldots, 2^n - 1$:



Moim zdaniem to już pokazuje równość. Dowolny zbiór z RHS wystarczy rozbić na wystarczająco dużo (nadal przeliczalnie) zbiorów postaci $[k2^{-n}, (k+1)2^{-n})$ dla odpowiednich n. Przejście z RHS do LSH jest nawet bardziej trywialne, bo każdy zbiór z $\sigma(X_n)$ można zapisać jako $\{X \in B\}$ dla $B \in Bor[0,1)$ tak jak na obrazku wyżej.

(b) Tak jak już ustaliliśmy w poprzednim podpunkcie, wiemy, że zbiory z ciała $\mathcal{F}_n$ są generowane przez zdarzenia $\{X \in [k2^{-m}, (k+1)2^{-m})\}$ dla $m = 0, 1, \ldots, n$ i $k = 0, 1, \ldots, 2^m - 1$. Weźmy więc jeden z takich zbiorów $A = \{X \in [k2^{-m}, (k+1)2^{-m}\} \in \mathcal{F}_n$ i popatrzmy jak powinno wyglądać wwo

$$\mathbb{E}\left[\mathbb{E}\left[h(X_{n+1}) \mid \mathcal{F}_n\right] \mathbb{1}_A\right] = \mathbb{E}\left[h(X_{n+1})\mathbb{1}_A\right]$$

Zauważmy, że jeśli $n+1 = m+l$, to X_{n+1} ma 2^l segmentów na A, tzn.

$$X_{n+1}\mathbb{1}_{\{X\in[k2^{-m},(k+1)2^{-m})\}} = \sum_{i=0}^{2^l-1} X_n \mathbb{1}_{\{X\in[k2^{-m}+i2^{-(n+1)},k2^{-m}+(i+1)2^{-(n+1)}\}}$$

i na każdym ze zbiorów $\{X \in [k2^{-m} + i2^{-(n+1)}.k2^{-m} + (i+1)2^{-(n+1)})\}$ zmienna X_n przyjmuje wartość $k2^{-m} + i2^{-(n+1)}$. Mając tę wiadomość z tyłu głowy, możemy wrócić do szukania wwo.

$$\begin{aligned}\mathbb{E}\left[h(X_{n+1})\mathbb{1}_A\right] &= \mathbb{E}\left[\sum_{i=0}^{2^l-1} h(X_{n+1})\mathbb{1}_{\{X\in[k2^{-m}+2^{-(n+1)},k2^{-m}+(i+1)2^{-(n+1)})\}}\right] = \\ &= \mathbb{E}\left[\sum_{i=0}^{2^l-1} h(k2^{-m}+i2^{-(n+1)})\mathbb{1}_{\{X\in[k2^{-m}+i2^{-(n+1)},k2^{-m}+(i+1)2^{-(n+1)})\}}\right] = \\ &= \sum_{i=0}^{2^l-1} h(k2^{-m}+i2^{-(n+1)})\mathbb{P}\left[X\in[k2^{-m}+i2^{-(n+1)},k2^{-m}+(i+1)2^{-(n+1)})\right] = \\ &= \sum_{i=0}^{2^l-1} h(k2^{-m}+i2^{-(n+1)})2^{-(n+1)}\end{aligned}$$

z drugiej strony, policzmy $\mathbb{E}\left[1/2(h(X_n)+h(X_n+2^{-(n+1)})\mathbb{1}_A\right]$

$$\begin{aligned}\mathbb{E}\left[1/2(h(X_n)+h(X_n+2^{-(n+1)})\mathbb{1}_A\right] &= \frac{1}{2}[\mathbb{E}\left[h(X_n)\mathbb{1}_A\right]+\mathbb{E}\left[h(X_n+2^{-(n+1)})\right]] = \\ &= \frac{1}{2}\sum_{i=0}^{2^{l-1}-1} 2^{-n}h(k2^{-m}+i2^{-n})+ \\ &\quad+\frac{1}{2}\sum_{i=0}^{2^{l-1}-1} 2^{-n}h(k2^{-m}+i2^{-n}+2^{-(n+1)}) = \\ &= \sum_{i=0}^{2^{l-1}-1} 2^{-(n+1)}h(k2^{-m}+(2i)2^{-(n+1)})+ \\ &\quad+\sum_{i=0}^{2^{l-1}-1} 2^{-(n+1)}h(k2^{-m}+(2i+1)2^{-(n+1)})\end{aligned}$$

pierwszy wyraz odpowiada za parzyste fragmenty $\mathbb{E}\left[h(X_{n+1})\mathbb{1}_A\right]$, a drugi odpowiada za nieparzyste. Stąd mogę powiedzieć, że

$$\mathbb{E}\left[h(X_{n+1}) \mid \mathcal{F}_n\right] = 1/2[h(X_n)+h(X_n+2^{-(n+1)})]$$

Jeśli teraz stwierdzimy, zamiast h podstawimy $f(X_n+2^{-n})-f(X_n)$, co jest ograniczone przez $L2^{-n}\leq L$ z definicji funkcji f (ale nie jestem świadoma gdzie konkretnie używałam ograniczoności h, może żeby wyciągnąć sumę przed całkę?), dostajemy

$$\begin{aligned}\mathbb{E}\left[Z_{n+1}\mid\mathcal{F}_n\right] &= \mathbb{E}\left[2^{n+1}[f(X_{n+1}+2^{-(n+1)})-f(X_{n+1})] \mid \mathcal{F}_n\right] = \\ &= 2^{n+1}[\mathbb{E}\left[f(X_{n+1}+2^{-(n+1)})\mid\mathcal{F}_n\right]-\mathbb{E}\left[f(X_{n+1})\mid\mathcal{F}_n\right]] = \\ &= 2^{n+1}[1/2[f(X_n+2^{-(n+1)})+f(X_n+2\cdot 2^{-(n+1)})]-1/2[f(X_n)+f(X_n+2^{-(n+1)})]] = \\ &= 2^n[f(X_n+2^{-n})-f(X_n)] = Z_n\end{aligned}$$

(c) Nie mam pojęcia jak to zrobić, ale z twierdzenia Dooba o zbieżności martyngałów wiemy, że istnieje zmienna losowa całkowalna Z taka, że $Z_n \to Z$, bo

$$\sup \mathbb{E}\left[Z_n^+\right] = \sup \mathbb{E}\left[2^n(f(X_n + 2^{-n}) - f(X_n))\right] \leq \sup \mathbb{E}\left[2^n L|(X_n + 2^{-n}) - X_n|\right] = \sup \mathbb{E}[L] = L < \infty.$$

Ale przecież $Z_n = g(X_n)$ dla pewnej funkcji ograniczonej g (jak wyżej), czyli $\lim Z_n = \lim g(X_n) = g(X)$, bo $X_n \to X$.

Zadanie 5.
Niech X_n będzie ciągiem zmiennych losowych przyjmujących wartości w $[0, 1]$ takim, że $X_0 = a \in [0, 1]$. Dla $n \in \mathbb{N}$ niech $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$. Załóżmy, że

$$\mathbb{P}\left[X_{n+1} = X_n/2 \mid \mathcal{F}_n\right] = 1 - X_n, \quad \mathbb{P}\left[X_{n+1} = (1 + X_n)/2 \mid \mathcal{F}_n\right] = X_n$$

(a) Pokaż, że (X_n) jest martyngałem. Wywnioskuj, że X_n jest zbieżne p.w. do pewnej zmiennej losowej Z.

(b) Pokaż, że $4\mathbb{E}\left[(X_{n+1} - X_n)^2\right] = \mathbb{E}\left[X_n(1 - X_n)\right]$

(c) Znajdź rozkład zmiennej losowej Z

Rozwiązanie.

Wykład Łańcuchy Markowa

Przykład(y) 8.1

1. Alicja i Bob zapisują na kartce kolejne wyniki rzutów DOPISAĆ ZE ZDJĘCIA

   Jakie jest prawdopodbieństwo, że zakład wygra Alicja?

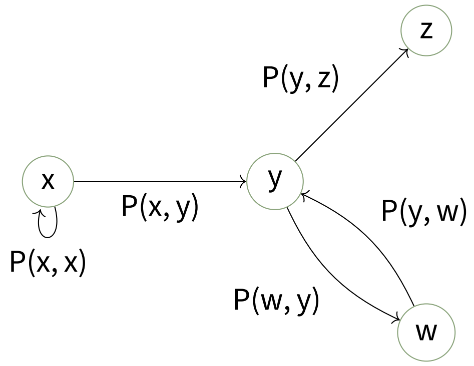
   Do określenia stanu gry wystarczy znajomość dwóch ostatnich znaków.

Definicja 8.1 : macierz stochastyczna.
Niech S będzie co najwyżej przeliczalną przestrzenią. Funkcję $P : S \times S \to [0, 1]$ nazywamy **macierzą przjeścia** (m. stochastyczną), jeżeli

$$\sum_{y \in S} P(x, y) = 1$$

dla każdego $x \in S$.

O P możemy myśleć jako o wagach krawędzi na grafie skierowanym



Wszystkie strzałki wchodzące do danego wierzchołka w tym grafie sumują się do jedynki.

Definicja 8.2 : łańcuch Markowa.
Jeśl iS jest zbiorem co najwyżej przeliczalny, to $\{X_n\}$ jest **łańcuchem Markowa z funkcją przejścia** P $\iff$ dla każdych $x_0, x_1, ..., x_n \in S$ takich, że $\mathbb{P}[X_n = x_n, ..., X_0 = x_0] > 0$ zachodzi

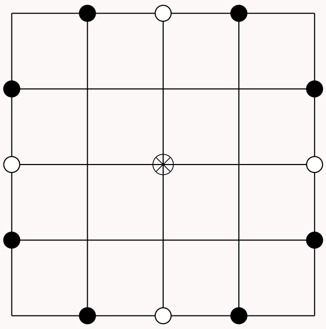
$$\mathbb{P}\left[X_{n+1} = x_{n+1} \mid X_0 = x_0, ..., X_n = x_n\right] = \mathbb{P}\left[X_{n+1} = x_{n+1} \mid X_n = x_n\right] = P(x_n, x_{n+1})$$

Łańcuch Markowa jest procesem bez pamięci, tzn. jeśli mamy wartość X_n to tak naprawdę możemy znaleźć wartość X_{n+k} dla każdego k, bo każdy X_n zależy tylko do X_{n-1}.

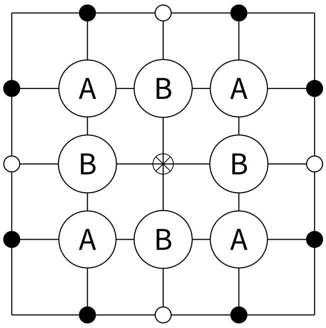
8.1 Zadania

Zadanie 1.
Policja w Nowym Yorku próbuje złapać przestępce znajdującego się w punkcie $\otimes$. Obstawiła część ulic, ale nie wszystkie. Przestępca w każdym korku porusza się losowo (tzn. z prawdopodobieństwem 1/4 w każdym z możliwych kierunków). Jeżeli wpadnie na policję $\bullet$ zostaje złapany, jeżeli dotrze do jednego z pól $\circ$ ucieka. Oblicz prawdopodobieństwo, że uda mu się

uciec.



Rozwiązanie.
Zacznę od oznaczenia pozycji literkami:



oraz $\otimes$ jest oznaczone S, $\bullet$ to 0 oraz $\circ$ będzie 1. Niech P_i oznacza prawdopodobieństwo ucieczki jeśli jesteśmy na polu i. Wtedy

$$\begin{cases} P_S = \frac{1}{4}(P_B + P_B + P_B + P_B) = P_B \\ P_B = \frac{1}{4}(P_1 + P_S + P_A + P_A) = \frac{1}{4}(P_1 + P_S) + \frac{1}{2}P_A \\ P_A = \frac{1}{4}(P_B + P_B + P_0 + P_0) = \frac{1}{2}(P_B + P_0) \\ P_1 = 1 \\ P_0 = 0 \end{cases}$$

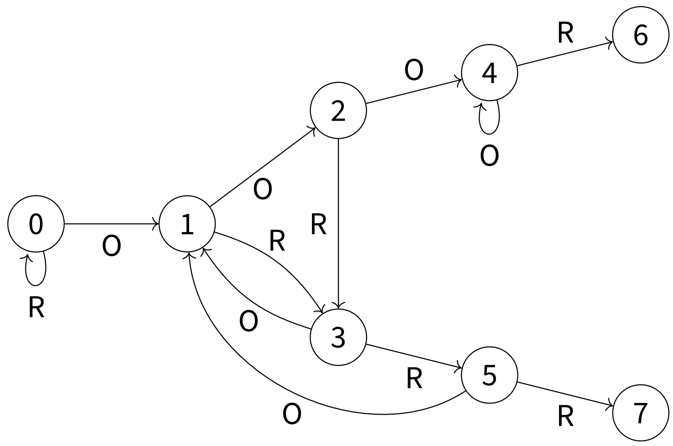
W takim razie $P_A = \frac{1}{2}P_B$ oraz

$$P_B = \frac{1}{4}(1 + P_S) + \frac{1}{4}P_B \Rightarrow 3P_B = 1 + P_S \Rightarrow 3P_B = 1 + P_B \Rightarrow P_B = \frac{1}{2}$$

ale ponieważ $P_S = P_B$, to odpowiedź końcowa wynosi $\frac{1}{2}$.

Zadanie 2.
Alicja i Robert rzucają symetryczną monetą tak długo aż wypadnie OOOR lub ORRR. Alicja wygrywa, gdy wzorzec OOOR wypadnie jako pierwszy, natomiast Robert, gdy wypadnie ORRR. Oblicz prawdopodobieństwo, że grę wygra Alicja.

Rozwiązanie.
Tak jak na wykładzie, rysujemy graf



Alicja wygrywa w wierzchołku 6 z $\mathbb{P}$ 1, a w wierzchołku 8 wygrywa z $\mathbb{P}$ równym 0. Rozpisując to cudeńko jako równanie rekurencyjne, gdzie P_i to prawdopodobieństwo, że będąc w wierzchołku i wygra Alicja, dostajemy:

$$\begin{cases} P_6 = 1 \\ P_8 = 0 \\ P_4 = \frac{1}{2}P_6 + \frac{1}{2}P_4 \\ P_2 = \frac{1}{2}P_4 + \frac{1}{2}P_3 \\ P_1 = \frac{1}{2}P_2 + \frac{1}{2}P_3 \\ P_3 = \frac{1}{2}P_1 + \frac{1}{2}P_5 \\ P_5 = \frac{1}{2}P_1 + \frac{1}{2}P_8 \end{cases}$$

Czyli $P_5 = \frac{1}{2}P_1$, a

$$P_3 = \frac{1}{2}P_1 + \frac{1}{2}P_5 = \frac{1}{2}P_1 + \frac{1}{4}P_1 = \frac{3}{4}P_1$$

idąc z tym do P_1

$$P_1 = \frac{1}{2}P_2 + \frac{1}{2}P_3 = \frac{1}{2}P_2 + \frac{3}{8}P_1$$

z P_4 dostaniemy

$$P_4 = \frac{1}{2}P_6 + \frac{1}{2}P_4 = \frac{1}{2} + \frac{1}{2}P_4 \Rightarrow \frac{1}{2}P_4 = \frac{1}{2} \Rightarrow P_4 = 1$$

i przechodzimy do P_2 na chwilę

$$P_2 = \frac{1}{2}P_4 + \frac{1}{2}P_3 = \frac{1}{2} + \frac{3}{8}P_1$$

podstawiając z powrotem do P_1

$$P_1 = \frac{1}{2}P_2 + \frac{3}{8}P_1 = \frac{1}{4} + \frac{3}{16}P_1 + \frac{6}{16}P_1 = \frac{1}{4} + \frac{9}{16}P_1$$

czyli $\frac{7}{16}P_1 = \frac{1}{4} \Rightarrow P_1 = \frac{4}{7}$, a ponieważ

$$P_0 = \frac{1}{2}P_0 + \frac{1}{2}P_1 \Rightarrow P_1 = P_0$$

to prawdopodobieństwo, że wygra Ala wynosi $\frac{4}{7}$.