Politechnika Gdańska
Wydział Fizyki Technicznej i Matematyki Stosowanej

Projekt licencjacki

Twierdzenie o dezintegracji miary

Autorzy:

Bartosz Bycul
Piotr Fonferek
Grzegorz Łucki
Jakub Sadowy

Opiekun projektu:

Dr Wojciech Czernous

Gdańsk, 2023

1 Podstawowe zagadnienia z teorii miary

Definicja 1 (π-układ)**.** Rodzinę zbiorów $\mathcal{H}$ nazywamy π-układem, jeśli $A, B \in \mathcal{H} \Rightarrow A \cap B \in \mathcal{H}$

Definicja 2 (generator)**.** Jeśli $(s, \mathcal{S})$ jest przestrzenią mierzalną, zaś $\mathcal{C} \subset \mathcal{S}$ jest π-układem, takim że $S \in \mathcal{C}$ i $\sigma(\mathcal{C}) = \mathcal{S}$, to $\mathcal{C}$ będziemy nazywać generatorem S.

Definicja 3 (λ-układ)**.** Niech X będzie niepustym zbiorem. Rodzinę zbiorów $\mathcal{H} \subseteq 2^X$ nazywamy λ-układem , gdy:

(i) $X \in \mathcal{H}$,

(ii) $\forall_{A,B\in\mathcal{H}} [B \subset A \Rightarrow A \setminus B \in \mathcal{H}]$,

(iii) $[\{A_1, A_2, A_3, ...\} \subseteq \mathcal{H}, \forall_{n\in\mathbb{N}} A_n \subseteq A_{n+1}] \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{H}$.

Powyższe obiekty pozwalają nam sformułować twierdzenie Sierpińskiego, szczególnie użyteczne do rozszerzania własności określonych na π-układach, do generowanych przez nie σ-ciał.

Twierdzenie 1 (o klasach monotonicznych, Sierpiński)**.** *Niech $\mathcal{C}$ będzie π-układem, zaś $\mathcal{D}$ będzie λ-układem w pewnej przestrzeni Ω i niech $\mathcal{C} \subset \mathcal{D}$. Wtedy $\sigma(\mathcal{C}) \subset \mathcal{D}$.*

Lemat 1 (jednoznaczność)**.** *Niech μ i ν będą miarami ograniczonymi na przestrzeni mierzalnej $(\Omega, \mathcal{A})$ oraz niech $\mathcal{C}$ będzie generatorem $\mathcal{A}$. Wtedy $\mu = \nu$, jeżeli $\mu A = \nu A$ dla każdego $A \in \mathcal{C}$.*

Dowód. Zakładając, że $\mu = \nu$ na $\mathcal{C}$, niech $\mathcal{D} = \{A \in \mathcal{A} : \mu A = \nu A\}$. Udowodnimy, że $\mathcal{D}$ jest λ-układem:

(i) z założenia, że $\Omega \in \mathcal{C}$ mamy $\Omega \in \mathcal{D}$;

(ii) niech $A, B \in \mathcal{D}$ i $A \in B$, wtedy z własności miary:

$$\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A),$$

czyli $B \setminus A \in \mathcal{D}$;

(iii) dla wstępującego ciągu zbiorów $A_n \in \mathcal{D}$, z twierdzenia 2 mamy

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \mu(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \nu(A_n) = \nu(\bigcup_{n=1}^{\infty} A_n),$$

czyli $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

Stąd i z twierdzenia 1 wynika $\mathcal{D} \supset \sigma(\mathcal{C}) = \mathcal{A}$, co znaczy, że $\mu = \nu$. □

Lemat 2 (ciągłość)**.** *Niech μ będzie miarą na $(\Omega, \mathcal{A})$ oraz $A_1, A_2, ... \in \mathcal{A}$. Wtedy:*

(i) jeżeli $A_n \uparrow A$, to $\mu A_n \uparrow \mu A$;

(ii) jeżeli $A_n \downarrow A$ i $\mu A_1 < \infty$, to $\mu A_n \downarrow \mu A$.

Dowód. Dla (1) możemy zastosować własność przeliczalnej addytywności dla $D_n = A_n \setminus A_{n-1}$ z $A_0 = \varnothing$. Aby otrzymać (2) zastosujmy (1) do zbiorów $B_n = A_1 \setminus A_n$. □

Poniżej określimy kilka przydatnych faktów dotyczących całek Lebesgue'a.

Definicja 4. Dla przestrzeni mierzalnych $(S, \mathcal{S})$ i $(T, \mathcal{T})$ oraz mierzalnego odwzorowania $f : S \to T$, mając określoną miarę μ na $(S, \mathcal{S})$, można zdefiniować miarę $\mu \circ f^{-1}$ na $(T, \mathcal{T})$ jako:

$$(\mu \circ f^{-1})B = \mu(f^{-1}B) = \mu\{s \in S; f(s) \in B\},\ B \in \mathcal{T}.$$

Przeliczalna addytywność $\mu \circ f^{-1}$ wynika z faktu, że μ jest miarą oraz f^{-1} zachowuje sumy i przekroje.

Uwaga 1 (Miara na funkcji)**.** Całkę funkcji mierzalnej f względem miary μ oznaczamy przez

$$\mu f = \int f d\mu = \int f(\omega)\mu(d\omega).$$

Lemat 3 (całkowanie przez podstawienie)**.** *Rozważmy przestrzeń z miarą $(\Omega, \mathcal{A}, \mu)$, przestrzeń mierzalną $(S, \mathcal{S})$ oraz dwa mierzalne odwzorowania, $f : \Omega \to S$ i $g : S \to \mathbb{R}$. Wtedy*

$$\mu(g \circ f) = (\mu \circ f^{-1})g,$$

gdy któraś ze stron istnieje.

Dowód. Niech $A \in \mathcal{A}$. Jeśli $g = \mathbb{1}_A$, to

$$\mu(g \circ f) = \mu(\mathbb{1}_A \circ f) = \mu\mathbb{1}\{f \in A\} = \mu\{\omega : f(\omega) \in A\} = (\mu \circ f^{-1})(A) = (\mu \circ f^{-1})\mathbb{1}_A = (\mu \circ f^{-1})g$$

Z liniowości całki, $\mu(g \circ f) = (\mu \circ f^{-1})$ zachodzi więc dla funkcji prostych g. Korzystając z twierdzenia 2, $g_n \uparrow g \Rightarrow \mu(g_n \circ f) = (\mu \circ f^{-1})g_n \uparrow (\mu \circ f^{-1})g = \mu(g \circ f)$, gdy g_n są funkcjami prostymi, a takowe można znaleźć dla dowolnej mierzalnej $g \geqslant 0$. W ogólnym przypadku, teza wynika z rozbicia $g = g^+ - g^-$ i z liniowości całki. $\square$

Twierdzenie 2 (Leviego o zbieżności monotonicznej)**.** *Niech $f_1,\ f_2,\ f_3,\ \ldots$ będą funkcjami mierzalnymi na $(\Omega, \mathcal{A}, \mu)$. Wtedy zachodzi*

$$0 \leqslant f_n \uparrow f \Longrightarrow \mu f_n \uparrow \mu f.$$

Definicja 5 (przestrzeń borelowska)**.** Przestrzenią borelowską nazywamy przestrzeń mierzalną, która jest borelowsko izomorficzna do borelowskiego podzbioru $[0, 1]$, tj. istnieje bijekcja $f : S \leftrightarrow A \in [0, 1]$, taka że zarówno f, jak i f^{-1}, są mierzalne.

Uwaga 2. Ponieważ przekątna S^2 jest domknięta dla przestrzeni Hausdorffa S (np. dla $[0, 1]$), a co za tym idzie, również dla jej podprzestrzeni $A \in \mathcal{B}[0, 1]$, to jest mierzalna również dla S borelowskiej. Izomorfizm mierzalny między S^2 a $[0, 1]^2$ budujemy jako produkt izomorfizmów. Rodzina zbiorów borelowskich w $[0, 1]^2$ jest σ-ciałem produktowym, zbudowanym jako produkt $\mathcal{B}([0, 1])$ (patrz twierdzenie 1.2 w [1]).

2 Jądra i warunkowość

Definicja 6 (Element losowy)**.** Rozważmy przestrzeń probabilistyczną $(\Omega, \mathcal{A}, P)$ i niech $(S, \mathcal{S})$ będzie przestrzenią mierzalną. Każde mierzalne przekształcenie $\xi : \ \Omega \to (S, \mathcal{S})$ czyli takie, że dla $B \in \mathcal{S}$ mamy $\xi^{-1}B \in \mathcal{A}$ nazywamy elementem losowym. Używamy oznaczenia:

$$\{\xi \in B\} := \xi^{-1}B \in \mathcal{A},$$

mamy więc

$$P\{\xi \in B\} = (P \circ \xi^{-1})B, \ B \in \mathcal{S}.$$

Miarę $P \circ \xi^{-1}$ będziemy nazywać rozkładem elementu losowego ξ i oznaczać przez $\mathcal{L}(\xi)$.

Uwaga 3. Jeżeli:

(i) $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, to ξ nazywamy zmienną losową;

(ii) $(S, \mathcal{S}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, to ξ nazywamy wektorem losowym;

(iii) $(S, \mathcal{S})$ jest przestrzenią funkcyjną, to ξ nazywamy procesem losowym;

(iv) $(S, \mathcal{S})$ jest klasą miar σ-skończonych na $(T, \mathcal{T})$ (tj. takich miar μ, że istnieje przeliczalne rozbicie A_n przestrzeni T i $\mu A_n < \infty$), przy czym $\mathcal{S}$ jest najmniejszym σ-ciałem, względem którego każda funkcja postaci $\mu \mapsto \mu B, B \in \mathcal{T}$, jest mierzalna; wówczas ξ nazywamy miarą losową.

Zajmiemy się bliżej tym ostatnim rodzajem elementu losowego. Charakteryzację przestrzeni miar losowych uzasadnimy w oparciu o następujące fakty.

Lemat 4 (funkcje mierzalne)**.** *Rozważamy odwzorowanie f pomiędzy dwiema przestrzeniami mierzalnymi $(S, \mathcal{S})$, $(T, \mathcal{T})$, i niech $\mathcal{C} \subset 2^T$ będzie takie, że $\sigma(\mathcal{C}) = \mathcal{T}$. Wówczas f jest mierzalna (względem $\mathcal{S}, \mathcal{T}$) wtedy i tylko wtedy, gdy $f^{-1}\mathcal{C} \subset \mathcal{S}$.*

Definicja 7 (σ-ciało generowane przez rodzinę przekształceń)**.** Weźmy rodzinę przestrzeni mierzalnych $(T_i, \mathcal{T}_i)$, $i \in I$, oraz funkcji $f_i : T \to T_i$. Najmniejsze σ-ciało podzbiorów T, względem którego wszystkie f_i są mierzalne, nazywamy indukowanym przez rodzinę $\{f_i\}$ i oznaczamy $\sigma(f)$. Z definicji, $\sigma(f) = \sigma\left(\bigcup_i f_i^{-1}\mathcal{T}_i\right)$.

Uwaga 4. Weźmy dwie przestrzenie mierzalne $(S, \mathcal{S})$ oraz $(T, \sigma(f))$, dla pewnej rodziny przekształceń $f_i : T \to T_i$. Wówczas funkcja $g : S \to T$ jest mierzalna wtedy i tylko wtedy, gdy dla każdego $i \in I$, złożenie funkcji $f_i \circ g : S \to T_i$ jest mierzalne. Rzeczywiście, z lematu 4, mierzalność g jest równoważna inkluzji $g^{-1}\left(\bigcup_i f_i^{-1}\mathcal{T}_i\right) \subset \mathcal{S}$, to znaczy relacji $g^{-1}f^{-1}\mathcal{T}_i = (f \circ g)^{-1}\mathcal{T}_i \subset \mathcal{S}, \ i \in I$.

Definicja 8 (Jądro)**.** Ustalmy dwie przestrzenie mierzalne $(S, \mathcal{S})$ oraz $(T, \mathcal{T})$. Przekształcenie $\mu : S \times \mathcal{T} \to \overline{\mathbb{R}}_+$ nazywamy jądrem z S do T, jeżeli funkcja $\mu_s B = \mu(s, B)$ jest $\mathcal{S}$-mierzalna dla każdego ustalonego $B \in \mathcal{T}$ i miarą probabilistyczną dla każdego ustalonego $s \in S$.

Uwaga 5 (operator całkowy)**.** Każdemu jądru μ odpowiada operator, który przekształca funkcję $f : T \to \mathbb{R}$ w jej całkę $\mu f : S \to \mathbb{R}$, $\mu f(s) = \int \mu(s, dt) f(t)$. Z liniowości i monotoniczności całki, funkcja μf jest $\mathcal{S}$-mierzalna.

Uwaga 6 (Jądro probabilistyczne)**.** Jeżeli jądro $\mu(s,T) = 1$ dla każdego $s \in S$, to będziemy nazywać je jądrem probabilistycznym.

Zauważmy, że przestrzeń $\mathcal{M}(T)$, σ-skończonych miar losowych na T, wyposażona jest w σ-ciało indukowane przez rodzinę przekształceń $\pi_S : \mu \mapsto \mu B,\ B \in \mathcal{T}$. Na mocy uwagi 4, funkcja $f : S \to \mathcal{M}(T)$ jest mierzalna wtedy i tylko wtedy, gdy $\pi_B \circ f$ jest funkcją mierzalną, dla każdego $B \in \mathcal{T}$, to jest wtedy, gdy f jest jądrem takim, że $f_s,\ s \in S$, są miarami σ-skończonymi.
Co ciekawe, zbiór miar probabilistycznych $\mathcal{P}(T)$ na T jest mierzalnym podzbiorem $\mathcal{M}(T)$, gdyż $\mathcal{P}(T) = \pi_T^{-1}(\{1\})$ jest zbiorem z σ-ciała indukowanego $\sigma(\pi_B :\ B \in \mathcal{T})$.

Lemat 5 (funkcjonał a miara)**.** *Dla przestrzeni mierzalnej $(S, \mathcal{S})$, niech S_+ będzie przestrzenią funkcji mierzalnych nieujemnych na S. Funkcjonał $F : S_+ \to \mathbb{R}_+$, spełniający warunki:*

(i) $F\xi = 0$, *gdy* $\xi \equiv 0$;

(ii) $F(\xi + \eta) = F\xi + F\eta$;

(iii) $F\xi_n \uparrow F\xi$, *gdy* $\xi_n \uparrow \xi$;

definiuje miarę μ na $\mathcal{S}$, daną wzorem $\mu A = F\mathbb{1}_A$. Ponadto, całka względem μ jest dana wzorem $\mu g = Fg$.

Dowód. Wynika to z definicji miary. Ponadto, Czytelnik zechce zauważyć, że z (ii) wynika $F(a\xi) = aF\xi$ dla $a \in \mathbb{Q}$, $a \geqslant 0$, zaś (iii) pozwala rozszerzyć tę własność na $a \in \mathbb{R}_+$. Wzór na całkę μg wynika więc z konstrukcji całki względem miary, dzięki liniowości i monotoniczności operatora F. □

Lemat 6 (miary losowe)**.** *Jeśli $\mathcal{C}$ jest generatorem $\mathcal{T}$, to dla odwzorowań $\pi_B : \mu \mapsto \mu B, B \in \mathcal{T}$, zachodzi: $\sigma(\pi_B : B \in \mathcal{T}) = \sigma(\pi_C : C \in \mathcal{C})$.*

Dowód. Rzeczywiście, załóżmy że $\mathcal{C} \subset \mathcal{D}$, gdzie

$$\mathcal{D} = \{A \in \mathcal{T} : \pi_A \text{ jest funkcją mierzalną}\}$$

Wówczas $T \in \mathcal{D}$, gdyż z założenia $T \in \mathcal{C} \subset \mathcal{D}$.
Po drugie, dla $A, B \in \mathcal{D}, A \subset B$, mamy

$$\pi_{B\backslash A}\mu = \mu(B\backslash A) = \mu B - \mu A = (\pi_B - \pi_A)\mu,$$

zatem $\pi_{B\backslash A}$ jest mierzalna, jako różnica takowych, czyli $B\backslash A \in \mathcal{D}$.
Po trzecie, dla $A_1, A_2, ... \in \mathcal{D}, A_n \uparrow A$, mamy

$$\pi_{A_n}\mu = \mu A_n \uparrow \mu A = \pi_A \mu,$$

zatem π_A jest mierzalna, będąc granicą funkcji mierzalnych. Stąd $\mathcal{D}$ jest $\lambda-$ układem i z twierdzenia 1, $\mathcal{T} = \sigma(\mathcal{C}) \subset D$, tj. każda π_B jest mierzalna, $B \in \mathcal{T}$. □

Wniosek 1. *By rodzina miar σ-skończonych μ_s, $s \in S$, była jądrem, wystarcza, gdy funkcje $s \mapsto \mu_s C$ są $\mathcal{S}$-mierzalne dla każdego $C \in \mathcal{C}$, o ile $\mathcal{C}$ jest generatorem $\mathcal{T}$.*

Lemat 7 (funkcjonał losowy)**.** *Dla przestrzeni mierzalnej $(S, \mathcal{S})$, niech $\xi : \Omega \times S_+ \to \mathbb{R}_+$ będzie takim odwzorowaniem, że dla każdego $\omega \in \Omega$, funkcjonał $\xi_\omega = \xi(\omega, \cdot)$ spełnia warunki lematu 5. Dalej, przypuśćmy, że $\mathcal{F}$ jest pewnym σ-ciałem w Ω, zaś $\mathcal{C}$ jest generatorem $\mathcal{S}$, i że dla każdego $C \in \mathcal{C}$, zmienna losowa $\omega \mapsto \xi_\omega \mathbb{1}_C$ jest $\mathcal{F}$-mierzalna. Wtedy ξ jest jądrem z $(\Omega, \mathcal{F})$ do $(S, \mathcal{S})$. W szczególności, dla dowolnej $f \in S_+$, zmienna losowa $\omega \mapsto \xi_\omega f$ jest $\mathcal{F}$-mierzalna.*

Dowód. Z lematu 5, każde ξ_ω jest miarą (skończoną, gdyż z definicji ξ mamy $\xi_\omega 1_S \in \mathbb{R}_+$). Zatem ξ jest jądrem, na mocy wniosku powyżej. Mierzalność $\omega \mapsto \xi_\omega f$ wynika z uwagi o operatorze całkowym. □

Lemat 8 (całka z jądra)**.** *Jeśli μ jest jądrem z $(S, \mathcal{S})$ do $(T, \mathcal{T})$, zaś ν jest miarą na $(S, \mathcal{S})$, to funkcja zbioru $\nu\mu : \mathcal{T} \to \mathbb{R}_+$, $(\nu\mu)B = \int \mu(s, B)\nu(ds)$, jest miarą.*

Dowód. Odwzorowanie

$$F : T_+ \to \mathbb{R}_+, \qquad F\xi = \int \left(\int \xi(t)\mu(s, dt) \right) \nu(ds),$$

spełnia warunki lematu 5, jest bowiem liniowe, a monotoniczność (iii) wynika z twierdzenia 2. □

Do dalszego badania zdefiniowanych przez nas obiektów, spróbujemy skonstruować $\mathcal{F}$-mierzalną zmienną losową, która będzie jednocześnie jądrem probabilistycznym pomiędzy dwoma przestrzeniami mierzalnymi. Zmienną tą będziemy dalej nazywać *prawdopodobieństwem warunkowym.*

Twierdzenie 3 (warunkowa wartość oczekiwana, Kołmogorow)**.** *Dla każdego σ-ciała $\mathcal{F} \subset \mathcal{A}$ istnieje jedyny operator liniowy $E^{\mathcal{F}} : L^1 \to L^1(\mathcal{F})$ (gdzie $L^1(\mathcal{F}) = \{\xi \in L^1 : \xi$ jest $\mathcal{F}$-mierzalne$\}$), którego wartościami są zmienne wyznaczone z dokładnością do zbiorów miary zero, taki że:*

(i) $E\left[E^{\mathcal{F}}\xi; A\right] = E\left[\xi; A\right]$, $\xi \in L^1, A \in \mathcal{F}$.

Operator $E^{\mathcal{F}}$ ma następujące własności (przy założeniu, że odpowiednie wyrażenia istnieją dla wartości bezwzględnych):

(ii) $\xi \geqslant 0$ *implikuje* $E^{\mathcal{F}}\xi \geqslant 0$ *p.n.;*

(iii) $E|E^{\mathcal{F}}\xi| \leqslant E|\xi|$;

(iv) $0 \leqslant \xi_n \uparrow \xi$ *implikuje* $E^{\mathcal{F}}\xi_n \uparrow E^{\mathcal{F}}\xi$ *p.n.;*

(v) $E^{\mathcal{F}}\xi\eta = \xi E^{\mathcal{F}}\eta$ *p.n., gdy ξ jest $\mathcal{F}$-mierzalne;*

(vi) $E(\xi E^{\mathcal{F}}\eta) = E(\eta E^{\mathcal{F}}\xi) = E(E^{\mathcal{F}}\xi)(E^{\mathcal{F}}\eta)$;

(vii) $E^{\mathcal{F}}E^{\mathcal{G}}\xi = E^{\mathcal{F}}\xi$ *p.n. dla każdego $\mathcal{F} \subset \mathcal{G}$.*

Definicja 9 (prawdopodobieństwo warunkowe)**.** Prawdopodobieństwo warunkowe zdarzenia $A \in \mathcal{S}$, pod warunkiem σ-ciała $\mathcal{F}$, definiujemy jako

$$P^{\mathcal{F}}A = E^{\mathcal{F}}\mathbb{1}_A \text{ lub } P[A|\mathcal{F}] = E[\mathbb{1}_A|\mathcal{F}],\ A \in \mathcal{A}.$$

Definicja 10 (regularny rozkład warunkowy)**.** Ustalmy σ-ciało $\mathcal{F} \subset \mathcal{A}$, przestrzeń mierzalną $(S, \mathcal{S})$ i element losowy $\xi : \Omega \to S$. Regularnym rozkładem warunkowym ξ pod warunkiem σ-ciała $\mathcal{F}$ nazywamy jądro probabilistyczne μ z $(\Omega, \mathcal{F})$ do $(S, \mathcal{S})$ (równoważnie, $\mathcal{F}$-mierzalną miarę losową na S), takie że dla każdego $B \in \mathcal{S}$

$$\mu(\omega, B) = (P^{\mathcal{F}}\{\xi \in B\})(\omega) \text{ dla p.w. } \omega \in \Omega.$$

Ogólniej, ustalmy przestrzeń mierzalną $(T,\mathcal{T})$ i element losowy $\eta:\Omega\to T$. Regularnym rozkładem warunkowym ξ pod warunkiem η nazywamy jądro probabilistyczne μ z T do S, takie że dla każdego $B\in\mathcal{S}$

$$\mu(\eta,B)=P^{\eta}\{\xi\in B\}\ \text{ p.n.}$$

Mówimy też, że μ jest wersją regularną $P^{\mathcal{F}}[\xi\in\cdot]$ (odpowiednio, $P^{\eta}[\xi\in\cdot]$). W przypadku, gdy ξ jest $\mathcal{F}$-mierzalna, $P[\xi\in B|\mathcal{F}]$ ma wersję regularną $\mathbb{1}\{\xi\in B\}$, natomiast kiedy ξ jest niezależna od $\mathcal{F}$, to ma wersję regularną postaci $P\{\xi\in B\}$.

Definiując warunkowy rozkład regularny za pomocą jądra probabilistycznego, nasuwa się pytanie o warunki jego istnienia. Odpowiedzi na nie udzieli nam twierdzenie o rozkładzie warunkowym. Jednak, aby je udowodnić, potrzebujemy poniższych lematów.

Lemat 9 (reprezentacja funkcyjna, Doob)**.** *Ustalmy funkcje mierzalne f i g, określone na przestrzeni Ω, o wartościach odpowiednio w przestrzeniach $(S,\mathcal{S})$, $(T,\mathcal{T})$, gdzie S jest borelowska. Wówczas f jest $g^{-1}\mathcal{T}$-mierzalna wtedy i tylko wtedy, gdy istnieje mierzalne odwzorowanie $h:T\to S$, takie że $f=h\circ g$. W szczególności, gdy $\mathcal{F}=\sigma(\eta)$, gdzie η jest zmienną losową, zaś ξ jest zmienną losową całkowalną, mamy istnienie funkcji borelowskiej h, takiej, że $E^{\mathcal{F}}\xi=h(\eta)$.*

Lemat 10 (miary Lebesgue'a - Stieltjesa)**.** *Niech F będzie prawostronnie ciągłą, niemalejącą funkcją na $\mathbb{R}$, taką, że $F(0)=0$. Wtedy istnieje lokalnie skończona miara μ na $\mathbb{R}$ (tj. taka, że $\mu B<\infty$ dla B ograniczonych), taka, że*

$$\mu(a,b]=F(b)-F(a),\ -\infty<a<b<\infty.$$

Twierdzenie 4 (rozkład warunkowy)**.** *Ustalmy przestrzeń borelowską $(S,\mathcal{S})$ i przestrzeń mierzalną $(T,\mathcal{T})$. Niech ξ i η będą elementami losowymi odpowiednio w S i T. Wtedy istnieje jądro probabilistyczne μ z T do S, takie że $P[\xi\in\cdot|\eta]=\mu(\eta,\cdot)$ p.n. i μ jest wyznaczone jednoznacznie $\mathcal{L}(\eta)$-p.w.*

Dowód. Bez straty ogólności, możemy przyjąć $S\in\mathcal{B}(\mathbb{R})$. Z lematu 9 dla każdego $r\in\mathbb{Q}$ możemy wybrać $\mathcal{T}$-mierzalną funkcję $f_r=f(\cdot,r):T\to[0,1]$ taką, że:

$$f(\eta,r)=P[\xi\leqslant r|\eta]\ p.n.,\ r\in\mathbb{Q}. \tag{1}$$

Zdefiniujmy zbiór $A=\{t\in T:\forall_{v,u\in\mathbb{Q},v\leqslant u}\lim_{r\to\infty}f_r=1,\ \lim_{r\to-\infty}f_r=0,f_v\leqslant f_u\}$. Zauważmy, że zbiór A został zdefiniowany dla f niemalejącej względem r. Zatem:

$$\begin{aligned}A&=\{t\in T:\forall_{v,u\in\mathbb{Q},v\leqslant u}\sup_r f_r=1,\inf_r f_r=0,f_v\leqslant f_u\}\\&=\bigcap_{v,u\in\mathbb{Q},v\leqslant u}\{t:f_v\leqslant f_u\}\cap\{t:\sup_r f_r=1\}\cap\{t:\inf_r f_r=0\}\end{aligned}$$

Z własności funkcji mierzalnych, $\{\sup_r f_r=1\},\{\inf_r f_r=0\}\in\mathcal{T}$. Podobnie, z mierzalności f_v-f_u, $\{f_v\leqslant f_u\}\in\mathcal{T}$. Zatem $A\in\mathcal{T}$ (jako przeliczany przekrój zbiorów mierzalnych). Ponieważ dla $u\leqslant v$ zachodzi $f_u(\eta)\leqslant f_v(\eta)$ p.n. (z twierdzenia 3 (ii)), a takich par $u,v\in Q$ jest przeliczalnie wiele, a także $f_r(\eta)\uparrow 1$ p.n. przy $r\to\infty$ z twierdzenia 3 (iv) i $f_r(\eta)\downarrow 0$ p.n. przy $r\to-\infty$ (co również wynika z twierdzenia 3 (iv), jako że wówczas $\mathbb{1}\{\xi>r\}\uparrow 1$, co implikuje $1-f(\eta,r)\uparrow 1$ p.n.), mamy więc $\eta\in A$ p.n. Ustalmy:

$$F(t,x)=\mathbb{1}_A(t)\inf_{r>x}f(t,r)+\mathbb{1}_{A^c}(t)\{x\geqslant 0\},\ x\in\mathbb{R},\ t\in T.$$

Zauważmy, że $F(t,\cdot)$ jest dystrybuantą na $\mathbb{R}$, dla każdego $t \in T$. Dalej, z twierdzenia 10, istnieje miara probabilistyczna $m(t,\cdot)$ na $\mathbb{R}$, spełniająca

$$m(t,(-\infty,x]) = F(t,x),\ x \in \mathbb{R},\ t \in T.$$

Widać, że funkcja F jest $\mathcal{T}$-mierzalna względem t dla każdego x. Rodzina $\mathcal{C}$ zbiorów $(-\infty,x],\ x \leqslant \infty$, jest generatorem $\mathcal{B}(\mathbb{R})$. Zauważmy, że funkcja m spełnia założenia wniosku 1 dla π-układu $\mathcal{C}$, jest zatem jądrem probabilistycznym z T do $\mathbb{R}$, gdyż $m(t,\mathbb{R}) = 1$. Korzystając z własności (iv) z twierdzenia 3 otrzymujemy:

$$m(\eta,(-\infty,x]) = F(\eta,x) = \inf_{r>x} f(\eta,r) = P[\xi \leqslant x|\eta]\ p.n.,\ x \in \mathbb{R}.$$

Dalej, ustalmy rodzinę zbiorów $\Lambda = \{B \in \mathcal{B}(\mathbb{R}) : m(\eta,B) = P[\xi \in B|\eta]\ p.n.\}$. Wykażmy, że Λ jest λ-układem:

(i) z własności miary probabilistycznej:

$$m(\eta,\mathbb{R}) = 1 = P[\xi \in \mathbb{R}|\eta]\ p.n.$$

więc $\mathbb{R} \in \Lambda$;

(ii) niech $A, B \in \Lambda$ oraz $A \subset B$, wówczas p.n.

$$\begin{aligned} m(\eta, B\setminus A) &= m(\eta,B) - m(\eta,A) = P[\xi \in B|\eta] - P[\xi \in A|\eta] = E^{\eta}[\mathbb{1}\{\xi \in B\} - \mathbb{1}\{\xi \in A\}] \\ &= P[\xi \in B \setminus A|\eta] \end{aligned}$$

(iii) ustalmy wstępujący ciąg $B_n \in \Lambda$ taki, że $B_n \uparrow B$. Korzystając ponownie z (iv) z twierdzenia 3, otrzymujemy:

$$m(\eta, \bigcup_{m=1}^{\infty} B_n) = \lim_{n\to\infty} m(\eta,B_n) = \lim_{n\to\infty} E^{\eta}\mathbb{1}\{\xi \in B_n\} = P[\xi \in B|\eta]\ p.n.$$

Wiedząc, że $\mathcal{C} \subset \Lambda$, z twierdzenia 1 mamy $\sigma(\mathcal{C}) \subset \Lambda$. Zatem $\Lambda = \mathcal{B}(\mathbb{R})$. Stąd dostajemy:

$$m(\eta,B) = P[\xi \in B|\eta]\ p.n.,\ B \in \mathcal{B}(\mathbb{R}).$$

W szczególności $m(\eta,S^c) = 0\ p.n.$, więc przyjmując jądro probabilistyczne

$$\mu(t,\cdot) = m(t,\cdot)\mathbb{1}\{m(t,S)=1\} + \delta_s\mathbb{1}\{m(t,S)<1\},\ t \in T,$$

gdzie $s \in S$ jest dowolnie ustalone, otrzymujemy:

$$\mu(\eta,B) = P[\xi \in B|\eta]\ p.n.,\ B \in \mathcal{B}(\mathbb{R}) \cap S.$$

Udowodnijmy, że μ jest jedyne. Niech μ' będzie jądrem probabilistycznym o powyższej własności. Wtedy:

$$\mu(\eta,(-\infty,r]) = P[\xi \leqslant r|\eta] = \mu'(\eta,(-\infty,r])\ p.n.,\ r \in \mathbb{Q}.$$

Z twierdzenia 1 dla π-układu będącego zbiorem półprostych postaci $(-\infty,r]$ oraz λ-układu analogicznego do Λ wynika, że $\mu(\eta,\cdot) = \mu'(\eta,\cdot)$ p.n. $\square$

3 Twierdzenie o dezintegracji miary

Celem tej sekcji jest rozszerzenie twierdzenia Fubiniego. Pokażemy, że wartości oczekiwane i warunkowe wartości oczekiwane można obliczyć, poprzez całkowanie po odpowiednich rozkładach regularnych, innymi słowy *dezintegrując* miarę produktową na jej jednowymiarowe składniki. Zaczniemy od wprowadzenia szczególnego przypadku twierdzenia o dezintegracji.

Lemat 11 (warunkowość)**.** *Niech ξ oraz η będą niezależnymi elementami losowymi w przestrzeniach mierzalnych $(S, \mathcal{S})$ i $(T, \mathcal{T})$. Weźmy mierzalne funkcje $f : S \times T \to \mathbb{R}$ oraz zdefiniujmy $g : S \to \mathbb{R}$ jako $g(s) = Ef(s, \eta)$, załóżmy też, że $E|f(\xi, \eta)| < \infty$. Wtedy $Ef(\xi, \eta) = Eg(\xi)$.*

Dowód. Niech μ oraz ν oznaczają odpowiednio rozkłady ξ i η. Zakładając, że $f \geqslant 0$ i przyjmując $g(s) = Ef(s, \eta)$, za pomocą lematu 3 oraz twierdzenia Fubiniego dostajemy:

$$Ef(\xi, \eta) = \int f(s,t)(\mu \otimes \nu)(dsdt) = \int \mu(ds) \int f(s,t)\nu(dt) = \int g(s)\mu(ds) = Eg(\xi) \tag{2}$$

W ogólności, (2) zachodzi dla funkcji f^+ i f^-, ponieważ obie są nieujemne. Ponadto mamy $|f^\pm| \leqslant |f|$, więc

$$E|f^\pm(s, \eta)| \leqslant E|f(s, \eta)| < \infty$$

Ponieważ $f = f^+ - f^-$, to (2) otrzymujemy z liniowości całki. $\square$

Twierdzenie 5 (dezintegracja)**.** *Ustalmy dwie przestrzenie mierzalne S i T, σ-algebrę $\mathcal{F} \subset \mathcal{A}$, i element losowy $\xi \in S$ taki, że $P[\xi \in \cdot|\mathcal{F}]$ ma wersję regularną ν. Następnie ustalmy $\mathcal{F}$-mierzalny element losowy η w T i mierzalną funkcję $f : S \times T$, gdzie $E|f(\xi, \eta)| < \infty$. Wtedy*

$$E[f(\xi, \eta)|\mathcal{F}] = \int \nu(ds) f(s, \eta) \ p.n. \tag{3}$$

Istnienie p.n oraz $\mathcal{F}$-mierzalność całki należy rozumieć jako część tezy. W szczególnym przypadku, kiedy $\mathcal{F} = \sigma(\eta)$ oraz $P[\xi \in \cdot|\eta] = \mu(\eta, \cdot)$ dla jądra probabilistycznego μ z T do S, (3) przybiera postać:

$$E[f(\xi, \eta)|\eta] = \int \mu(\eta, ds) f(s, \eta) \ p.n. \tag{4}$$

Całkując (3) oraz (4), otrzymujemy często stosowane równości:

$$Ef(\xi, \eta) = E \int \nu(ds) f(s, \eta) = E \int \mu(\eta, ds) f(s, \eta). \tag{5}$$

Jeżeli $\xi \perp\!\!\!\perp \eta$, to możemy przyjąć $\mu(\eta, \cdot) \equiv \mathcal{L}(\xi)$, zaś (5) sprowadza się do tezy z lematu 11.

Dowód. Niech $B \in \mathcal{S}$, $C \in \mathcal{T}$ oraz niech ν będzie regularną wersją $P[\xi \in B|\mathcal{F}]$. Możemy skorzystać z punktu (i) twierdzenia 3, aby otrzymać

$$\begin{aligned} P\{\xi \in B, \eta \in C\} &= E\left[\mathbb{1}_{\{\xi \in B\}}; \eta \in C\right] \overset{\text{(i)}}{=} E[E[\mathbb{1}_{\{\xi \in B\}}|\mathcal{F}]; \eta \in C] \\ &= E[P[\xi \in B|\mathcal{F}]; \eta \in C] = E[\nu(B); \eta \in C] = E \int \nu(ds) \mathbb{1}_{\{s \in B, \eta \in C\}}, \end{aligned}$$

gdzie ostatnia równość wynika z tożsamości dla funkcji podcałkowej:

$$\omega \mapsto \nu_\omega(B)\mathbb{1}\{\eta(\omega) \in C\} = \int \mathbb{1}\{s \in B,\ \eta(\omega) \in C\}\nu_\omega(ds),$$

zaś $\nu_\omega(B)$ oraz $\mathbb{1}\{\eta(\omega) \in C\}$ są $\mathcal{F}$-mierzalne względem ω. Zdefiniujmy odwzorowanie $\zeta : \Omega \times (S \times T)_+ \to \mathbb{R}_+$ następująco:

$$\zeta(\omega, f) = \int f(s, \eta(\omega))\nu_\omega(ds).$$

Zauważmy, że ζ spełnia założenia lematu o funkcjonale losowym: liniowość i monotoniczność względem f, a także $\mathcal{F}$-mierzalność funkcji $\omega \mapsto \zeta_\omega \mathbb{1}_{B\times C}$. Ta ostatnia własność wynika stąd, że

$$\zeta(\omega, 1_{B\times C}) = \int \mathbb{1}_{B\times C}(s, \eta(\omega))\nu_\omega(ds) = \nu_\omega(B)1\{\eta(\omega) \in C\}$$

jest iloczynem funkcji $\mathcal{F}$-mierzalnych. A zatem, ζ jest jądrem z Ω do $S \times T$.

Z lematu o funkcjonale losowym wynika również mierzalność ζf. Ponieważ rodzina zbiorów

$$\mathcal{C} = \{B \times C : B \in \mathcal{S}, C \in \mathcal{T}\}$$

jest generatorem $\mathcal{S} \otimes \mathcal{T}$, lemat o funkcjonale losowym daje $\mathcal{F}$-mierzalność ζf, czyli prawej strony równości (3), dla dowolnej $\mathcal{S} \otimes \mathcal{T}$-mierzalnej funkcji $f : S \times T \to \mathbb{R}_+$.

Zajmiemy się teraz dowodem pierwszej równości w (5). Wykazaliśmy, że:

$$E\mathbb{1}_{B\times C}(\xi, \eta) = E\int \nu(ds)\mathbb{1}_{B\times C}(s, \eta).$$

Innymi słowy, biorąc $A = B \times C$, mamy

$$\mu_1(A) = \mu_2(A),$$

gdzie $\mu_1 = E1\{(\xi, \eta) \in \cdot\} = P \circ (\xi, \eta)^{-1}$ jest rozkładem elementu losowego (ξ, η) w $S\times T$; natomiast $\mu_2 = P\zeta$ jest jądrem ζ, scałkowanym względem P. Oczywiście μ_1 jest miarą na $S \times T$; z lematu o całkowaniu jądra wynika, że taką miarą jest również μ_2. Łatwo zauważyć, że mamy tu miary probabilistyczne: μ_1, jest bowiem rozkładem, zaś μ_2, bowiem $\zeta_\omega(S \times T) = 1$. Ponieważ miary te są sobie równe na każdym zbiorze o tej postaci, co A, a więc na każdym elemencie $\mathcal{C}$ (generatora $\mathcal{S} \otimes \mathcal{T}$), z lematu o jednoznaczności miary wynika $\mu_1 = \mu_2$ na $\mathcal{S} \otimes \mathcal{T}$.

W szczególności, $\mu_1 f = \mu_2 f$, czyli otrzymujemy (5) dla wszystkich $\mathcal{S}\otimes\mathcal{T}$-mierzalnych, nieujemnych funkcji f:

$$Ef(\xi, \eta) = E\int \nu(ds)f(s, \eta).$$

Teraz ustalmy mierzalną funkcję $f : S\times T \to \mathbb{R}_+$, $Ef(\xi, \eta) < \infty$, oraz niech $A \in \mathcal{F}$ będzie dowolne. Zdefiniujmy, dla $(t, z) \in T \times \{0, 1\}$, $s \in S$:

$$f_1(s,(t, z)) = f(s, t),\ \ f_2(s,(t, z)) = z$$

Wówczas, gdy f jest $\mathcal{S} \otimes \mathcal{T}$-mierzalna, to f_1 i f_2 są mierzalne względem $S \otimes (\mathcal{T} \otimes \sigma(\{0, 1\})) = (\mathcal{S} \otimes \mathcal{T}) \otimes \sigma(\{0, 1\})$; mianowicie, dla $B \in \mathcal{B}(\mathbb{R})$, zachodzi:

$$f_1^{-1}(B) = f^{-1}(B) \times \{0, 1\},\ \ f_2^{-1}(B) = S \times T \times (\{0, 1\} \cap B);$$

co za tym idzie, $\tilde{f} = f_1 f_2$ jest $\mathcal{S} \otimes \mathcal{T} \otimes \sigma(\{0,1\})$-mierzalna.
Mamy ponadto

$$\tilde{f}(s, (\eta, \mathbb{1}_A)) = f(s, \eta)\mathbb{1}_A.$$

Traktując $(\eta, \mathbb{1}_A)$ jako $\mathcal{F}$-mierzalny element losowy w $T \times \{0,1\}$, z (5) otrzymujemy

$$E\tilde{f}(\xi, (\eta, \mathbb{1}_A)) = E \int \nu(ds)\tilde{f}(s, (\eta, \mathbb{1}_A)),$$

czyli

$$Ef(\xi, \eta)\mathbb{1}_A = E \int \nu(ds) f(s, \eta)\mathbb{1}_A, \ A \in \mathcal{F}.$$

Biorąc $f = f^+ - f^-$, dostajemy tą samą równość dla dowolnych f mierzalnych. Z definicji warunkowej wartości oczekiwanej wynika teza (3). Na koniec wystarczy zauważyć, że, z twierdzenia 4 istnieje $\mu(\eta, \cdot) = P(\xi \in \cdot | \eta)$ spełniające (4). $\square$

4 Zastosowania twierdzenia o dezintegracji miary

Definicja 11 (Równość według rozkładu)**.** Elementy losowe ξ, η są równe według rozkładu, co oznaczamy $\xi \stackrel{\mathrm{d}}{=} \eta$, gdy $\mathcal{L}(\xi) = \mathcal{L}(\eta)$.

Uwaga 7. Niech ξ, η będą elementami losowymi w przestrzeni mierzalnej S. Wówczas

$$\xi \stackrel{\mathrm{d}}{=} \eta \Leftrightarrow \forall_{f \in \hat{C}} Ef(\xi) = Ef(\eta),$$

gdzie $\hat{C}$ jest przestrzenią funkcji mierzalnych i ograniczonych o wartościach rzeczywistych. Wynika to z lematu 3.

Pokażemy teraz, jak regularne rozkłady warunkowe mogą posłużyć do konstrukcji elementów losowych o zadanych własnościach. W większości przypadków potrzebujemy tylko zmiennej losowej ϑ o rozkładzie $U(0,1)$, która jest niezależna od wszystkich wcześniej wprowadzonych elementów losowych i σ-ciał. Zatem zakładamy, że na podstawowej przestrzeni probabilistycznej jesteśmy w stanie określić taką zmienną. Możemy to zrobić bez straty ogólności, ponieważ powyższy warunek będzie spełniony dla rozszerzenia naszej przestrzeni. Wystarczy przyjąć

$$\hat{\Omega} = \Omega \times [0,1], \;\; \hat{\mathcal{A}} = \mathcal{A} \otimes \mathcal{B}[0,1], \;\; \hat{P} = P \otimes \lambda,$$

gdzie λ jest miarą Lebesgue'a na $[0,1]$. Wtedy $\vartheta(\omega, t) \equiv t$ ma rozkład $U[0,1]$ na $\hat{\Omega}$ i $\vartheta \perp\!\!\!\perp \mathcal{A}$. Ponadto, dowolny element losowy ξ na Ω, może być traktowany jako funkcja na $\hat{\Omega}$. Formalnie rzecz biorąc, zastępujemy ξ elementem losowym $\hat{\xi}(\omega, t) = \xi(\omega)$, mającym oczywiście ten sam rozkład. Zachodzi też następujący fakt.

Lemat 12 (Reprodukcja)**.** *Istnieją pewne mierzalne funkcje $f_1, f_2, \dots$ na [0,1], takie, że gdy ϑ ma rozkład $U(0,1)$, zmienne losowe $\vartheta_n = f_n(\vartheta)$ są niezależne o jednakowym rozkładzie $U(0,1)$.*[1]

Z lematu 12 możemy stworzyć ciąg niezależnych zmiennych losowych $\vartheta_1, \vartheta_2, ...$ Ich użyteczność ilustruje poniższy lemat.

Lemat 13 (Jądra i randomizacja)**.** *Niech μ będzie jądrem probabilistycznym z przestrzeni mierzalnej S na przestrzeń Borelowską T. Wtedy istnieje mierzalna funkcja $f : S \times [0,1] \to T$ taka, że jeżeli ϑ ma rozkład U(0,1), to $f(s, \vartheta)$ ma rozkład $\mu(s, \cdot)$ dla każdego $s \in S$.* [2]

Łatwo jest zauważyć, że mając elementy losowe $\xi, \tilde{\xi}, \eta, \tilde{\eta}$, takie, że zachodzi równość $(\xi, \eta) \stackrel{\mathrm{d}}{=} (\tilde{\xi}, \tilde{\eta})$, to zachodzi $\xi \stackrel{\mathrm{d}}{=} \tilde{\xi}$ oraz $\eta = \tilde{\eta}$. Jednak implikacja odwrotna nie musi być prawdziwa. Twierdzenie o transferze pozwoli nam na wprowadzenie warunków, tak, aby można było przeprowadzić to rozumowanie w drugą stronę.

Twierdzenie 6 (transfer)**.** *Dla pewnej przestrzeni Borelowskiej $(T, \mathcal{T})$ i przestrzeni mierzalnej $(S, \mathcal{S})$ ustalmy elementy losowe $\xi \stackrel{\mathrm{d}}{=} \tilde{\xi}$ oraz η odpowiednio w S i T. Wtedy istnieje element losowy $\tilde{\eta}$ w T, taki że $(\tilde{\xi}, \tilde{\eta}) \stackrel{\mathrm{d}}{=} (\xi, \eta)$. W szczególności, istnieje mierzalna funkcja $f : S \times [0,1] \to T$, taka, że możemy wziąć $\tilde{\eta} = f(\tilde{\xi}, \vartheta)$, gdy ma rozkład $U(0,1)$ oraz $\vartheta \perp\!\!\!\perp \tilde{\xi}$.*

[1]Dowód można znaleźć w [1](Lemma 3.21)
[2]Dowód można znaleźć w [1](Lemma 3.22)

Dowód. Z twierdzenia 4, istnieje jądro probabilistyczne μ z S do T, które spełnia

$$\mu(\xi, B) = P[\eta \in B|\xi] \text{ p.n.}, \ B \in \mathcal{T}.$$

Z lematu 13, możemy wybrać mierzalną funkcję $f : S \times [0,1] \to T$ taką, że $f(s, \vartheta)$ ma rozkład $\mu(s, \cdot)$ dla każdego $s \in S$. Zdefiniujmy $\tilde{\eta} = f(\tilde{\xi}, \vartheta)$. Dla dowolnej mierzalnej funkcji $g : S \times T \to \mathbb{R}_+$ dostajemy:

$$\begin{aligned} Eg(\tilde{\xi}, \tilde{\eta}) &= Eg(\tilde{\xi}, f(\tilde{\xi}, \vartheta)) \overset{(1)}{=} E \int g(\xi, f(\xi, u))du \\ &\overset{(2)}{=} E \int g(\xi, t)\mu(\xi, dt) \overset{(3)}{=} Eg(\xi, \eta), \end{aligned}$$

gdzie kolejne równości wynikają:

(1) z lematu 11 zastosowanego dla niezależnych elementów losowych $\tilde{\xi}$ oraz ϑ; funkcja $h : (x, y) \to g(x, f(x, y))$ jest mierzalna jako złożenie mierzalnej funkcji g z funkcją $F(x, y) = (F_1, F_2)$, gdzie $F_1 = id_x : (x, y) \to x$, $F_2 = f$. Jej mierzalność wynika z tego, że funkcja o wartościach w przestrzeni produktowej jest mierzalna wtedy i tylko wtedy, gdy mierzalna jest każda funkcja składowa.

(2) z lematu 3. Ustalmy ω, a tym samym $\xi(\omega)$. Połóżmy $f := f(\xi, \cdot)$ oraz $g := g(\xi, \cdot)$. Wiemy, że

$$\int g(\xi, f(\xi, u))du = \lambda(g \circ f) = (\lambda \circ f^{-1})g = \int g(\xi, t)\mu(\xi, dt),$$

gdzie ostatnia równość wynika z faktu, że $f \circ \vartheta$ ma rozkład $\mu(\xi, \cdot)$, czyli

$$\mu(\xi, \cdot) = P \circ (f \circ \vartheta)^{-1} = P \circ \vartheta^{-1} \circ f^{-1}.$$

Z faktu, że ϑ ma rozkład $U(0, 1)$ otrzymujemy

$$P \circ \vartheta^{-1} = \lambda,$$

gdzie λ jest miarą Lebesgue'a. Zatem:

$$\lambda \circ f^{-1}(dt) = \mu(\xi, dt).$$

(3) wprost z twierdzenia 5.

Z powyższego dostajemy $(\tilde{\xi}, \tilde{\eta}) \overset{d}{=} (\xi, \eta)$, gdyż $P\{(\tilde{\xi}, \tilde{\eta}) \in B \times C\} = E\mathbb{1}_{B\times C}(\tilde{\xi}, \tilde{\eta}) = E\mathbb{1}_{B\times C}(\xi, \eta) = P\{(\xi, \eta) \in B \times C\}$. □

Uwaga 8. Dla elementów losowych ξ oraz $\tilde{\xi}$ o wartościach w przestrzeni mierzalnej $(S, \mathcal{S})$ i dla funkcji mierzalnej f z S do przestrzeni mierzalnej $(T, \mathcal{T})$, równość $\xi \overset{d}{=} \tilde{\xi}$ pociąga za sobą równość $f(\xi) \overset{d}{=} f(\tilde{\xi})$.

Dowód. Istotnie, na mocy uwagi 7, wystarczy zauważyć, że gdy f jest mierzalna, zaś $g \in \hat{C}$, to $g \circ f \in \hat{C}$. □

Wniosek 2. *(równania stochastyczne) Ustalmy dwie przestrzenie borelowskie S i T, mierzalne odwzorowanie $f : T \to S$ oraz elementy losowe ξ w S i η w T, takie że $\xi \stackrel{\mathrm{d}}{=} f(\eta)$. Wtedy istnieje element losowy $\tilde{\eta} \stackrel{\mathrm{d}}{=} \eta$ w T, dla którego $\xi = f(\tilde{\eta})$ p.n.*

Dowód. Z twierdzenia 6, istnieje zmienna losowa $\tilde{\eta}$ w T taka, że $(\xi, \tilde{\eta}) \stackrel{\mathrm{d}}{=} (f(\eta), \eta)$. W szczególności $\tilde{\eta} \stackrel{\mathrm{d}}{=} \eta$ i $(\xi, f(\tilde{\eta})) \stackrel{\mathrm{d}}{=} (f(\eta), f(\eta))$, gdzie druga równość wynika z uwagi 8, a także z mierzalności $(s,t) \mapsto (s, f(t))$, uzasadnionej analogicznie jak w dowodzie twierdzenia 6. Dzięki mierzalności przekątnej S^2, dostajemy $P\{\xi = f(\tilde{\eta})\} = P\{f(\eta) = f(\eta)\} = 1$, a z tego $\xi = f(\tilde{\eta})$ p.n. □

Twierdzenie 7. *(coupling, Skorohod, Dudley) Rozważmy $\xi, \xi_1, \xi_2, \ldots$ będące elementami losowymi w przestrzeni metrycznej ośrodkowej (S, ρ) takie, że $\xi_n \xrightarrow{d} \xi$. Wtedy istnieje przestrzeń probabilistyczna z elementami losowymi $\eta \stackrel{\mathrm{d}}{=} \xi$ i $\eta_n \stackrel{\mathrm{d}}{=} \xi_n$, $n \in \mathbb{N}$ takimi, że $\eta_n \to \eta$ p.n.*[3]

Wniosek 3. *(Skorohod, coupling rozszerzony) Ustalmy mierzalne funkcje $f, f_1, f_2, \ldots$ z przestrzeni borelowskiej S do przestrzeni polskiej T oraz elementy losowe $\xi, \xi_1, \xi_2, \ldots$ w S takie, że $f_n(\xi_n) \xrightarrow{d} f(\xi)$. Wtedy istnieją elementy losowe $\tilde{\xi} \stackrel{\mathrm{d}}{=} \xi$ oraz $\tilde{\xi}_n \stackrel{\mathrm{d}}{=} \xi_n$ takie, że $f_n(\tilde{\xi}_n) \to f_n(\tilde{\xi})$ p.n.*

Dowód. Z twierdzenia 7, istnieją $\eta \stackrel{\mathrm{d}}{=} f(\xi)$ i $\eta_n \stackrel{\mathrm{d}}{=} f_n(\xi_n)$ takie, że $\eta_n \to \eta$ p.n. Co więcej, z wniosku 2 możemy położyć $\tilde{\xi} \stackrel{\mathrm{d}}{=} \xi$ i $\tilde{\xi}_n \stackrel{\mathrm{d}}{=} \xi_n$ takie, że p.n. $f(\tilde{\xi}) = \eta$ oraz $f_n(\tilde{\xi}_n) = \eta_n$ dla wszystkich n. Wtedy $f_n(\tilde{\xi}_n) \to f(\tilde{\xi}) = \eta$ p.n. □

Prawdopodobieństwo warunkowe pozwala nam na rozszerzenie definicji niezależności zdarzeń do jej warunkowej postaci. Uzasadnimy również kilka własności, które będą potrzebne do sformułowania twierdzenia o warunkowej niezależności i randomizacji.

Uwaga 9. Najmniejsze σ-ciało zawierające $\mathcal{F}$ i $\mathcal{G}$ oznaczamy jako $(\mathcal{F}, \mathcal{G})$ lub $\mathcal{F}, \mathcal{G}$.

Definicja 12 (warunkowa niezależność)**.** Mówimy, że σ-ciała $\mathcal{F}$ i $\mathcal{H}$ są niezależne pod warunkiem σ-ciała $\mathcal{G}$ (co zapisujemy jako $\mathcal{F} \perp\!\!\!\perp_{\mathcal{G}} \mathcal{H}$), jeśli

$$P^{\mathcal{G}}(A \cap B) = (P^{\mathcal{G}}A)(P^{\mathcal{G}}B) \text{ p.n.},$$

dla każdego zbioru $A \in \mathcal{F}$ i $B \in \mathcal{H}$.

Uwaga 10. Jeśli $\mathcal{F} \subset \mathcal{F}'$, to z definicji powyższej wynika

$$\mathcal{F}' \perp\!\!\!\perp_{\mathcal{G}} \mathcal{H} \Rightarrow \mathcal{F} \perp\!\!\!\perp_{\mathcal{G}} \mathcal{H}.$$

Definicja 13. Mówimy, że $P^{\mathcal{F}} = P^{\mathcal{G}}$ na $\mathcal{H}$, gdy dla każdego $H \in \mathcal{H}$ zachodzi

$$P^{\mathcal{F}}H = P^{\mathcal{G}}H \text{ p.n. .}$$

Twierdzenie 8 (Warunkowa niezależność, Doob)**.** *Dla dowolnych σ-ciał $\mathcal{F}$, $\mathcal{G}$ oraz $\mathcal{H}$ zachodzi równoważność*

$$\mathcal{F} \perp\!\!\!\perp_{\mathcal{G}} \mathcal{H} \Leftrightarrow P^{\mathcal{F},\mathcal{G}} = P^{\mathcal{G}} \textit{ p.n. na } \mathcal{H}^4.$$

[3]Dowód można znaleźć w [1](Theorem 4.30)
[4]Dowód można znaleźć w [1](Proposition 6.6)

Lemat 14. *Dla dowolnych σ-ciał $\mathcal{F}$, $\mathcal{G}$ oraz $\mathcal{H}$ zachodzi*

$$\mathcal{F} \perp\!\!\!\perp_{\mathcal{G}} \mathcal{H} \Longleftrightarrow \mathcal{F} \perp\!\!\!\perp_{\mathcal{G}} (\mathcal{G}, \mathcal{H}).$$

Dowód. Z twierdzenia 8, oba wyrażenia są równoważne

$$P^{\mathcal{G},\mathcal{H}} = P^{\mathcal{G}} \text{ p.n. na } \mathcal{F}.$$

□

Lemat 15 (reguła łańcuchowa)**.** *Dla każdego σ-ciała $\mathcal{G}$, $\mathcal{H}$ oraz $\mathcal{F}$ i $\mathcal{F}'$ zachodzi*

$$\mathcal{F}, \mathcal{F}' \perp\!\!\!\perp_{\mathcal{G}} \mathcal{H} \Rightarrow \mathcal{F}' \perp\!\!\!\perp_{\mathcal{F},\mathcal{G}} \mathcal{H}$$

Dowód. Teza wynika z twierdzenia 8 i z uwagi 10, iż $\mathcal{F} \perp\!\!\!\perp_{\mathcal{F},\mathcal{G}} \mathcal{H}$. □

Uwaga 11. Dla trywialnego σ-ciała $\mathcal{G} = \{\varnothing, \Omega\}$, mamy $\perp\!\!\!\perp_{\mathcal{G}} \equiv \perp\!\!\!\perp$, gdyż wówczas $P^{\mathcal{G}}A = PA$ dla dowolnego mierzalnego $A \subset \Omega$.

Wniosek 4. *Wobec uwagi 11, z lematu 15, zastosowanego dla $\mathcal{G} = \{\varnothing, \Omega\}$, wynika*

$$\mathcal{F}, \mathcal{F}' \perp\!\!\!\perp \mathcal{H} \Longrightarrow \mathcal{F}' \perp\!\!\!\perp_{\mathcal{F}} \mathcal{H}.$$

Twierdzenie 9 (warunkowa niezależność i randomizacja)**.** *Ustalmy odpowiednio elementy losowe ξ, η oraz ζ w przestrzeniach mierzalnych S, T oraz U, gdzie S jest borelowska. Wówczas $\xi \perp\!\!\!\perp_{\eta} \zeta$ zachodzi wtedy i tylko wtedy, gdy $\xi = f(\eta, \vartheta)$ p.n. dla pewnej funkcji mierzalnej $f : T \times [0,1] \to S$ i pewnej zmiennej losowej $\vartheta \perp\!\!\!\perp (\eta, \zeta)$ o rozkładzie $U(0,1)$.*

Dowód. ("$\Rightarrow$") Załóżmy, że $\xi = f(\eta, \vartheta)$ p.n., gdzie dzięki f jest funkcją mierzalną oraz $\vartheta \perp\!\!\!\perp (\eta, \zeta)$. Wtedy, z lematu 15, mamy $\vartheta \perp\!\!\!\perp_{\eta} \zeta$, więc korzystając z lematu 14 mamy $(\eta, \vartheta) \perp\!\!\!\perp_{\eta} \zeta$, a ponadto

$$(\eta, \vartheta)^{-1}(f^{-1}(\mathcal{S})) \subset (\eta, \vartheta)^{-1}(\mathcal{T} \otimes B[0,1]), \text{ skąd } \sigma(f(\eta, \vartheta)) \subset \sigma(\eta, \vartheta).$$

Zatem $\xi \perp\!\!\!\perp_{\eta} \zeta$.
("$\Leftarrow$") Przypuśćmy, że $\xi \perp\!\!\!\perp_{\eta} \zeta$ i ustalmy $\vartheta \perp\!\!\!\perp (\eta, \zeta)$ o rozkładzie $U(0,1)$. Z twierdzenia 6 istnieje pewna mierzalna funkcja $f : T \times [0,1] \to S$ taka, że element losowy $\tilde{\xi} = f(\eta, \vartheta)$ spełnia $(\tilde{\xi}, \eta) \stackrel{\mathrm{d}}{=} (\xi, \eta)$, więc:

$$P\{\xi \in B\} = P\{\xi \in B, \eta \in T\} = P\{(\xi, \eta) \in B \times T\} = P\{(\tilde{\xi}, \eta) \in B \times T\} = P\{\tilde{\xi} \in B\},$$

czyli $\tilde{\xi} \stackrel{\mathrm{d}}{=} \xi$. Korzystając z pierwszej części dowodu, mamy $\tilde{\xi} \perp\!\!\!\perp_{\eta} \zeta$. Stąd, na mocy twierdzenia 8:

$$P\left[\tilde{\xi} \in \cdot | \eta, \zeta\right] = P\left[\tilde{\xi} \in \cdot | \eta\right] = P\left[\xi \in \cdot | \eta\right] = P\left[\xi \in \cdot | \eta, \zeta\right] \text{ } p.n.,$$

gdzie środkowa równość wynika z definicji P^{η} oraz z faktu, że $P\{\tilde{\xi} \in B, \eta \in A\} = P\{\xi \in B, \eta \in A\}$, gdyż $(\tilde{\xi}, \eta) \stackrel{\mathrm{d}}{=} (\xi, \eta)$. Ponieważ $\xi, \tilde{\xi}$ mają wartości w przestrzeni borelowskiej, twierdzenie 4 zapewnia nam istnienie regularnej wersji $\mu(\eta, \cdot)$ rozkładu warunkowego $P[\tilde{\xi} \in \cdot | \eta, \xi] = P[\xi \in \cdot | \eta, \xi]$. Dalej, dla dowolnej mierzalnej i ograniczonej funkcji $f : S \times T \times U \to \mathbb{R}_+$, z twierdzenia 5:

$$Ef(\xi, (\eta, \zeta)) = E \int \mu((\eta, \zeta), ds) f(s, (\eta, \zeta)) = Ef(\tilde{\xi}, (\eta, \zeta)),$$

gdzie ostatnia równość wynika z poniższego:

$$\mu((\eta,\zeta),B) = E(\mathbb{1}\{\xi \in B\}|\eta,\zeta) = E(\mathbb{1}\{\tilde{\xi} \in B\}|\eta,\zeta) \text{ p.n.}$$

No, dobrze, zatem $(\tilde{\xi},\eta,\zeta) \stackrel{\mathrm{d}}{=} (\xi,\eta,\zeta)$ na mocy uwagi 7. Z twierdzenia 6, istnieje pewne $\tilde{\vartheta} \stackrel{\mathrm{d}}{=} \vartheta$, takie, że $(\xi,\eta,\zeta,\tilde{\vartheta}) \stackrel{\mathrm{d}}{=} (\tilde{\xi},\eta,\zeta,\vartheta)$. W szczególności, $\tilde{\vartheta} \perp\!\!\!\perp (\eta,\zeta)$ (bo z założenia $\vartheta \perp\!\!\!\perp (\eta,\zeta)$, czyli $\mathcal{L}(\eta,\zeta,\vartheta) = \mathcal{L}(\eta,\zeta) \otimes \mathcal{L}(\vartheta)$, ale $\mathcal{L}(\tilde{\vartheta}) = \mathcal{L}(\vartheta)$ oraz $\mathcal{L}(\eta,\zeta,\vartheta) = \mathcal{L}(\eta,\zeta,\tilde{\vartheta})$ oraz

$$P\{\tilde{\xi} \in A, f(\eta,\vartheta) \in B\} = P\{(\tilde{\xi},\eta,\vartheta) \in A \times f^{-1}(B)\} = P\{\xi \in A, f(\eta,\tilde{\vartheta}) \in B\}$$

czyli, z lematu 1 dla π-układu $\mathcal{C} = \{A \times B : \ A, B \in \mathcal{S}\}$, $(\xi, f(\eta,\tilde{\vartheta})) \stackrel{\mathrm{d}}{=} (\tilde{\xi}, f(\eta,\vartheta))$. Ponieważ $\tilde{\xi} = f(\eta,\vartheta)$, to $(\xi, f(\eta,\tilde{\vartheta})) \stackrel{\mathrm{d}}{=} (\tilde{\xi}, f(\eta,\vartheta)) = (f(\eta,\vartheta), f(\eta,\vartheta))$ oraz z mierzalności przekątnej w S^2 otrzymujemy:

$$P\{\xi = f(\eta,\tilde{\vartheta})\} = P\{f(\eta,\vartheta) = f(\eta,\vartheta)\} = 1,$$

czyli $\xi = f(\eta,\tilde{\vartheta})$ p.n. $\square$

Literatura

[1] Olav Kallenberg: *Foundations of Modern Probability.* Springer, New York, 2002.