Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Dana Křepinská

Radon-Nikodymova derivace v pravděpodobnosti

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	Mgr. Petr Dostál, Ph.D.
Studijní program:	Matematika
Studijní obor:	Obecná matematika

Praha 2012

Chtěla bych velmi poděkovat vedoucímu této bakalářské práce, Mgr. Petru Dostálovi, Ph.D., za čas a ochotu, s jakou se mi věnoval, a také za materiály a rady, bez kterých by tato práce nebyla nikdy dokončena.

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne .... Podpis autora

Název práce: Radon-Nikodymova derivace v pravděpodobnosti

Autor: Dana Křepinská

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Mgr. Petr Dostál, Ph.D.

Abstrakt: Tato bakalářská práce se zabývá Radon-Nikodymovou derivací, jejími vlastnostmi, souvislostí s derivací míry a následně jejím využitím v teorii pravděpodobnosti. Je zde podrobně definované podmíněné pravděpodobnostní rozdělení a vyřešena otázka jednoznačnosti při spojitém podmiňování jevem nulové pravděpodobnosti. Dále je v textu pomocí podmíněného rozdělení definována podmíněná střední hodnota a dokázané některé její vlastnosti. Práce se také zmiňuje o borelovsky izomorfních prostorech a okrajově o podmíněném rozptylu a kovarianci. Závěr práce je věnován aplikaci podmiňování při konstrukci Brownova mostu z Wienerova procesu a následnému využití Brownova mostu ve statistice.

Klíčová slova: Radon-Nikodymova derivace, podmíněné pravděpodobnostní rozdělení, Brownův most

Title: Radon-Nikodym Derivate in Probability Theory

Author: Dana Křepinská

Department: Department od Probability and Mathematical Statictics

Supervisor: Mgr. Petr Dostál, Ph.D.

Abstract: This thesis concerns the Radon-Nikodym derivate, its properties, connection with measure derivative and its applications in the probability theory. The text defines the conditional probability distribution and solves the problem of uniqueness in the case of conditioning of an event which has zero probability of occuring. Next part of the text is about the conditional expactation, which is defined by the conditional distribution, and some of its properties. There are also few words about the Borel isomorphic spaces and the conditional variability and covariance. Last section of this work is about construction of the Brownian Bridge from the Wiener process and about its applications is the statistics.

Keywords: the Radon-Nikodym derivate, the conditional probability distribution, the Brownian bridge

Obsah

Úvod

Tato bakalářská práce se věnuje pojmu Radon-Nikodymova derivace. V první kapitole je seznámení s klíčovou Radon-Nikodymovou větou a souvislost Radon-Nikodymovy derivace s pojmem derivace míry.

V další části práce je vylíčena aplikace v teorii pravděpodobnosti. Je zde shrnuto zavedení podmíněného rozdělení a také rozebrána problematika spojitého podmiňování jevem nulové pravděpodobnosti. Pomocí podmíněného rozdělení je také zavedena podmíněná střední hodnota a jsou zde dokázány její základní vlastnosti. Tato kapitola také přibližuje problematiku borelovsky izomorfních prostorů, na kterých máme zajištěnou existenci podmíněného rozdělení. Okrajově se text zmiňuje o podmíněném rozptylu a podmíněné kovarianci.

Poslední kapitola práce je věnována Wienerově procesu a Brownově mostu, který lze právě pomocí podmíněného rozdělení zkonstruovat z Wienerova procesu.

V celé práci se předpokládají základní znalosti z teorie míry, pravděpodobnosti a statistiky.

Kapitola 1

Radon-Nikodymova derivace

1.1 Radon-Nikodymova věta

V této kapitole je uvedena Radon-Nikodymova věta včetně důkazu a některých jejích vlastností. V celé této části se budeme pohybovat na měřitelném prostoru $(X, \mathcal{A})$.

Definice 1.1

Nechť μ je σ-konečná míra na $\mathcal{A}$, λ je komplexní míra na $\mathcal{A}$. Nechť platí

$$\lambda = \lambda_a + \lambda_s, \quad \lambda_a \ll \mu, \quad \lambda_s \perp \mu,$$

Pak míry λ_a a λ_s nazýváme *Lebesgueovým rozkladem* míry λ.

Poznámka

Pro komplexní funkci $f = f_1 + if_2$ zavedeme značení $f \in L^1(\mu)^2$, pokud pro reálné funkce platí $f_1, f_2 \in L^1(\mu)$.

Definice 1.2

Nechť μ je σ-konečná míra na $\mathcal{A}$, λ je komplexní nebo znaménková míra na $\mathcal{A}$. Funkci $h \in L^1(\mu)^2$ nazveme *Radon-Nikodymovou derivací* míry λ vzhledem k míře μ (značíme také $\frac{d\lambda}{d\mu}$), pokud splňuje, že

$$\lambda(E) = \int_E h \, d\mu, \quad E \in \mathcal{A}.$$

Lemma 1.1

Nechť μ je konečná míra na $(X, \mathcal{A})$, $f \in L^1(\mu)$ a nechť $-\infty < a \leq b < \infty$. Nechť platí

$$A_E(f) := \frac{1}{\mu(E)} \int_E f \, d\mu \in [a, b], \quad E \in \mathcal{A}, \quad \mu(E) > 0.$$

Potom $f(x) \in [a, b]$ pro skoro všechna $x \in X$.

Důkaz

Označíme $G := \mathbb{R} \backslash [a, b]$, pak G je otevřená. Zvolme $r > 0$ a $z_0 \in \mathbb{R}$ a definujme množinu B jako uzavřené okolí bodu z_0 o poloměru r, tedy

$$B := \{z \in \mathbb{R}, |z - z_0| \leq r\}.$$

Nyní stačí dokázat, že pokud $B \subset G$, pak $\mu(f^{-1}(B)) = 0$. Označíme $E = f^{-1}(B)$ a budeme dokazovat sporem. Nechť tedy $B \subset G$ a $\mu(E) > 0$. Pak

$$\begin{aligned}|A_E(f) - z_0| &= \left|\frac{1}{\mu(E)}\int_E f\ d\mu - \frac{1}{\mu(E)}\int_E z_0\ d\mu\right| = \left|\frac{1}{\mu(E)}\int_E (f - z_0)\ d\mu\right| \leq \\ &\leq \frac{1}{\mu(E)}\int_E |f - z_0|\ d\mu \leq r,\end{aligned}$$

neboli $A_E(f) \in B \subset \mathbb{R} \setminus [a, b]$. Což se spor s předpokladem $A_E(f) \in [a, b]$.

□

Tvrzení 1.2

Lebesgueův rozklad $\lambda = \lambda_a + \lambda_s$ a Radon-Nikodymova derivace h jsou pro dané míry μ a λ určeny jednoznačně.

Důkaz

Nejprve ověříme jednoznačnost Lebesgueova rozkladu. Využijeme zde jednoznačnosti Radon-Nikodymovy derivace h, kterou pak ověříme ve druhé části důkazu. Nechť $\lambda = \lambda'_a + \lambda'_s$ je jiný Lebesgueův rozklad, nechť B, B' jsou množiny takové, že

$$\mu(B) = \mu(B') = 0,$$

$$\lambda_s(E) = \lambda(E \cap B), \quad \lambda'_s(E) = \lambda(E \cap B'), \quad E \in \mathcal{A}.$$

Potom pro $C = B \cup B'$ platí, že $\mu(C) = 0$. Z předpokladu máme, že

$$\lambda_s(E) + \int_E h\ d\mu = \lambda(E) = \lambda'_s(E) + \int_E h\ d\mu, \quad E \in \mathcal{A}.$$

Tedy $\lambda_s(E) = \lambda'_s(E)$ pro každé $E \in \mathcal{A}$, a tudíž $\lambda_a = \lambda'_a$.

Nyní dokážeme jednoznačnost Radon-Nikodymovy derivace. Nechť h a h' jsou dvě Radon-Nikodymovy derivace $\frac{d\lambda}{d\mu}$. Definujeme funkci $g := h - h'$. Podle Lemmatu 1.1 pak platí

$$A_E(g) = 0 \quad \Rightarrow \quad g = 0\ s.v., \quad E \in \mathcal{A}, \quad \mu(E) > 0.$$

$$\begin{aligned}A_E(g) &= \frac{1}{\mu(E)}\int_E (h - h')\ d\mu = \frac{1}{\mu(E)}\int_E h\ d\mu\ - \ \frac{1}{\mu(E)}\int_E h'\ d\mu = \\ &= \frac{\lambda(E)}{\mu(E)}\ - \ \frac{\lambda(E)}{\mu(E)} = 0.\end{aligned}$$

□

Lemma 1.3

Nechť μ je σ-konečná míra. Pak existuje $w \in L^1(\mu)$, $0 < w < 1$.

Věta 1.4 (Rieszova věta o reprezentaci)

Nechť F je spojitý lineární funkcionál na Hilbertově prostoru H. Potom existuje právě jeden prvek $y \in H$ takový, že $F(x) = (x, y)$, $x \in H$.

Poznámka

Důkazy Lemmatu 1.3 a Věty 1.4 v této práci vynecháme, lze je nalézt v literatuře, např. v Rudinově knize [6].

Věta 1.5 (Radon-Nikodymova věta, Lebesgueův rozklad)

Nechť μ je σ-konečná míra na $\mathcal{A}$, λ je konečná míra na $\mathcal{A}$. Potom existuje Lebesgueův rozklad míry $\lambda = \lambda_a + \lambda_s$ a Radon-Nikodymova derivace

$$h = \frac{d\lambda_a}{d\mu}.$$

Důkaz

Nechť w je funkce z Lemmatu 1.3, definujeme pomocnou míru $\psi := \lambda + w\mu$, tedy

$$\psi(E) = \lambda(E) + \int_E w \, d\mu, \quad E \in \mathcal{A}.$$

Pak ψ je konečná míra. Pro $f \geq 0$ měřitelnou platí:

$$\int_X f \, d\psi = \int_X f \, d\lambda + \int_X fw \, d\mu.$$

Nechť $f \in L^2(\psi)$, označíme $c := \sqrt{\psi(X)} < \infty$, pak z Hölderovy nerozvnosti dostaneme odhad:

$$\left| \int_X f \, d\lambda \right| \leq \int_X |f| \, d\lambda \leq \int_X |f| 1 \, d\psi \leq \left(\int_X |f|^2 \, d\psi \right)^{\frac{1}{2}} \left(\int_X 1^2 \, d\psi \right)^{\frac{1}{2}} = c\|f\|_2,$$

zaručující spojitost lineárního funkcionálu

$$f \mapsto \int_X f \, d\lambda.$$

Z Rieszovy věty o reprezentaci víme, že existuje $g \in L^2(\psi)$ takový, že

$$\int_X f \, d\lambda = \int_X fg \, d\psi = \int_X fg \, d\lambda + \int_X fgw \, d\mu, \tag{1}$$

speciálně pro $f = 1_E$, $E \in \mathcal{A}$, $\psi(E) > 0$ platí:

$$\lambda(E) = \int_E g \, d\psi \quad \Rightarrow \quad 0 \leq \frac{1}{\psi(E)} \int_E g \, d\psi = \frac{\lambda(E)}{\psi(E)} \leq 1.$$

Podle Lemmatu 1.1 platí, že $0 \leq g(x) \leq 1$ pro s.v. $x \in X$, BÚNO předpokládejme, že $0 \leq g(x) \leq 1$ všude.

Z (1) plyne, že

$$\int_X (1-g) f \, d\lambda = \int_X fgw \, d\mu, \quad f \in L^2(\psi). \tag{2}$$

Definujeme $A := \{g < 1\}$, $S := \{g = 1\}$, potom $A, S \in \mathcal{A}$, $A \cap S = \emptyset$, definujeme

$$\lambda_a(E) := \lambda(A \cap E), \quad \lambda_s(E) := \lambda(S \cap E), \quad E \in \mathcal{A},$$

a dosadíme do (2) $f = 1_S$, pak

$$0 = \int_S 1w\, d\mu \quad \Rightarrow \quad \mu(S) = 0 \quad \Rightarrow \quad \lambda_s \perp \mu.$$

Dosadíme do (2) $f = (1 + g + g^2 + \ldots + g^n)1_E$, $n \in \mathbb{N}$, $E \in \mathcal{A}$, pak

$$\int_E (1 - g^{n+1})\, d\lambda = \int_E g(1 + g + g^2 + \ldots + g^n)w\, d\mu. \tag{3}$$

Je-li $x \in S$, pak levá strana (3) je rovna 0. Navíc pro $x \in A$ platí $g^n(x) \searrow 0$, tedy limita levé strany (3) je rovna $\lambda(E \cap A)$. Definujeme

$$h := g \sum_{n=0}^{\infty} g^n w 1_A,$$

použijeme Lebesgueovu větu.

$$\lim_{n\to\infty} \int_E g(1 + g + g^2 + \ldots + g^n)w d\mu = \int_E h\, d\mu,$$

$$\lambda_a(E) = \int_E h\, d\mu, \quad E \in \mathcal{A},$$

navíc $\lambda_a(X) = \int_X h\, d\mu < \infty$, tedy $h \in L^1(\mu)$.

$\square$

Poznámka

Obecněji lze důkaz Radon-Nikodymovy věty rozšířit pro λ komplexní, pak můžeme rozepsat $\lambda = \lambda_r + i\lambda_i$, kde λ_r a λ_i jsou reálné míry, které můžeme rozložit na kladné a záporné části. Pro ty už platí výše uvedená věta.

Pro znaménkovou míru λ provedeme analogické kroky, zde ale využijeme Hahnův rozklad $\lambda = \lambda_1 - \lambda_2$.

1.2 Derivace míry

V následující sekci se podíváme na derivaci míry a ukážeme si její souvislost s Radon-Nikodymovou derivací. V celé této kapitole se pohybujeme na prostoru $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ a symbolem λ budeme označovat Lebesgueovu míru na tomto prostoru. Také budeme užívat označení $B(r, x)$ pro otevřenou kouli v $\mathbb{R}^k$ o poloměru r, tedy $B(r, x) = \{y \in \mathbb{R}^k; |y - x| < r\}$.

Definice 1.3

Nechť μ je komplexní borelovská míra v $\mathbb{R}^k$, pak *derivaci míry* definujeme předpisem

$$(D\mu)(x) = \lim_{r\to 0} \frac{\mu(B(x, r))}{\lambda(B(x, r))},$$

pokud tato limita existuje.

Poznámka

Pro pozdější účely si definujeme tzv. maximální funkci $M\mu$. Pro nezápornou míru ji definujeme předpisem

$$(M\mu)(x) = \sup_{r>0} \left(\frac{\mu(B(x,r))}{\lambda(B(x,r))} \right),$$

pro komplexní míru je maximální funkce rovna maximální funkci variace míry $|\mu|$.

Pro $f \in L^1(\mathbb{R}^k)$ definujeme maximální funkci $Mf : \mathbb{R}^k \to [0,\infty]$ jako

$$(Mf)(x) = \sup_{r>0} \left(\frac{1}{\lambda(B(x,r))} \int_{B(x,r)} |f| \; d\lambda \right).$$

Lemma 1.6

Nechť $\{B(x_i,r_i)\}$ je konečný systém otevřených koulí v $\mathbb{R}^k$, $i = 1,\dots,n$, nechť $W = \bigcup_{i=1}^N B(x_i,r_i)$. Potom existuje $S \subset \{1,\dots,N\}$ taková, že platí

a) $B(x_i,r_i) \cap B(x_j,r_j) = \emptyset$, pro $i \neq j$, $i,j \in S$,

b) $W \subset \bigcup_{i\in S} B(x_i,3r_i)$,

c) $\lambda(W) \leq 3^k \sum_{i\in S} \lambda(B(x_i,r_i))$.

Důkaz

Seřadíme koule $B_i := B(x_i,r_i)$ tak, aby $r_1 \geq r_2 \geq \ldots \geq r_N$, položíme $i_1 = 1$. Vyřadíme všechny koule B_j, pro které $B_j \cap B_i \neq \emptyset$.

Nechť B_{i_2} je první koule ze zbývajících, pokud takové existují. Vyřadíme všechny koule B_j, pro které $j > i_2$ a $B_j \cap B_{i_2} \neq \emptyset$. Označíme první kouli ze zbývajících B_{i_3} a opakujeme dále.

Protože postup se zastaví po konečně mnoha krocích, získáme konečnou množinu $S = \{i_1, i_2, \ldots\}$. Pak a) platí z postupu, jakým jsme vytvořili S. Pro každou vyřazenou kouli platí $B_j \subset B(x_i,3r_i)$ pro některé $i \in S$, neboť víme, že

$$r' \leq r, \quad B(x',r') \cap B(x,r) \neq \emptyset \quad \Rightarrow \quad B(x',r') \subset B(x,3r),$$

tedy platí b). Navíc $\lambda(B(x,3r)) = 3^k \lambda(B(x,r))$ v $\mathbb{R}^k$, tedy c) plyne z b).

□

Věta 1.7

Nechť μ je komplexní borelovská míra v $\mathbb{R}^k$, $\alpha > 0$, potom platí, že

$$\lambda(\{x \in \mathbb{R}^k; (M\mu)(x) > \alpha\}) = \lambda(\{M\mu > \alpha\}) \leq 3^k \alpha^{-1} \|\mu\|,$$

kde $\|\mu\| = |\mu|(\mathbb{R}^k)$.

Důkaz

Nechť $K \subset \{M\mu > \alpha\}$ je kompaktní. Každý bod $x \in K$ je středem otevřené koule B, pro niž $|\mu|(B) > \alpha\lambda(B)$.

Protože K je kompaktní množina, pak existuje konečný systém těchto koulí, který K pokrývá. Z předchozího Lemmatu 1.6 máme, že pro disjunktní podsystém $\{B_1, B_2, \dots, B_n\}$ platí:

$$\lambda(K) \leq 3^k \sum_{i=1}^{n} \lambda(B_i) \leq 3^k \alpha^{-1} \sum_{i=1}^{n} |\mu|(B_i) \leq 3^k \alpha^{-1} \|\mu\|.$$

Nyní přejdeme k supremu přes všechny K kompaktní množiny, tedy dostáváme, že

$$\lambda(\{M\mu > \alpha\}) \leq 3^k \alpha^{-1} \|\mu\|.$$

□

Poznámka

Je-li $f \in L^1(\mathbb{R}^k)$, $\alpha > 0$ a $E = \{|f| > \alpha\}$ pak platí:

$$\lambda(\{|f| > \alpha\}) \leq \alpha^{-1} \|f\|_1.$$

Definice 1.4

Nechť $f \in L^1(\mathbb{R}^k)$, $x \in \mathbb{R}^k$, pak x nazveme *Lebesgueovým bodem* funkce f, pokud platí:

$$\lim_{r \to 0} \frac{1}{\lambda(B_r)} \int_{B(x,r)} |f(y) - f(x)| \, d\lambda(y) = 0.$$

Poznámka

(i) $\lambda(B(x,r))$ závisí pouze na r, neboť Lebesgueova míra je invariantní vůči posunutí, tedy můžeme psát pouze $\lambda(B_r)$.

(ii) Z definice Lebesgueova bodu vidíme, že každý bod spojitosti f je zároveň Lebesgueovým bodem. Pokud se zamyslíme, nad interpretací vztahu v definici, jsou to takové body, ve kterých funkce neosciluje příliš mnoho z hlediska průměrů.

Věta 1.8

Nechť $f \in L^1(\mathbb{R}^k)$, potom skoro každý bod $x \in \mathbb{R}^k$ je Lebesgueův bod f.

Důkaz

Nechť $x \in \mathbb{R}^k$ a $r > 0$. Definujeme

$$(T_r f)(x) = \frac{1}{\lambda(B_r)} \int_{B(x,r)} |f - f(x)| \, d\lambda,$$
$$(Tf)(x) = \limsup_{r \to 0} (T_r f)(x).$$

Chceme dokázat, že $Tf = 0$ skoro všude. Nechť $y > 0$, pak pro $n \in \mathbb{N}$ existuje funkce $g \in \mathcal{C}(\mathbb{R}^k)$ taková, že $\|f - g\|_1 < \frac{1}{n}$. Položme $h = f - g$.

Funkce g je spojitá, tedy $Tg = 0$, a protože platí:

$$(T_r h)(x) \leq \frac{1}{\lambda(B_r)} \int_{B(x,r)} |h| \, d\lambda + |h(x)|,$$

pak také $Th \leq Mh + |h|$. Protože $T_r f \leq T_r g + T_r h$, dostáváme, že $Tf \leq Mh + |h|$, tedy platí:

$$\{Tf > 2y\} \subset \underbrace{\{Mh > y\} \cup \{|h| > y\}}_{E(y,n)}.$$

Protože $\|h\|_1 < \frac{1}{n}$, pak z předchozí Věty 1.7 a Poznámky za ní plyne, že

$$\lambda(E(y,n)) \leq \frac{3^k + 1}{yn}.$$

Levá strana nezávisí na n, tedy

$$\{Tf > 2y\} \subset \bigcap_{n=1}^{\infty} E(y,n),$$

navíc platí, že $\lambda(\bigcap_{n=1}^{\infty} E(y,n)) = 0$, tedy existuje měřitelná množina N, pro kterou $\lambda(N) = 0$. Ale Lebesgueova míra je úplná, tedy $\lambda(\{Tf > 2y\}) = 0$ pro každé $y > 0$, takže $Tf = 0$ skoro všude.

$\square$

Věta 1.9

Nechť μ je komplexní borelovská míra na $\mathbb{R}^k$ a $\mu \ll \lambda$. Nechť f je Radon-Nikodymova derivace $\frac{d\mu}{d\lambda}$. Potom derivace míry $D\mu = f$ skoro všude a platí:

$$\mu(E) = \int_E (D\mu)\ d\lambda, \quad E \in \mathcal{B}(\mathbb{R}^k).$$

Důkaz

Nechť x je Lebesgueův bod funkce f, pak platí:

$$f(x) = \lim_{r \to 0} \frac{1}{\lambda(B_r)} \int_{B(x,r)} f\ d\lambda = \lim_{r \to 0} \frac{\mu(B(x,r))}{\lambda(B(x,r))},$$

tedy $(D\mu)(x)$ existuje a je rovna $f(x)$. Toto platí dle Věty 1.8 skoro všude.

$\square$

1.3 Výpočet Radon-Nikodymovy derivace

Nyní si formulujeme a dokážeme větu, která nám dává alternativní postup pro výpočet Radon-Nikodymovy derivace.

Věta 1.10

Nechť F je zprava spojitá, neklesající funkce na $\mathbb{R}$, nechť μ_F je příslušná Lebesgue-Stieltjesova míra. Nechť $h \in L^1(\mu_F)$, pak pro μ_F-skoro všechna $x \in \mathbb{R}$ platí, že

$$\lim_{y \to x-} \frac{G(x) - G(y)}{F(x) - F(y)} = h(x), \tag{4}$$

kde $G(x) = \int_{-\infty}^{x} h dF = \int_{(-\infty,x]} h d\mu_F$.

Důkaz

Pokud $\mu_F \equiv 0$, pak věta platí triviálně. Nechť $\mu_F \not\equiv 0$, předpokládejme nejprve, že F je distribuční funkce, nechť U je náhodná veličina s rovnoměrným rozdělením na $[0,1]$, pak náhodná veličina $Y = F^{-1}(U)$ má distribuční funkci F, kde

$$F^{-1}(u) = \inf\{x \in \mathbb{R}; F(x) \geq u\}, \quad u \in (0,1),$$

Pak také

$$G(x) = \int_{-\infty}^{x} h\, dF = E\left[h(Y); Y \leq x\right] = E\left[h(F^{-1}(U)); U \leq F(x)\right] = \\ = \int_0^{F(x)} h(F^{-1}(u))\, du.$$

Pokud $F(x) > F(x-)$, pak

$$\lim_{y \to x-} \frac{G(x) - G(y)}{F(x) - F(y)} = \frac{G(x) - G(x-)}{F(x) - F(x-)} = h(x).$$

Pokud $F(x) = F(x-)$, definujeme absolutně spojitou funkci H předpisem

$$H(u) = \int_0^u h(F^{-1}(v))\, dv,$$

pro kterou $H'(u) = h(F^{-1}(u))$ pro skoro všechna $u \in (0,1)$. Chceme dokázat, že (4) platí pro $x = F(u)$, pro skoro všechna $u \in (0,1)$. Vidíme, že množina všech $u \in (0,1)$, pro které $card(F^{-1}\{u\}) > 1$, je spočetná, neboť každý takový $F^{-1}\{u\}$ obsahuje různé racionální číslo. Navíc platí, že pokud je $F^{-1}\{u\} = \emptyset$, pak $F(x) > F(x-)$ pro $x = F^{-1}(u)$, což je vyřešeno výše. Předpokládejme, že $x = F^{-1}(u) \in \mathbb{R}$ takové, že $card(F^{-1}\{u\}) = 1$, $F(x) = F(x-) = u$ a navíc $H'(u) = h(F^{-1}(u))$, pak

$$\frac{G(x) - G(y)}{F(x) - F(y)} = \frac{H(F(x)) - H(F(y))}{F(x) - F(y)} \xrightarrow{y \to x-} H'(F(x)) = h(F^{-1}(F(x))) = \\ = h(F^{-1}(u)) = h(x).$$

Nyní rozebereme případ, kdy F je neklesající, zprava spojitá a platí

$$-\infty < F(-\infty) < F(\infty) < \infty.$$

Definujeme funkci

$$\tilde{F}(x) := \frac{F(x) - F(-\infty)}{F(\infty) - F(-\infty)}.$$

Funkce $\tilde{F}$ je zprava spojitá a neklesající, navíc je to distribuční funkce. Pro Lebesgue-Stieltjesovu míru $\mu_{\tilde{F}}$ pak platí

$$\mu_{\tilde{F}}((a,b]) = \tilde{F}(b) - \tilde{F}(a) = \frac{F(b) - F(a)}{F(\infty) - F(-\infty)}.$$

Nechť $h \in L^1(\mu_F)$, pak

$$\int |h|\, d\mu_{\tilde{F}} = \frac{\int |h|\, d\mu_F}{\mu_F(\mathbb{R})} < \infty \quad \Rightarrow \quad h \in L^1(\mu_{\tilde{F}}).$$

Potom pro skoro všechna $x \in \mathbb{R}$ platí

$$h(x) = \lim_{y \to x-} \frac{\int_{(y,x]} h \, d\tilde{F}}{\tilde{F}(x) - \tilde{F}(y)} = \lim_{y \to x-} \frac{\int_{(y,x]} h \, dF}{F(x) - F(y)} = \lim_{y \to x-} \frac{G(x) - G(y)}{F(x) - F(y)}.$$

Zbývá už jen situace, kdy F je neklesající, zprava spojitá a $F(\infty) - F(-\infty) = \infty$. Definujeme funkce

$$F_n(x) = F((-n) \vee x \wedge n), \quad n \in \mathbb{N}.$$

Protože $\mu_{F_n} \leq \mu_F$, pak také pro $h \in L^1(\mu_F)$ platí

$$\int |h| \, dF_n \leq \int |h| \, dF \quad \Rightarrow \quad h \in L^1(\mu_{F_n}).$$

Nechť $x \in \mathbb{R}$, pak existuje $n \in \mathbb{N}$, že $x \in (-n, n)$, navíc (4) platí pro F_n. Tedy (4) platí pro F.

$\square$

Kapitola 2

Podmiňování

2.1 Podmíněné rozdělení

V této kapitole se budeme zabývat podmiňováním náhodných veličin. Než k tomu ale přikročíme, probereme si nejprve základy podmiňování jevů.

Na prostoru $(\Omega, \mathcal{A}, P)$ máme dva náhodné jevy A a B, přičemž platí, že $P(B) > 0$. Pak podmíněnou pravděpodobnost jevu A za podmínky, že nastal jev B definujeme jako

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0.$$

Pokud jevy A a B jsou nezávislé, pak víme, že $P(A \cap B) = P(A)P(B)$, tedy

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A).$$

Tímto vztahem lze také definovat nezávislost dvou náhodných jevů s nenulovou pravděpodobností.

Stejné úvahy lze aplikovat rovněž pro práci s náhodnými veličinami, ovšem stále jen v případě, že pravděpodobnost podmínky je kladná. Nechť náhodná veličina $X : (\Omega, \mathcal{A}) \to (S, \mathcal{B}(S))$. Předpokládejme nejprve o náhodné veličině Y, že má diskrétní rozdělení. Pak definujeme podmíněné rozdělení následovně

$$P(X \in A|Y = y) = \frac{P(x \in A, Y = y)}{P(y = y)}, \quad P(Y = y) > 0, \quad A \in \mathcal{B}(S).$$

Pokud ale má veličina Y absolutně spojité rozdělení, pak víme, že jednoho bodu nabývá s nulovou pravděpodobností, tedy nelze použít stejné úvahy. V následující části textu definujeme obecný tvar podmíněného rozdělení a rozebereme otázku jednoznačnosti v případě spojitého podmiňování. Pro jednoduchost se nejprve zaměříme na reálnou náhodnou veličinu, později pak výsledky zobecníme.

Definice 2.1

Nechť $X : (\Omega, \mathcal{A}) \to (S, \mathcal{S})$, $Y : (\Omega, \mathcal{A}) \to (H, \mathcal{H})$ jsou náhodné veličiny. *Podmíněným rozdělením* X při $Y = y$ rozumíme funkci

$$P_{X|Y=y} : \mathcal{S} \times H \to [0, 1], \quad (B, y) \mapsto P_{X|Y=y}(B), \quad B \in \mathcal{S}, \quad y \in H,$$

která splňuje:

- pro každé $y \in H$ je funkce $B \mapsto P_{X|Y=y}(B)$ pravděpodobnostní míra v $\mathcal{S}$,
- pro každou $B \in \mathcal{S}$ je funkce $y \mapsto P_{X|Y=y}(B)$ $\mathcal{H}$-měřitelná v proměnné y,
- pokud $B \in \mathcal{S}$, $C \in \mathcal{H}$, pak

$$P(X \in B, Y \in C) = \int_C P_{X|Y=y}(B)\ dP_Y(y).$$

Poznámka

Symbolem $\mathbb{L}(\mathcal{H})$ označíme množinu všech $\mathcal{H}$-měřitelných reálných náhodných veličin, kdykoliv $\mathcal{H}$ je σ-algebra.

Symbolem $\mathbb{L}(\mathcal{H}, \mathcal{S})$ budeme dále značit množinu všech náhodných veličin $X : (H, \mathcal{H}) \to (S, \mathcal{S})$, kde $H = \bigcup \mathcal{H}$ a $S = \bigcup \mathcal{S}$, kdykoliv $\mathcal{H}, \mathcal{S}$ jsou σ-algebry.

Poznámka

Je-li F distribuční funkce, pak $G = F|_{\mathbb{Q}}$ je neklesající, zprava spojitá a platí

$$\lim_{x\to\infty} F(x) = 1, \quad \lim_{x\to-\infty} = 0.$$

Naopak pokud funkce $G : \mathbb{Q} \to \mathbb{R}$ splňuje výše uvedené vlastnosti, pak předpis

$$F(x) = \inf\{G(w); x \leq w, w \in \mathbb{Q}\}$$

definuje distribuční funkci rozšiřující funkci G.

Poznámka

Nechť $X \in \mathbb{L}(\mathcal{A})$, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ jsou náhodné veličiny. Definujeme míry

$$\mu(B, C) = P(X \in B, Y \in C), \quad B \in \mathcal{B}(\mathbb{R}), \quad C \in \mathcal{H},$$

$$\mu_x := \mu((-\infty, x], \cdot), \quad x \in \mathbb{Q}.$$

Definujeme funkci F_y jako Radon-Nikodymovu derivaci

$$F_y(x) = \frac{d\mu_x}{dP_Y}(y), \quad y \in H, \quad x \in \mathbb{Q}.$$

Platí, že $P_Y(\{y \in H; \exists F \text{ distribuční funkce, že } F(x) = \frac{d\mu_x}{dP_Y}(y), x \in \mathbb{Q}\}) = 1$. Ukážeme, že funkce F_y je neklesající, zprava spojitá a pro její limity platí

$$\lim_{x\to\infty} F(x) = 1, \quad \lim_{x\to-\infty} = 0.$$

Tyto poznatky využijeme v následujícím důkazu existence podmíněného rozdělení.

Vidíme, že F_y je neklesající skoro jistě, neboť pro $x, z \in \mathbb{Q}$, $x < z$, máme

$$\mu_x \leq \mu_z \quad \Rightarrow \quad F_y(x) \leq F_y(z).$$

Nyní ukážeme, že F_y je spojitá zprava skoro jistě. Platí skoro jistě

$$0 \leq \mu_x \leq P_Y \quad \Rightarrow \quad F_y(x) \in [0, 1].$$

Nechť $\{x_n\}$ je monotónní posloupnost prvků z $\mathbb{Q}$, pak také $F_y(x_n)$ je monotónní a $F_y(x_n) \in [0,1]$ skoro jistě. Tedy posloupnost $\{F_y(x_n)\}$ je konvergentní.

$$\forall C \in \mathcal{H}: \quad \lim_{n\to\infty} P(X \le x_n, Y \in C) = \lim_{n\to\infty} \int_C F_y(x_n)\ dP_Y(y) =$$
$$= \int_C \lim_{n\to\infty} F_y(x_n)\ dP_Y(y).$$

Tedy z definice Radon-Nikodymovy derivace a F_y máme, že skoro jistě platí

$$\lim_{n\to\infty} F_y(x_n) = F_y(\lim_{n\to\infty} x_n),$$

takže pro $\{x_n\}$ nerostoucí získáváme spojitost zprava.

Nechť $x_n \to \infty$, $x_n \in \mathbb{Q}$, pak máme, že

$$P(X \le x_n, Y \in C) \xrightarrow{n\to\infty} P(Y \in C) = \int_C dP_Y(y) = \int_C \lim_{n\to\infty} F_y(x_n)\ dP_Y(y),$$

Tedy $\lim_{x\to\infty} F_y(x)|_{\mathbb{Q}} = 1$. Podobně pro $x_n \to -\infty$, $x_n \in \mathbb{Q}$, dostáváme

$$P(X \le x_n, Y \in C) \xrightarrow{n\to-\infty} 0 = \int_C 0\ dP_Y(y) = \int_C \lim_{n\to-\infty} F_y(x_n)\ dP_Y(y),$$

potom $\lim_{x\to-\infty} F_y(x)|_{\mathbb{Q}} = 0$.

Věta 2.1 (Existence podmíněného rozdělení)

Nechť $X \in \mathbb{L}(\mathcal{A})$, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ jsou náhodné veličiny. Pak existuje podmíněné rozdělení $P_{X|Y=y}$.

Důkaz

Pro libovolné $B \in \mathcal{B}(\mathbb{R})$ a $C \in \mathcal{H}$ definujeme míru μ předpisem

$$\mu(B, C) = P(X \in B, Y \in C).$$

Pak pro $C \in \mathcal{H}$ je funkce $B \mapsto \mu(B, C)$ konečná míra na $\mathcal{H}$.

Nechť pro $x \in \mathbb{Q}$ je $B_x = (-\infty, x]$, ukážeme, že $\mu_x := \mu((-\infty, x], \cdot) \ll P_Y$ na $(H, \mathcal{H})$. Nechť $C \in \mathcal{H}$ takové, že $P_Y(C) = P(Y \in C) = 0$, pak

$$\mu_x(C) = \mu((-\infty, x], C) = P(X \in (-\infty, x], Y \in C) = 0.$$

Tedy $\mu_x \ll P_Y$ na $(H, \mathcal{H})$. Z Radon-Nikodymovy věty víme:

$$\exists F_y(x) := \frac{d\mu_x}{dP_Y}(y), \quad y \in H, \quad x \in \mathbb{Q}.$$

Definujeme $S := \{y \in H | \exists$ distribuční funkce F, že $F_y(x) = F(x), \forall x \in \mathbb{Q}\} \in \mathcal{H}$, pak z předchozí Poznámky plyne, že existuje rozšíření F_y na distribuční funkci, tedy $P_Y(S) = 1$.

Definujeme pro $x \in \mathbb{Q}$ funkci

$$\overline{F_y}(x) = \begin{cases} F_y(x), & y \in S \\ F_X(x), & y \notin S. \end{cases}$$

Z Radon-Nikodymovy věty plyne:

$$P(X \leq x, Y \in C) = \mu_x(C) = \int_C F_y(x)\ dP_Y(y), \quad x \in \mathbb{Q}.$$

Definujeme míru ν_y následujícím vztahem

$$\nu_y((-\infty, x]) = \begin{cases} F_y(x), & y \in S \\ F_X(x), & y \notin S. \end{cases}$$

Pak platí, že $\overline{F_y}(x) = \nu_y((-\infty, x])$, $x \in \mathbb{Q}$. Tvrdíme, že $\nu_y = P_{X|Y=y}$. Stačí už jen ověřit tři vlastnosti podmíněného rozdělení z definice.

- pro každé $y \in H$ je ν_y pravděpodobnost na $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, toto plyne z definice ν_y.
- Chceme ukázat, že pro každou $B \in \mathcal{B}(\mathbb{R})$ je funkce $\nu_y(B) \in \mathbb{L}(\mathcal{H})$. Nejprve tuto vlastnost ověříme na množinách B_x pro $x \in \mathbb{Q}$. Protože $S \in \mathcal{H}$, platí

  $$y \mapsto \nu_y(B_x) = 1_S(y)\frac{d\mu_x}{dP_Y}(y) + 1_{H\setminus S}(y)F_X(x) \in \mathbb{L}(\mathcal{H}).$$

  Tedy jsme vlastnost ověřili pro systém $\{B_x, x \in \mathbb{Q}\}$.

  Zároveň ale víme, že tento systém je uzavřený na konečné průniky. Nyní se zaměříme na systém

  $$\mathcal{D} = \{B \in \mathcal{B}(\mathbb{R}), \nu_y(B) \in \mathbb{L}(\mathcal{H})\}.$$

  Z definice snadno ověříme, že tento systém je Dynkinův. Navíc z předchozího platí, že $\{B_x, x \in \mathbb{Q}\} \subseteq \mathcal{D}$. Protože ale

  $$\sigma\{B_x, x \in \mathbb{Q}\} = \mathcal{B}(\mathbb{R}),$$

  pak z Dynkinova lemmatu vyplývá, že $\nu_y \in \mathbb{L}(\mathcal{H})$ pro každou $B \in \mathcal{B}(\mathbb{R})$.
- Chceme ukázat, že pro $B \in \mathcal{B}(\mathbb{R}), C \in \mathcal{H}$ platí:

  $$\mu(B, C) = P(X \in B, Y \in C) = \int_C \nu_y(B)\ dP_Y(y).$$

  Použijeme stejný postup jako u předchozího bodu, tedy nejprve vlastnost ověříme na $\{B_x, x \in \mathbb{Q}\}$:

  $$\mu(B_x, C) = P(X \leq x, Y \in C) = \int_C F_y(x)\ dP_Y(y) = \\ = \int_C \nu_y((-\infty, x])\ dP_Y(y).$$

  Nyní definujeme systém

  $$\mathcal{D}' = \{B \in \mathcal{B}(\mathbb{R}), \mu(B, C) = \int_C \nu_y(B)\ dP_Y(y)\}.$$

  I tento systém je Dynkinův, jak se můžeme snadno z definice přesvědčit.

Stejnými úvahami jako v předchozím bodě dojdeme k závěru, že třetí vlastnost je splněna.

□

Věta 2.2 (Jednoznačnost podmíněného rozdělení)
Nechť $X \in \mathbb{L}(\mathcal{A})$, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ jsou náhodné veličiny. Pak podmíněné rozdělení $P_{X|Y=y}$ je určeno P_Y-skoro jistě jednoznačně.

Důkaz
Nechť existují dvě podmíněná rozdělení X za podmínky $Y = y$, označíme je μ_y a ν_y. Víme, že obě rozdělení splňují pro $B \in \mathcal{B}(\mathbb{R})$, $C \in \mathcal{H}$:

$$P(X \in B, Y \in C) = \int_C \mu_y(B)\ dP_Y(y) = \int_C \nu_y(B)\ dP_Y(y).$$

Tedy pro $B \in \mathcal{B}(\mathbb{R})$ je $\mu_y(B) = \nu_y(B)$, P_Y-skoro všude, protože $\mathbb{Q}$ je spočetná, máme:

$$\mu_y((-\infty, q]) = \nu_y((-\infty, q]), \quad q \in \mathbb{Q}, \quad P_Y - s.v.$$

Dále platí pro $x \in \mathbb{R}$, že $\mu_y((-\infty, x]) = \nu_y((-\infty, x])$, P_Y-skoro všude. Tedy máme, že $\mu_y = \nu_y$, P_Y-skoro všude, neboť pravděpodobnostní míra na $\mathbb{R}$ je dána distribuční funkcí.

□

Definice 2.2
Řekneme, že prostor $(S, \mathcal{S})$ je *izomorfní* prostoru $(T, \mathcal{T})$, pokud existuje prosté $f : (S, \mathcal{S}) \to (T, \mathcal{T})$, takové, že $f^{-1} : (T, \mathcal{T}) \to (S, \mathcal{S})$, značíme $(S, \mathcal{S}) \overset{f}{\sim} (T, \mathcal{T})$.

Poznámka
Pokud $(S, \mathcal{S}) \overset{f}{\sim} (T, \mathcal{T})$, pak můžeme značit $f : (S, \mathcal{S}) \leftrightarrow (T, \mathcal{T})$, navíc pro $A \in \mathcal{S}$ označíme $f(A)$ jako vzor množiny A přes zobrazení f^{-1}.

Věta 2.3
Nechť $X \in \mathbb{L}(\mathcal{A}, \mathcal{S})$, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ jsou náhodné veličiny, nechť platí, že $(S, \mathcal{S}) \sim (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Potom existuje podmíněné rozdělení $P_{X|Y=y}$.

Důkaz
Z předpokladů víme, že existuje prosté zobrazení $f : (S, \mathcal{S}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, takové, že $f^{-1} : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (S, \mathcal{S})$.

Definujeme náhodnou veličinu $Z := f(X)$, pak $Z \in \mathbb{L}(\mathcal{A})$. Z Věty 2.1 o existenci podmíněného rozdělení víme, že existuje $P_{Z|Y=y}$. Chceme definovat $P_{X|Y=y}$. Nechť $A \in \mathcal{S}$, označíme $B := f(A) \in \mathcal{B}(\mathbb{R})$, tedy

$$X = f^{-1}(Z) \in A \quad \Leftrightarrow \quad Z \in f(A) \quad \Leftrightarrow \quad Z \in B,$$

$$\begin{aligned} P_{X|Y=y}(A) &= P(X \in A|Y = y) = P(Z \in B|Y = y) = P_{Z|Y=y}(B) = \\ &= P_{Z|Y=y}(f(A)) = (P_{Z|Y=y} \circ f)(A). \end{aligned}$$

Nyní ukážeme, že zobrazení

$$(A, y) \in \mathcal{S} \times H \mapsto (P_{Z|Y=y} \circ f)(A) \in [0, 1]$$

splňuje axiomy z definice podmíněného rozdělení $P_{X|Y=y}$.

- pro každé $y \in H$: $P_{X|Y=y}(A)$ je dle definice pravděpodobnost v $\mathcal{S}$.
- nechť $A \in \mathcal{S}$. Protože

  $$\mathcal{S} = \{f^{-1}(B), B \in \mathcal{B}(\mathbb{R})\} = \{[f \in B], B \in \mathcal{B}(\mathbb{R})\},$$

  existuje $B \in \mathcal{B}(\mathbb{R})$ takové, že $A = f^{-1}(B)$. Pak

  $$P_{X|Y=y}(A) = (P_{Z|Y=y} \circ f)(A) = P_{Z|Y=y}(B), \quad B \in \mathcal{B}(\mathbb{R}),$$

  tudíž $P_{X|Y=y}(A) \in \mathbb{L}(\mathcal{H})$.
- Nechť $A \in \mathcal{S}$, $C \in \mathcal{H}$:

  $$P(X \in A, Y \in C) = P(Z \in B, Y \in C) = \int_C P_{Z|Y=y}(B)\ dP_Y(y) = \\ = \int_C P_{X|Y=y}(B)\ dP_Y(y).$$

□

Poznámka

Uvedeme si protipříklad, kdy neexistuje podmíněné rozdělení. Buď $C \subseteq [0, 1]$ s vnitřní Lebesgueovou mírou 0 a vnější 1. Zvolme $\Omega = [0, 1]$, $\mathcal{A} = \sigma\{\mathcal{B} \cup \{C\}\}$, kde $\mathcal{B}$ značí borelovskou σ-algebru na $[0, 1]$. Označme $Q = R[0, 1]$ Lebesgueovu míru na $\mathcal{B}$. Definujme míru P následujícím vztahem

$$P(B \cap C + D \setminus C) := \frac{Q(B) + Q(D)}{2}, \quad B, D \in \mathcal{B},$$

kde + značí disjunktní sjednocení. Ověření toho, že výše uvedená formule korektně definuje pravděpodobnostní míru na $\mathcal{B}$ rozšiřující míru Q, necháváme na čtenáři.

Protože z definice je $P(C) = \frac{1}{2}$, pak pro $B \in \mathcal{B}$ platí

$$P(B \cap C) = \frac{P(B)}{2} = P(B)P(C),$$

tedy C je nezávislá s $\mathcal{B}$ a stejně tak $[0, 1] \setminus C$ je nezávislá s $\mathcal{B}$.

Definujeme náhodné veličiny $X : (\Omega, \mathcal{A}) \to (\Omega, \mathcal{B})$, X má rovnoměrné rozdělení na $(0, 1)$, a $Y : (\Omega, \mathcal{A}) \to (\Omega, \mathcal{A})$, $Y(\omega) = \omega$. Pak tvrdíme, že neexistuje podmíněné rozdělení $P_{Z|X}$.

Intuice nám říká, že pokud známe hodnotu X, měli bychom mít $Z = X$. Jediný možný kandidát na podmíněné rozdělení je tedy Diracova míra: $P_{Z|X=x} \sim N(x, 0)$. Tento systém rozdělení ale nesplňuje

$$\frac{1}{2} = P(C) = \int_0^1 P(C|X = x)\ dx,$$

protože pravá strana není dobře definována, neboť v x je $P(C|X=x)=1_C(x)$ lebesgueovsky neměřitelná funkce, tedy není borelovská.

Poznámka

K předchozí poznámce ještě doplníme zajímavou situaci, kdy jedna náhodná veličina je funkcí druhé, přesto je s ní nezávislá. Definujeme náhodné veličiny $X : (\Omega, \mathcal{A}) \to (\Omega, \mathcal{B})$ s rovnoměrným rozdělením na $(0,1)$ a $Y = 1_C(X)$, která má diskrétní rovnoměrné rozdělení na $\{0,1\}$. Pak veličiny X a Y jsou nezávislé. Tato situace může nastat proto, že náhodná veličina Y sice závisí funkčně na X, ale tato funkce je neměřitelná.

Definice 2.3

Je-li $(S, \mathcal{S})$ měřitelný prostor, řekneme, že $(T, \mathcal{T})$ je jeho *měřitelný podprostor*, pokud

(i) $T \in \mathcal{S}$,

(ii) $\mathcal{T} = \{B \cap T; B \in \mathcal{S}\}$.

Poznámka

Nyní tedy víme o existenci podmíněného rozdělení pro reálné náhodné veličiny a také pro prostory, které jsou izomorfní s podprostorem $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. V dalším textu budeme směřovat k důkazu toho, že každý separabilní úplný metrický prostor je izomorfní podprostoru $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, tedy i na něm je zajištěna existence podmíněného rozdělení.

Definice 2.4

Nechť (P, ρ) je metrický prostor. Řekneme, že je *polský*, pokud existuje metrika d na P taková, že $d \sim \rho$ a prostor (P, d) je úplný a separabilní.

Poznámka

Víme, že např. prostor $\mathbb{R}$ s eukleidovskou metrikou je polský, neboť je úplný a separabilní ($\mathbb{Q}$ je spočetná, hustá podmnožina). Pokud definujeme prostor se součinouvou metrikou

$$\mathbb{R}^{\mathbb{N}} = \prod_{n=1}^{\infty} \mathbb{R}, \quad d(u,v) = \sum_{n=1}^{\infty} 2^{-k} \wedge |u_k - v_k|, \quad u, v \in \mathbb{R}^{\mathbb{N}}.$$

Pak i tento prostor je polský, což lze ověřit z definice.

Věta 2.4

Nechť P je polský prostor. Pak podprostor $Y \subseteq P$ je polský právě tehdy, když je typu G_δ v P.

Důkaz

Nejprve dokážeme implikaci "$\Leftarrow$". Nechť $\{U_n\}$ je posloupnost otevřených podmnožin P, nechť $Y = \bigcap_{n=1}^{\infty} U_n$. Protože každá otevřená podmnožina polského prostoru je polský prostor (důkaz lze nalézt např. v knize [2], str. 251, Proposition 8.1.1), pak U_n je polský pro každé $n \in \mathbb{N}$, a také $\prod_{n=1}^{\infty} U_n$ je polský prostor (opět důkaz neuvádíme, lze jej vyhledat v knize [2], str. 253, Proposition 8.1.3). Definujeme $\Delta \subseteq \prod_{n=1}^{\infty} U_n$:

$$\Delta = \left\{ \{u_n\} \in \prod_{n=1}^{\infty} U_n : u_j = u_k, \forall j, k \right\}.$$

Potom Δ je uzavřená podmnožina $\prod_{n=1}^{\infty} U_n$, a tedy také polská (opět Proposition 8.1.1). Navíc Y je homeomorfní Δ, neboť můžeme definovat zobrazení

$$f : Y \to \Delta, \quad y \mapsto \{y\}_{n=1}^{\infty}.$$

Tedy také Y je polský prostor.

Nyní dokážeme obrácenou implikaci "$\Rightarrow$". Předpokládejme, že podprostor Y prostoru P je polský. Nechť d je metrika na P a d_0 je úplná metrika na Y. Pro $n \in \mathbb{N}$ definujeme

$$V_n = \bigcup_{n=1}^{\infty} \left\{ W \subseteq P : W \cap Y \neq \emptyset, diam_{d_0}(W) \leq \frac{1}{n} \right\}.$$

Nyní dokážeme, že

$$Y = \overline{Y} \cap \left(\bigcap_{n=1}^{\infty} V_n \right).$$

Zřejmě platí

$$Y \subset \overline{Y} \cap \left(\bigcap_{n=1}^{\infty} V_n \right),$$

zaměříme se tedy na opačnou inkluzi.

Nechť $x \in \overline{Y} \cap (\bigcap_{n=1}^{\infty} V_n)$. Protože $x \in \bigcap_{n=1}^{\infty} V_n$, můžeme najít posloupnost $\{W_n\}$ takovou, že

$$W_n \cap Y \neq \emptyset, \quad diam_{d_0}(W_n) < \frac{1}{n}, \quad x \in W_n, \quad n \in \mathbb{N}.$$

BÚNO předpokládejme, že $\{W_n\}$ je klesající a že $diam_d(W_n) < \frac{1}{n}$. Protože Y je úplný prostor s metrikou d_0, pak platí, že existuje právě jeden prvek $y \in Y$ takový, že $y \in \overline{W_n \cap Y}$. Navíc $diam_d(\overline{W_n}) < \frac{1}{n}$ a $\{x, y\} \subseteq \overline{W_n}$. Tedy $x = y$.

Protože každá uzavřená podmnožina metrického prostoru P je typu G_δ (důkaz opět v knize [2], str. 347, D24), pak ze vztahu, který jsme dokázali, plyne, že Y je typu G_δ v P.

$\square$

Věta 2.5 (Urysohnova věta)

Nechť (S, ρ) je polský prostor. Pak existuje množina $B \in G_\delta([0,1]^{\mathbb{N}})$ taková, že S a B jsou homeomorfní, tedy existuje $f : S \to B$ spojitá bijekce taková, že $f^{-1} : B \to S$ je spojité zobrazení.

Důkaz

BÚNO předpokládejme, že metrika ρ nabývá hodnot v $[0, 1]$, jinak použijeme metriku $\rho_1 = 1 \wedge \rho$. Nechť $\{x_n\}$ je hustá posloupnost v S, položme

$$f : x \in S \mapsto \{\rho(x, x_n)\}, \quad f(x) \in [0,1]^{\mathbb{N}},$$

tedy každému x přiřadíme posloupnost vzdáleností $\rho(x, x_n)$ od $\{x_n\}$. Označíme $B = f(S) \subseteq [0,1]^{\mathbb{N}}$. Ověříme, že se jedná o homeomorfismus.

Nejprve ukážeme, že f je prosté. Nechť $x, y \in S$ takové, že $\rho(x, x_n) = \rho(y, x_n)$, $n \in \mathbb{N}$. Z hustoty $\{x_n\}$ v S máme, že $\forall \epsilon > 0 \exists n \in \mathbb{N}$ takové, že $\rho(x, x_n) < \frac{\epsilon}{2}$. Tedy z trojúhelníkové nerovnosti plyne :

$$\rho(x, y) \leq \rho(x, x_n) + \rho(x_n, y) = 2\rho(x, x_n) < \epsilon.$$

Tedy $\rho(x, y) = 0$, takže $x = y$, z čehož plyne, že f je prosté.

Nyní dokážeme, že f je spojité zobrazení. Nechť $y_k \to y$ v S, pak také platí, že $\rho(y_k, x_n) \to \rho(y, x_n)$, $k \to \infty$, $n \in \mathbb{N}$, neboť

$$|\rho(y_k, x_n) - \rho(y, x_n)| \leq \rho(y, y_k) \xrightarrow{k\to\infty} 0.$$

Konvergence složek $f(y_k)_n = \rho(y_k, x_n) \to \rho(y, x_n) = f(y)_n$ dává konvergenci $f(y_k) \to f(y)$ v metrickém prostoru $[0,1]^{\mathbb{N}}$. Tedy f je spojité.

Dalším krokem bude dokázat spojitost f^{-1}. Nechť $y_k, y \in S$ jsou takové, že $f(y_k)_n = \rho(y_k, x_n) \to \rho(y, x_n) = f(y)_n$, $n \in \mathbb{N}$. Ukážeme, že $y_k \to y$ v S, tedy že $\rho(y_k, y) \to 0$. Buď $\epsilon > 0$, $n \in \mathbb{N}$ takové, že $\rho(x_n, y) < \frac{\epsilon}{3}$. Najdeme $k_0 \in \mathbb{N}$, aby $\forall k \geq k_0$ platilo

$$|f(y_k)_n - f(y)_n| = |\rho(y_k, x_n) - \rho(y, x_n)| < \frac{\epsilon}{3}.$$

Pak z trojúhelníkové nerovnosti plyne, že

$$\rho(y, y_k) \leq \rho(y, x_n) + \rho(x_n, y_k) \leq \frac{\epsilon}{3} + 2\rho(y, x_n) < \epsilon.$$

Tedy z již dokázaného plyne, že f je homeomorfismus S na $f(S)$ Nyní dokážeme, že $f(S)$ je polský prostor. Na tomto prostoru uvažujme nejprve metriku přenesenou z (S, ρ) pomocí f, tedy $r(u, v) = \rho(f^{-1}(u), f^{-1}(v))$, pak tato metrika zdědí jak separabilitu, tak úplnost.

Chceme tedy ověřit, že je r ekvivalentní se součinovou metrikou

$$d(u, v)|_{f(S)} = \sum_{n=1}^{\infty} 2^{-k} \wedge |u_k - v_k|.$$

Nechť $y_k \to y$, tedy $\rho(y_k, y) \to 0$, $k \to \infty$. Navíc platí, že

$$\rho(y_k, y) = \rho(f^{-1}(f(y_k)), f^{-1}(f(y))) = r(f(y_k), f(y)).$$

Tedy vidíme, že $y_k \to y$ v $\rho \Leftrightarrow f(y_k) \to f(y)$ v r. Protože f je homeomorfismus, pak $y_k \to y$ v $\rho \Leftrightarrow f(y_k) \to f(y)$ v $d_{f(S)}$. Tedy dohromady $d_{f(S)} \sim r$.

Dokázali jsme, že $f(S)$ je polský prostor, pak z předchozí Věty 2.4 vyplývá, že navíc $f(S) \in G_\delta([0,1]^{\mathbb{N}})$.

$\square$

Poznámka

Z předchozí Věty 2.5 speciálně plyne, že $f(S) \in G_\delta([0,1]^{\mathbb{N}}) \subseteq \mathcal{B}([0,1]^{\mathbb{N}})$, tedy platí, že prostory $(S, \mathcal{B}(S))$, $(B, \mathcal{B}(B))$ jsou izomorfní, měřitelné prostory. Nyní dokážeme, že $(B, \mathcal{B}(B)) \sim ([0,1], \mathcal{B}([0,1]^{\mathbb{N}}))$.

Poznámka

Nechť $f \in \mathbb{L}(\mathcal{A}, \mathcal{S})$, $g \in \mathbb{L}(\mathcal{B}, \mathcal{T})$, pak

(i) $f \otimes g \in \mathbb{L}(\mathcal{A} \otimes \mathcal{B}, \mathcal{S} \otimes \mathcal{T})$, kde $(f \otimes g)(x, y) = (f(x), g(y))$ je tenzorový součin funkcí f a g,

(ii) $f^{\mathbb{N}} \in \mathbb{L}(\mathcal{A}^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$, kde $f^{\mathbb{N}}$ je zkrácený zápis pro tenzorovou mocninu $f^{\otimes\mathbb{N}}$, tedy

$$f^{\mathbb{N}}(x_n, n \in \mathbb{N}) = (f(x_n))_{n=1}^{\infty},$$

a $\mathcal{A}^{\mathbb{N}}$ je zkrácený zápis pro mocninovou σ-algebru $\mathcal{A}^{\otimes\mathbb{N}}$.

Důkaz

(i) σ-algebra $\mathcal{S} \otimes \mathcal{T}$ je generována množinami typu $S \times T$, $S \in \mathcal{S}$, $T \in \mathcal{T}$, stačí tedy ukázat, že

$$(f \otimes g)^{-1}(S \times T) \in \mathcal{A} \otimes \mathcal{B}.$$

Zřejmě

$$\begin{aligned}(f \otimes g)^{-1}(S \times T) &= [f \otimes g \in S \times T] = [f \in S] \times [g \in T] = \\ &= f^{-1}S \times g^{-1}T \in \mathcal{A} \otimes \mathcal{B},\end{aligned}$$

neboť $[f \in S] = f^{-1}S \in \mathcal{A}$ a $[g \in T] = g^{-1}T \in \mathcal{B}$ dle předpokladu.

(ii) Podobně σ-algebra $\mathcal{S}^{\mathbb{N}}$ je generována projekcemi do souřadnic. Stačí ukázat, že $f_n^{\mathbb{N}} \in \mathbb{L}(\mathcal{A}^{\mathbb{N}}, \mathcal{S})$, kde

$$f_n^{\mathbb{N}}(x_k, k \in \mathbb{N}) = f(x_n)$$

je n-tá souřadnice funkce $f^{\mathbb{N}}(x_k, k \in \mathbb{N})$. Zde

$$(f_n^{\mathbb{N}})^{-1}S = \left[f_n^{\mathbb{N}} \in S\right] = \mathbb{A}^{n-1} \times [f_n \in S] \times \mathbb{A}^{\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}, \quad S \in \mathcal{S},$$

kde $\mathbb{A} = \bigcup \mathcal{A}$. Tedy $f_n^{\mathbb{N}} \in \mathbb{L}(\mathcal{A}^{\mathbb{N}}, \mathcal{S})$, tedy $f^{\mathbb{N}} \in \mathbb{L}(\mathcal{A}^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$.

$\square$

Důsledek

Označíme $\mathbb{A} := \bigcup \mathcal{A}$, $\mathbb{B} := \bigcup \mathcal{B}$, $\mathbb{S} := \bigcup \mathcal{S}$, $\mathbb{T} := \bigcup \mathcal{T}$. Je-li f měřitelný izomorfismus mezi $(\mathbb{A}, \mathcal{A})$ a $(\mathbb{S}, \mathcal{S})$, tedy $f \in \mathbb{L}(\mathcal{A}, \mathcal{S})$ a $f^{-1} \in \mathbb{L}(\mathcal{S}, \mathcal{A})$ a g mezi $(\mathbb{B}, \mathcal{B})$ a $(\mathbb{T}, \mathcal{T})$, pak

(i) $f \otimes g \in \mathbb{L}(\mathcal{A} \otimes \mathcal{B}, \mathcal{S} \otimes \mathcal{T})$ a $(f \otimes g)^{-1} = f^{-1} \otimes g^{-1} \in \mathbb{L}(\mathcal{S} \otimes \mathcal{T}, \mathcal{A} \otimes \mathcal{B})$, což znamená, že $f \otimes g$ je měřitelný izomorfismus mezi $(\mathbb{A}, \mathcal{A}) \otimes (\mathbb{B}, \mathcal{B})$ a $(\mathbb{S}, \mathcal{S}) \otimes (\mathbb{T}, \mathcal{T})$.

(ii) $f^{\mathbb{N}} \in \mathbb{L}(\mathcal{A}^{\mathbb{N}}, \mathcal{S}^{\mathbb{N}})$, a také $(f^{\mathbb{N}})^{-1} = (f^{-1})^{\mathbb{N}} \in \mathbb{L}(\mathcal{A}^{\mathbb{N}}, \mathcal{A}^{\mathbb{N}})$, což znamená, že $f^{\mathbb{N}}$ je izomorfismus mezi $(\mathbb{A}, \mathcal{A})^{\mathbb{N}} := (\mathbb{A}, \mathcal{A})^{\otimes\mathbb{N}}$ a $(\mathbb{S}, \mathcal{S})^{\mathbb{N}} := (\mathbb{S}, \mathcal{S})^{\otimes\mathbb{N}}$.

Věta 2.6

Nechť $B \in \mathcal{B}(P)$, kde P je polský prostor. Pak existuje $C \in \mathcal{B}([0,1])$, která je borelovsky izomorfní B.

Důkaz

Z předchozí Věty 2.5 vidíme, že $B \overset{f}{\sim} G$, kde $G \in \mathcal{B}([0,1]^{\mathbb{N}})$, definujeme množinu $J \subseteq \{0,1\}^{\mathbb{N}}$ následujícím způsobem

$$J = \left\{ (0,0,\ldots) \cup \{x \in [0,1]^{\mathbb{N}}, \sum_{n=1}^{\infty} x_n = \infty\} \right\},$$

a definujeme funkci $F : [0,1] \to J$,

$$x = \sum_{n=1}^{\infty} \frac{x_n}{2^n} \in [0,1] \mapsto (x_1, x_2, \ldots),$$

kde $x_n \in \{0,1\}$, $n \in \mathbb{N}$, $\sum_{n=1}^{\infty} x_n = \infty$.

Nyní ukážeme, že F je borelovské prosté na J. Nechť $x = (x_1, x_2, \ldots) \in J$, pak existuje vzor $F^{-1}(x) = \sum_{n=1}^{\infty} \frac{x_n}{2^n} \in [0,1]$. Nechť navíc $y = (y_1, y_2, \ldots)$ takový, že $x = y$, pak

$$\sum_{n=1}^{\infty} \frac{x_n}{2^n} = \sum_{n=1}^{\infty} \frac{y_n}{2^n},$$

tedy F je bijekce mezi $[0,1]$ a J. Z předchozí Poznámky pak plyne, že

$$[0,1] \overset{F}{\sim} J \quad \Rightarrow \quad [0,1]^{\mathbb{N}} \overset{F^{\mathbb{N}}}{\sim} J^{\mathbb{N}} \in \mathcal{B}(\{0,1\}^{\mathbb{N}^2}).$$

Protože existuje bijekce π mezi $\mathbb{N}$ a $\mathbb{N}^2$, můžeme definovat funkci g následujícím předpisem

$$g : x \in \{0,1\}^{\mathbb{N}^2} \mapsto (x_{\pi(n)}, n \in \mathbb{N}) \in \{0,1\}^{\mathbb{N}}.$$

Protože v obou prostorech $\{0,1\}^{\mathbb{N}^2}$ a $\{0,1\}^{\mathbb{N}}$ se bere konvergence po složkách, je zobrazení g homeomorfismus a tedy máme

$$\{0,1\}^{\mathbb{N}^2} \overset{g}{\sim} \{0,1\}^{\mathbb{N}}.$$

Definujeme zobrazení $h : \{0,1\}^{\mathbb{N}} \to \mathcal{C}$, $(x_1, x_2, \ldots) \mapsto \sum_{n=1}^{\infty} \frac{2x_n}{3^n}$, kde

$$\mathcal{C} = \left\{ \sum_{n=1}^{\infty} \frac{2x_n}{3^n}, \forall n \in \mathbb{N} : x_n \in \{0,1\} \right\} \subset [0,1].$$

$\mathcal{C}$ je tzv. Cantorovo diskontinuum. Pak h je homeomorfismus, tedy borelovské prosté zobrazení na $\mathcal{C}$. Nechť $x = (x_1, x_2, \ldots)$, $y = (y_1, y_2, \ldots) \in \{0,1\}^{\mathbb{N}}$ takové, že $h(x) = h(y)$, pak zřejmě $x = y$, navíc $\forall x \in \mathcal{C}$ takové, že $x = \sum_{n=1}^{\infty} \frac{2x_n}{2^n}, n \in \mathbb{N}$, platí, že $h^{-1}(x) = (x_1, x_2, \ldots) \in \{0,1\}^{\mathbb{N}}$.

Nechť $\{(x_1^{(n)}, x_2^{(n)}, \ldots\}_n$ je posloupnost prvků z $\{0,1\}^{\mathbb{N}}$ konvergující k prvku $(x_1, x_2, \ldots)$. Pak $\forall k \in \mathbb{N}$ platí, že $k_k^{(n)} \to x_k$, a tedy

$$\lim_{n\to\infty} \left| \sum_{k=1}^{\infty} \frac{2(x_k^{(n)} - x_k)}{3^k} \right| \leq \sum_{k=1}^{\infty} \frac{2 \lim_{n\to\infty} |x_k^{(n)} - x_k|}{3^k} = 0.$$

Tedy h je homeomorfismus a podle předchozí věty platí, že $\mathcal{C} \in \mathcal{B}([0,1])$. Celkem máme, že

$$B \stackrel{h \circ g \circ F^{\mathbb{N}} \circ f}{\sim} \mathcal{C}.$$

□

Věta 2.7 (O transformaci)

Nechť $X \in \mathbb{L}(\mathcal{A}, \mathcal{S})$, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ jsou náhodné veličiny takové, že existuje podmíněné rozdělení $P_{X|Y=y}$, nechť $f : (S, \mathcal{S}) \to (T, \mathcal{T})$ měřitelná. Pak podmíněné rozdělení $f(X)$ za podmínky $Y = y$ je

$$P_{f(X)|Y=y} = P_{X|Y=y} \circ f^{-1}.$$

Důkaz

Víme, že $P_{X|Y=y}$ je podmíněné rozdělení, ověříme vlastnosti z definice pro podmíněné rozdělení $P_{f(X)|Y=y}$.

- pro každé $y \in H$ je funkce $B \mapsto P_{f(X)|Y=y}(B)$ pravděpodobnost pro $B \in \mathcal{T}$, neboť
$$P_{f(X)|Y=y}(B) = (P_{X|Y=y} \circ f^{-1})(B) = P_{X|Y=y}(f^{-1}(B)).$$
- pro každou $B \in \mathcal{T}$ je funkce $y \mapsto P_{f(X)|Y=y}(B)$ $\mathcal{H}$-měřitelná pro $y \in H$, neboť
$$P_{f(X)|Y=y}(B) = (P_{X|Y=y} \circ f^{-1})(B) = P_{X|Y=y}(f^{-1}(B)).$$
- Pro $B \in \mathcal{T}$, $D \in \mathcal{H}$ platí:
$$P(f(X) \in B, Y \in D) = \int_D P_{X|Y=y}(f^{-1}(B))\ dP_Y(y) =$$
$$= \int_D (P_{X|Y=y} \circ f^{-1})(B)\ dP_Y(y) = \int_D P_{f(X)|Y=y}(B)\ dP_Y(y).$$

□

2.2 Podmíněná střední hodnota

Nyní se budeme zabývat podmíněnou střední hodnotou. Zavedeme ji úplně analogicky, jako se zavádí v nepodmíněném případě a ukážeme si její základní vlastnosti.

V celé této kapitole budeme vždy předpokládat existenci podmíněného rozdělení, neboť tato problematika byla již dostatečně rozebrána výše.

Definice 2.5

Nechť $X \in \mathbb{L}(\mathcal{A}, \mathcal{S})$, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ jsou náhodné veličiny. Pak *podmíněnou střední hodnotu* X při $Y = y$ definujeme jako

$$E[X|Y = y] = \int_S x\ dP_{X|Y=y}(x).$$

Poznámka

Pro podmíněnou střední hodnotu platí, že $E[X|Y=y] = f(y)$, kde f je měřitelná funkce. Pak $E[X|Y]$ je náhodná veličina $f(Y)$.

Poznámka

Z Věty 2.7 víme, že $P_{f(X)|Y} = P_{X|Y} \circ f^{-1}$, pro f měřitelnou. Z toho můžeme odvodit vztah

$$E[f(X)|Y] = \int f(x)\ dP_{X|Y}(x),$$

neboť pro $Z := f(X)$ platí:

$$E[f(X)|Y] = E[Z|Y] = \int z\ dP_{Z|Y}(z) = \int z\ dP_{f(X)|Y}(z) =$$

$$= \int z\ d(P_{X|Y} \circ f^{-1})(z) = \int f(x)\ dP_{X|Y}(x).$$

Věta 2.8 (Základní vlastnosti podmíněné střední hodnoty)

Nechť $X, Z \in \mathbb{L}(\mathcal{A}, \mathcal{S})$, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ jsou náhodné veličiny. Pak platí:

a) $a, b, c \in \mathbb{R}$, pak $E[aX + bZ + c|Y=y] = aE[X|Y=y] + bE[Z|Y=y] + c$,

b) $X \leq Z$ skoro jistě, pak $E[X|Y=y] \leq E[Z|Y=y]$,

c) pro f měřitelnou platí $X = f(Y)$, pak $E[X|Y] = X$,

d) X a Y nezávislé, pak $E[X|Y=y] = EX$,

e) pro f měřitelnou platí $E\left[E\left[f(X)|Y\right]\right] = Ef(X)$, specielně $E\left[E\left[X|Y\right]\right] = EX$.

Důkaz

a) Definujeme funkce $f(x,z) := ax + bz + c$, $g_1(x,z) := x$, $g_2(x,z) := z$, pak z Věty 2.7 dostáváme:

$$E[aX+bZ+c|Y=y] = E[f(X,Z)|Y=y] = \int_S f(x,z)\ dP_{(X,Z)|Y=y}(x,z) =$$

$$= \int_S ax + bz + c\ dP_{(X,Z)|Y=y}(x,z) = a\int_S x\ dP_{(X,Z)|Y=y}(x,z) +$$

$$+ b\int_S z\ dP_{(X,Z)|Y=y}(x,z) + c\int_S\ dP_{(X,Z)|Y=y}(x,z) = aE[g_1(X,Z)|Y=y] +$$

$$+ bE[g_2(X,Z)|Y=y] + c = aE[X|Y=y] + bE[Z|Y=y] + c.$$

b) Definujeme funkce $g_1(x,z) := x$, $g_2(x,z) := z$, pak

$$g_1(X,Z) = X \leq Z = g_2(X,Z),$$

$$E[X|Y=y] = E[g_1(X,Z)|Y=y] = \int_S g_1(x,z)\ dP_{(X,Z)|Y=y}(x,z) \leq$$

$$\leq \int_S g_2(x,z)\ dP_{(X,Z)|Y=y}(x,z) = E[g_2(X,Z)|Y=y] = E[Z|Y=y].$$

c)

$$E[X|Y] = \int_S x \; dP_{X|Y}(x) = \int_S f(y) \; dP_{X|Y}(x) = f(Y) = X$$

d) Z nezávislosti víme: $P(X \in B, Y \in C) = P(X \in B)P(Y \in C)$, pro $B \in \mathcal{S}$, $C \in \mathcal{H}$. Dále platí:

$$P(X \in B)P(Y \in C) = \int_B \; dP_X(x) \int_C \; dP_Y(y) = \int_C \int_B \; dP_X(x) \; dP_Y(y),$$

z definice podmíněného rozdělení máme:

$$P(X \in B, Y \in C) = \int_C P_{X|Y=y}(B) \; dP_Y(y) = \int_C \int_B \; dP_{X|Y=y}(x) \; dP_Y(y),$$

tedy pro každou $B \in \mathcal{S}$ je $P_X(B) = P_{X|Y=y}(B)$, tedy

$$E[X|Y=y] = \int_S x \; dP_{X|Y=y}(x) = \int_S x \; dPX(x) = EX.$$

e) Nechť nejprve $f = 1_B$, $B \in \mathcal{S}$, pak

$$P(X \in B, Y \in C) = \int_C P_{X|Y=y}(B) \; dP_Y(y) = E[P_{X|Y}(B); Y \in C],$$

označíme $P_{X|Y}(B) = P(X \in B|Y)$, pak $P(X \in B) = E[P(X \in B)]$.

$$Ef(X) = E1_B(X) = P(X \in B),$$

$$E\left[E\left[1_B(X)|Y\right]\right] = E\int_B \; dP_{X|Y}(x) = EP_{X|Y}(B) = E[P(X \in B|Y)],$$

tedy $E\left[E\left[f(X)|Y\right]\right] = Ef(X)$.
Nechť nyní f je jednoduchá, tedy $f = \sum_{k=1}^n c_k 1_{B_k}$, $c_k \in \mathbb{R}$, $B_k \in \mathcal{S}$, pak

$$\begin{aligned} Ef(X) = E\sum_{k=1}^n c_k 1_{B_k}(X) = \sum_{k=1}^n c_k E1_{B_k}(X) = \sum_{k=1}^n c_k E[E[1_{B_k}(X)|Y]] = \\ = E\left[E\left[\sum_{k=1}^n c_k 1_{B_k}(X)|Y\right]\right] = E\left[E\left[f(X)|Y\right]\right], \end{aligned}$$

tedy $E\left[E\left[f(X)|Y\right]\right] = Ef(X)$.
Pro f nezápornou měřitelnou pak existuje posloupnost $\{f_n\}$ jednoduchých funkcí, takových, že $f_n \nearrow f$, $n \to \infty$. Pak z Leviho věty o monotónní konvergenci plyne:

$$Ef_n(X) = \int_S f_n(x) \; dP_X(x) \xrightarrow{n\to\infty} \int_S f(x) \; dP_X(x) = Ef(X),$$

$$E[f_n(X)|Y] = \int_S f_n(x) \; dP_{X|Y}(x) \xrightarrow{n\to\infty} \int_S f(x) \; dP_{X|Y}(x) = E[f(X)|Y],$$

tedy také platí, že $E\left[E\left[f_n(X)|Y\right]\right] \to E\left[E\left[f(X)|Y\right]\right], n \to \infty$.
Víme, že platí $Ef_n(X) = E\left[E\left[f_n(X)|Y\right]\right]$, pak nutně také musí platit, že

$Ef(X) = E\left[E\left[f(X)|Y\right]\right]$.
Nakonec uvažujme f obecnou měřitelnou, pak ji lze rozložit na $f = f^+ - f^-$, kde f^+ a f^- jsou nezáporné, měřitelné, navíc $Ef(X) = Ef^+(X) - Ef^-(X)$ a $E[f(X)|Y] = E[f^+(X)|Y] - E[f^-(X)|Y]$.
Tedy získáváme $E\left[E\left[f(X)|Y\right]\right] = Ef(X)$. Speciálně pro $f(x) = x$ máme: $E\left[E\left[X|Y\right]\right] = EX$.

□

Věta 2.9

Nechť X_n jsou reálné náhodné veličiny, které splňují $0 \leq X_1 \leq X_2 \leq \ldots$, skoro jistě, nechť $X_n \to X$, $n \to \infty$, skoro jistě, $n \in \mathbb{N}$. Nechť $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ je náhodná veličina. Potom $E[X_n|Y = y] \to E[X|Y = y]$, $n \to \infty$.

Důkaz

Označme náhodnou veličinu $T = (X_n)$, pak $T : (\Omega, \mathcal{A}) \to (\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. Protože prostor $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ je polský, víme že existuje $P_{T|Y=y}$. Definujeme funkce $f_n : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, $f_n(t) = t_n$, $t \in \mathbb{R}^{\mathbb{N}}$, a funkci $f := \lim_{n\to\infty} f_n$. Pak platí, že $f_n(T) = X_n$, $f_n(T) \to f(T) = X$.

Z Leviho věty a věty o transformaci dostáváme:

$$E[X_n|Y = y] = \int_S f_n(t)\, dP_{T|Y=y}(t) \xrightarrow{n\to\infty} \int_S f(t)\, dP_{T|Y=y}(t) = E[X|Y = y].$$

□

Věta 2.10

Nechť X_n, $n \in \mathbb{N}$, Z jsou reálné náhodné veličiny, $|X_n| \leq Z$ skoro jistě, $\forall n \in \mathbb{N}$, $Z \in L_1(\Omega, \mathcal{A}, P)$, $X_n \to X$, $n \to \infty$, skoro jistě. Pak $E[X_n|Y = y] \to E[X|Y = y]$, $n \to \infty$.

Důkaz

Postup důkazu je naprosto stejný jako v předchozí větě, jen místo Leviho věty použijeme Lebesgueovu věty a dostáváme, že

$$E[X_n|Y = y] = \int_S f_n(t)\, dP_{T|Y=y}(t) \xrightarrow{n\to\infty} \int_S f(t)\, dP_{T|Y=y}(t) = E[X|Y = y].$$

□

Věta 2.11 (Jensenova nerovnost)

Nechť $D \subseteq \mathbb{R}^n$, $n \in \mathbb{N}$, nechť $X : (\Omega, \mathcal{A}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ je reálný náhodný vektor a $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ je náhodná veličina. Nechť $X(\Omega) \subseteq D$. Pak platí:

- $E[X|Y] = (E[X_1|Y], E[X_2|Y], \ldots, E[X_n|Y]) \in D$ skoro jistě.
- Pro $f : D \to \mathbb{R}$ konvexní, měřitelnou platí: $E[f(X)|Y] \geq f(E[X|Y])$.

Důkaz

$$1 = P(X \in D) = E(P(X \in D|Y)) = P_{X|Y}(D),$$

tedy $P_{X|Y}(D) = 1$, takže $E[X|Y] \in D$ skoro jistě.

Z nepodmíněné Jensenovy nerovnosti víme, že pro f měřitelnou, konvexní platí:

$$Ef(X) = \int_S f(x)\ dP_X(x) \geq f\left(\int_S x\ dP_X(x)\right) = f(EX).$$

Tedy také z definice podmíněného rozdělení platí:

$$E[f(X)|Y] = \int_S f(x)\ dP_{X|Y}(x) \geq f\left(\int_S x\ dP_{X|Y}(x)\right) = f(E[X|Y]).$$

$\square$

Věta 2.12

Nechť X, Y jsou reálné náhodné veličiny, pro něž platí, že Y má hustotu f_Y a $X \in L_1(\Omega, \mathcal{A}, P)$, pak existuje borelovská integrovatelná funkce $h_{X|Y}$ taková, že

- $$H_{X|Y}(y) := E[X1_{[Y<y]}] = \int_{[Y<y]} X\ dP = \int_{-\infty}^{y} h_{X|Y}(z)\ dz,$$
- $$E[X|Y=y] = \frac{h_{X|Y}(y)}{f_Y(y)}, \quad P_Y - s.v., \quad y \in \mathbb{R}.$$

Pak lze přímo počítat pro P_Y-skoro všechna $y \in \mathbb{R}$:

$$E[X|Y=y] = \frac{\frac{d}{dy}E[X1_{[Y<y]}]}{\frac{d}{dy}P(Y<y)}.$$

Důkaz

Definujeme $h(y) := E[X|Y=y]$, pak platí:

$$\begin{aligned} H_{X|Y}(y) &= \int_{[Y<y]} X\ dP = \int_{[Y<y]} h(Y)\ dP = \int_{(-\infty,y)} h(y)\ dP_Y(y) = \\ &= \int_{-\infty}^{y} h(y) f_Y(y)\ dy. \end{aligned}$$

Platí:

$$\int_{-\infty}^{\infty} |h(y)f_Y(y)|\ dy = \int |h(Y)|\ dP = E[E[X|Y]] \leq E|X| < \infty.$$

Tedy funkce $h(y)f_Y(y)$ je integrovatelná, pak dostáváme, že funkce $H_{X|Y}$ je absolutně spojitá. Má tedy skoro všude derivaci a ta je skoro všude rovna $h(y)f_Y(y)$. Protože $f_Y(y) > 0$, P_Y-skoro všude, dostaneme, že P_Y-skoro všude platí:

$$E[X|Y=y] = h(y) = \frac{H'_{X|Y}(y)}{f_Y(y)} = \frac{\frac{d}{dy}E[X1_{[Y<y]}]}{\frac{d}{dy}P(Y<y)},$$

neboť $F_Y(y)$ je absolutně spojitá funkce, a tedy $F'_Y(y) = f_Y(y)$ platí pro skoro všechna $y \in \mathbb{R}$, tedy i pro P_Y-skoro všechna $y \in \mathbb{R}$.

$\square$

Poznámka

Pro Radon-Nikodymovu derivaci λ vzhledem k μ používáme občas diferenciálový zápis

$$h = \frac{d\lambda}{d\mu},$$

což lze převést do součinového tvaru: $d\lambda = hd\mu$. Předpokládejme, že X má absolutně spojité rozdělení a $f \in L^1(P_X)$, pak také funkce

$$y \mapsto E[f(X); Y \leq y]$$

je absolutně spojitá a platí

$$\frac{d}{dy} E[f(X); Y \leq y] = E[f(X)|Y = y] \frac{d}{dy} P(Y \leq y),$$

cože lze psát také jako

$$dE[f(X); Y \leq y] = E[f(X)|Y = y] \; dP(Y \leq y).$$

Pokud posuneme diferenciál k proměnné, získáme ještě přehlednější zápis

$$E[f(X); Y \in dy] = E[f(X)|Y = y] P(Y \in dy).$$

Nechť nyní Y má absolutně spojité rozdělení s hustotou f_Y, pak

$$P(Y \in dy) = f_Y(y) \; dy.$$

Předpokládejme dále, že $Y = |X|$, nechť F je distribuční funkce veličiny X. Pak využitím Věty 2.12 máme, že

$$\begin{aligned} E[f(X)|Y = y] &= \frac{E[g(X); Y \in dy]}{P(Y \in dy)} = \frac{E[g(X); |X| \in dy]}{P(|X| \in dy)} = \\ &= \frac{E[g(X); X \in dy] + E[g(X); -X \in dy]}{P(X \in dy) + P(-X \in dy)} = \end{aligned}$$

$$= E[g(X)|X = y] \frac{dP_X}{dP_X + dP_{-X}}(y) + E[g(X)|X = -y] \frac{dP_{-X}}{dP_X + dP_{-X}}(y).$$

Příklad 1

Nechť X je náhodná veličina s hustotou

$$f_X(x) = \begin{cases} \frac{1}{2}(x+1), & x \in (-1, 1) \\ 0, & x \notin (-1, 1). \end{cases}$$

Spočítejte $E[X||X| = y]$ pro $y \in (-1, 1)$.

Řešení: Použijeme vzorec z předchozí Věty 2.12, kde $Y = |X|$.

$$\begin{aligned} E[X||X| = y] &= \frac{\frac{d}{dy} E[X 1_{[|X|<y]}]}{\frac{d}{dy} P(|X| < y)} = \frac{\frac{d}{dy}\left(\int_{-y}^{y} x \frac{1}{2}(x+1) \; dx\right)}{\frac{d}{dy}\left(\int_{-y}^{y} \frac{1}{2}(x+1) \; dx\right)} = \\ &= \frac{\frac{d}{dy}\left(\frac{1}{2}\left[\frac{x^3}{3}\right]_{-y}^{y}\right)}{\frac{d}{dy}\left(\frac{1}{2}[x]_{-y}^{y}\right)} = \frac{\frac{d}{dy}\left(\frac{y^3}{3}\right)}{\frac{d}{dy} y} = y^2, \quad y \in (-1, 1). \end{aligned}$$

Příklad 2

Nechť X a Y jsou nezávislé náhodné veličiny s rovnoměrným rozdělením na $(0, 2\pi)$. Spočítejte $E[\sin(X+Y)|X]$.

Řešení: V následujících úpravách využijeme vlastnosti c) a d) z Věty 2.8.

$$\begin{aligned} E[\sin(X+Y)|X] &= E[\sin X \ \cos Y + \cos X \ \sin Y|X] = \\ &= \sin X \ E[\cos Y|X] + \cos X \ E[\sin Y|X] = \\ &= \sin X \ E(\cos Y) + \cos X \ E(\sin Y). \end{aligned}$$

Spočítáme klasické střední hodnoty

$$E(\cos Y) = \int_0^{2\pi} \frac{1}{2\pi} \cos y \ dy = \frac{1}{2\pi} [\sin y]_0^{2\pi} = 0,$$

$$E(\sin Y) = \int_0^{2\pi} \frac{1}{2\pi} \sin y \ dy = \frac{1}{2\pi} [-\cos y]_0^{2\pi} = \frac{1}{2\pi}(-1+1) = 0,$$

Tedy celkem máme $E[\sin(X+Y)|X] = 0$.

Příklad 3

Nechť X má rovnoměrné rozdělení na $(0,1)$ a Y má Poissonovo rozdělení $Po(1)$. Spočítejte $E[|X-Y||Y^3]$.

Řešení: Protože ve chvíli, kdy známe hodnotu Y^3, je už jednoznačně dána hodnota Y, můžeme podmíněnou střední hodnotu přepsat:

$$E[|X-Y||Y^3] = E[|X-Y||Y].$$

Dále upravíme

$$E[|X-Y||Y] = \frac{E[|X-k|; Y=k]}{P(Y=k)} = E(|X-k|),$$

pro $k = 0$ máme:

$$E(|X-k|) = E|X| = EX = \frac{1}{2},$$

pro $k \geq 1$ máme:

$$E(|X-k|) = E(k-X) = k - EX = k - \frac{1}{2}.$$

Celkem tedy dostáváme výsledek

$$E[|X-Y||Y=k] = |k - \frac{1}{2}| \quad \Rightarrow \quad E[|X-Y||Y] = |Y - \frac{1}{2}|.$$

Příklad 4

Nechť G je schodovitá Cantorova funkce, nechť V je náhodná veličina s distribuční funkcí G. Spočítejte $E[h(V)||V-\frac{1}{2}|]$, kde $h \in L^1(P_V)$.

Řešení: Víme, že Cantorova funkce je spojitá, není ale absolutně spojitá, nelze ji tedy derivovat. Definujeme náhodnou veličinu $X := V - \frac{1}{2}$. Pak pro její distribuční funkci F platí

$$F(x) = P(X \leq x) = P(V - \frac{1}{2} \leq x) = P(V \leq x + \frac{1}{2}) = G(x + \frac{1}{2}).$$

Snadno lze ověřit, že $F(x) + F(-x) = 1$, tedy X má symetrické spojité rozdělení, takže $P_X = P_{-X}$. Pro P_X-skoro všechna x platí

$$\frac{dP_X}{d(P_X + P_{-X})} = \frac{dP_{-X}}{d(P_X + P_{-X})} = \frac{P(X \in dx)}{P(X \in dx) + P(-X \in dx)} = \frac{1}{2}.$$

Tedy z Poznámky před příklady dostáváme výsledek

$$E\left[h(V)|\left|V - \frac{1}{2}\right|\right] = E\left[h\left(X + \frac{1}{2}\right)||X|\right] = \frac{h\left(x + \frac{1}{2}\right) + h\left(\frac{1}{2} - x\right)}{2}.$$

Příklad 5

Nechť X má rovnoměrné rozdělení na $(-1, 2)$, spočítejte $E[X||X|]$.

Řešení: Výpočet bude za použití Poznámky před příklady velice rychlý. Protože platí, že

$$\frac{P(X \in dx)}{P(|X| \in dx)} = \frac{f_X(x)}{f_X(x) + f_X(-x)} \; 1_{[1,2)}(x) + \frac{1}{2} \; 1_{(-1,1)}(x),$$

pak také

$$E[X||X|] = |X| \; 1_{[1,2)}(|X|).$$

2.3 Podmíněný rozptyl, kovariance

Podobně jako podmíněnou střední hodnotu i podmíněný rozptyl a podmíněnou kovarianci budeme definovat analogicky nepodmíněnému případu. V této práci se ale pro jednoduchost omezíme pouze na případ reálných náhodných vektorů.

Definice 2.6

Nechť $X, Z : (\Omega, \mathcal{A}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ jsou reálné náhodné vektory a nechť $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ je náhodná veličina. Pak *podmíněný rozptyl* X při Y definujeme jako

$$var(X|Y) = E[(X - E[X|Y])^{\otimes 2}|Y] = E[(X - E[X|Y])(X - E[X|Y])^T|Y].$$

Podmíněnou kovarianci X a Z při Y definujeme jako

$$cov(X, Z|Y) = E[(X - E[X|Y])(Z - E[Z|Y])^T|Y].$$

Věta 2.13

Nechť $X : (\Omega, \mathcal{A}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ je náhodný vektor, $Y \in \mathbb{L}(\mathcal{A}, \mathcal{H})$ je náhodná veličina. Je-li $X \in L_2(\Omega, \mathcal{A}, P)$, pak platí:

$$var(X|Y) = E(var(X|Y)) + var(E[X|Y])$$

Důkaz

Použijeme následující trik:

$$varX = E(X - EX)^{\otimes 2} = E(\underbrace{(X - E[X|Y])}_{Q_1} + \underbrace{(E[X|Y] - EX)}_{Q_2})^{\otimes 2} = EQ_1^{\otimes 2} +$$
$$+ EQ_2^{\otimes 2} + EQ_1Q_2^T + EQ_2Q_1^T,$$

pak $EQ_1^{\otimes 2} = E(var(X|Y))$ a $EQ_2^{\otimes 2} = var(E[X|Y])$.
Tedy zbývá dokázat, že zbývající členy jsou nulové:

$$EQ_1Q_2^T = E[E[Q_1Q_2^T|Y]] = E[E[(X - E[X|Y])(E[X|Y] - EX)^T|Y]] =$$
$$= E[\underbrace{E[X - E[X|Y]|Y]}_{0}(E[X|Y] - EX)^T] = 0.$$

Druhý člen vyřešíme analogicky.

□

2.4 Spojité podmiňování

Nyní se znovu vrátíme k podmíněnému rozdělení. Víme, že v případě absolutně spojitých náhodných veličin je nabývání jedné hodnoty jev s nulovou pravděpodobností. Proto není příliš vhodné používat zápis $P(X \in A|Y = y)$ ve chvíli, kdy $P(Y = y) = 0$. Zamyslíme se tedy nad interpretací takového podmíněného rozdělení. Intuitivně můžeme chápat takovou podmínku jako

$$\lim_{\epsilon \to 0+} P(X \in A|Y \in (y - \epsilon, y + \epsilon)).$$

Zavedeme si pro tento případ označení $P(X \in A|Y \in dy)$.

Ve Větě 2.2 jsme dokázali jednoznačnost podmíněného rozdělení s ohledem na míru P_Y. Pokud ale platí, že $P(Y = y) = 0$, neříká výše uvedená věta o jednoznačnosti $P_{X|Y \in dy}$ nic. Tuto skutečnost tedy rozebereme v následující větě.

Tvrzení 2.14

Buď (S, d) separabilní úplný metrický prostor, pak existuje spočetný systém spojitých omezených funkcí na (S, d) určujících borelovskou pravděpodobnost.

Důkaz

Buď $S' \subseteq S$ hustá, spočetná podmnožina. Pak systém

$$\mathcal{B}\{B(s, \epsilon), s \in S', 0 < \epsilon \in \mathbb{Q}\}$$

určuje pravděpodobnostní míru na $\mathcal{B}(S)$, neboť každou otevřenou množinu G v S lze zapsat ve tvaru

$$G = \bigcup_{G \supseteq B \in \mathcal{B}} B,$$

přičemž otevřené množiny tvoří systém uzavřený na konečné průniky generující borelovskou σ-algebru $\mathcal{B}(S)$. Nyní položíme

$$\mathcal{F}\{f_{s,\epsilon,n}; n \in \mathbb{N}, s \in S', 0 < \epsilon \in \mathbb{Q}\}, \quad f_{s,\epsilon,n}(x) = (n(\epsilon - d(x,s))_+) \wedge 1.$$

Funkce $f \in \mathcal{F}$ jsou spojité s hodnotami v $[0,1]$ a platí, že

$$0 \leq f_{s,\epsilon,n} \nearrow 1_{B(s,\epsilon)}, \quad n \to \infty.$$

Z Leviho věty o monotónní konvergenci pak dostáváme tvrzení.

□

Věta 2.15

Nechť $X = \{X_t, t \in [0,1]\}$ je spojitý náhodný proces, $Y \in \mathbb{L}(\mathcal{A})$ je absolutně spojitá reálná náhodná veličina. Označme

$$supp(Y) = \{y \in \mathbb{R} : \exists \epsilon > 0, P(|Y - y| < \epsilon) > 0\}.$$

Nechť $y_0 \in supp(Y)$. Nechť pro $y \in \mathbb{R}$ existuje $P_{X|Y \in dy}$, takové, že pro $y_n \to y_0$

$$P_{X|Y \in dy_n} \xrightarrow{w} P_{X|Y \in dy_0},$$

Pak je $P_{X|Y \in dy_0}$ určeno jednoznačně.

Důkaz

Tvrzení dokážeme sporem. Nechť $P_{X|Y \in dy}$ a $P'_{X|Y \in dy}$, $y \in \mathbb{R}$, jsou dvě podmíněná rozdělení splňující předpoklady věty. Předpokládejme, že

$$P_{X|Y \in dy_0} \neq P'_{X|Y \in dy_0}.$$

Víme, že $P_{X|Y \in dy}$ a $P'_{X|Y \in dy}$ jsou spojitá v distribuci v bodě y_0, tedy existuje $\epsilon > 0$ takové, že

$$P_{X|Y \in dy} \neq P'_{X|Y \in dy}, \quad y \in (y_0 - \epsilon, y_0 + \epsilon).$$

Označíme jev $A = [|Y - y_0| < \epsilon]$, pak platí, že $P(A) > 0$. Nechť $\{f_n, n \in \mathbb{N}\}$ je spočetný systém spojitých omezených funkcí, určující míru na $\mathbb{C}([0,1])$. Existenci tohoto systému jsme získali z Tvrzení 2.14, neboť prostor $\mathbb{C}([0,1])$ je separabilní a úplný. Pak $A = \bigcup_{n=1}^{\infty} A_n$, kde

$$A_n = \left\{ \omega \in A : \int f_n \, dP_{X|Y \in dy}(\omega) \neq \int f_n \, dP'_{X|Y \in dy}(\omega) \right\},$$

tedy existuje $n \in \mathbb{N}$, že $P(A_n) > 0$. Současně platí, že

$$\int f_n \, dP_{X|Y} \overset{s.j.}{=} E[f_n(X)|Y] \overset{s.j.}{=} \int f_n \, dP'_{X|Y}.$$

Dostáváme tedy spor s tím, že

$$\forall \omega \in A_n, \quad P(A_n > 0) \quad : \quad \int f_n \, dP_{X|Y}(\omega) \neq \int f_n \, dP'_{X|Y}(\omega).$$

□

Příklad (Borel-Kolmogorovův paradox)

Nechť $(X,Y,Z) \sim R\{x^2+y^2+z^2=1\}$. Chceme najít podmíněné rozdělení (X,Y,Z) za podmínky $X=0$ (poledník), $Y=0$ (poledník) a $Z=0$ (rovník). Jak za chvíli uvidíme, tak různé postupy nám dají různé odpovědi.

1) První myšlenka je podmiňovat jevem $|Z|<\epsilon$ a poslat $\epsilon \to 0+$, tím v limitě získáme kandidáta pro rozdělení za podmínky $Z=0$, tedy

$$(X,Y,Z)|Z=0 \quad \sim \quad R\{x^2+y^2=1\}.$$

2) Druhá odpověď vychází z interpretace kružnice $[X=0]$ či $[Y=0]$ jako poledníku. Převedeme veličiny X,Y,Z do sférických souřadnic, tedy

$$\begin{aligned} X &= \cos\Phi\cos\Psi, \quad && \Phi \in [-\pi,\pi], \\ Y &= \sin\Phi\cos\Psi, \quad && \Psi \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right], \\ Z &= \sin\Psi, \end{aligned}$$

budeme podmiňovat jevem $\left||\Phi|-\frac{\pi}{2}\right|<\epsilon$ a opět pošleme $\epsilon \to 0+$. V tomto případě také existuje limita

$$(X,Y,Z)\Big|\left||\Phi|-\frac{\pi}{2}\right|<\epsilon \xrightarrow{d} \mu,$$

ale $\mu \neq R\{y^2+z^2=1\}$, konkrétně body kolem rovníku $[Z=0]$ mají větší váhu než v okolí pólů $[X=0,Y=0]$.

Vidíme tedy, že dostáváme různá podmíněná rozdělení v závislosti na tom, zda kružnici obepínající sféru interpretujeme jako rovník nebo jako poledník.

Nyní se omezíme na dvourozměrný případ. Nechť $(X,Y) \sim R\{x^2+y^2<1\}$, tedy (X,Y) je vektor s hustotou

$$f_{(X,Y)} = \frac{1}{\pi}\, 1_{(x^2+y^2<1)}.$$

Řekněme, že nás zajímá podmíněné rozdělení $P_{(X,Y)|X=0}$. Jenže $P(X=0)=0$. Můžeme spočítat slabou limitu $P_{(X,Y)||X|<\epsilon}$ pro $\epsilon \to 0+$. Příkladem podmíněného rozdělení za podmínky $X=0$ je

$$P_{(X,Y)|X=x} = N(x,0) \otimes R(-\sqrt{1-x^2},\sqrt{1-x^2}), \quad x \in (-1,1).$$

Protože tento systém je spojitý vzhledem ke slabé konvergenci v $x \in (-1,1)$, dostáváme v rámci spojitého podmiňování pro $x=0$:

$$P_{(X,Y)|X\in dx} = N(0,0) \otimes R(-1,1).$$

Převedeme veličiny X,Y na polární souřadnice, tedy

$$\begin{aligned} X &= R\cos\Phi, \\ Y &= R\sin\Phi. \end{aligned}$$

Platí, že $[X=0] = \left[|\Phi| = \frac{\pi}{2}\right]$. Nyní spočítáme pro $\epsilon \to 0+$ slabou limitu

$$P_{(X,Y)|\left||\Phi|-\frac{\pi}{2}\right|<\epsilon}.$$

Protože

$$\left[\left||\Phi| - \frac{\pi}{2}\right| < \epsilon\right] = \left[|Y| > |X| \tan\left(\frac{\pi}{2} - \epsilon\right)\right],$$

pak pro malá ϵ máme ve slabé konvergenci

$$\lim_{\epsilon \to 0+} P_{(X,Y)|\left||\Phi| - \frac{\pi}{2}\right| < \epsilon} = \lim_{K \to \infty} R\{1 > |y| > |x|K\} = N(0,0) \otimes SB(2,1),$$

kde rozdělení $SB(2,1)$ vznikne z $B(2,1)$ symetrizací. Zde bychom mohli podobně jako v předchozí úvaze najít verzi $P_{(X,Y)||\Phi|=\varphi}$, která je slabě spojitá v bodě $\varphi = \frac{\pi}{2}$. Pak bychom dostali, že pro $\varphi = \frac{\pi}{2}$ je

$$P_{(X,Y)||\Phi| \in d\varphi} = N(0,0) \otimes SB(2,1).$$

Vidíme tedy, že

$$P_{(X,Y)|X \in dx} = N(0,0) \otimes R(-1,1) \quad \neq \quad N(0,0) \otimes SB(2,1) = P_{(X,Y)||\Phi| \in d\varphi},$$

přestože $[X = x] = [|\Phi| = \varphi]$ pro $x = 0$ a $\varphi = \frac{\pi}{2}$.

Kapitola 3

Wienerův proces a Brownův most

V této kapitole se seznámíme s Wienerovým procesem, konstrukcí Brownova mostu a následným využitím ve statistice.

Definice 3.1

Wienerův proces $\{W_t, t \geq 0\}$, někdy také nazývaný *standartní Brownův pohyb*, je gaussovský náhodný proces se spojitými trajektoriemi a nezávislými přírůstky, který splňuje, že $W_t - W_s \sim N(0, |t-s|)$ a $W_0 = 0$.

Tvrzení 3.1

Wienerův proces je centrovaný a pro jeho kovarianční funkci platí, že

$$cov(W_s, W_t) = s \wedge t.$$

Důkaz

Centrovanost snadno ověříme jednoduchým výpočtem.

$$EW_t = E(W_t - W_0) = 0, \quad t \geq 0.$$

Protože přírůstky Wienerova procesu jsou nezávislé, jsou také nekorelované.

$$\begin{aligned} s \leq t: \quad cov(W_s, W_t) &= cov(W_s - W_0, W_t - W_s + W_s - W_0 = \\ &= cov(W_s - W_0, W_t - W_s) + var(W_s - W_0) = s \end{aligned}$$

tedy $cov(W_s, W_t) = s \wedge t$.

□

Příklad

Definujeme pomocí Wienerova procesu proces $B^0 = \{B_t^0, t \in [0,1]\}$ vztahem

$$B_t^0 = W_t - tW_1.$$

Pak zřejmě proces B^0 má spojité trajektorie a je gaussovský, protože

$$\begin{pmatrix} B_{t_1}^0 \\ B_{t_2}^0 \\ \vdots \\ B_{t_n}^0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & -t_1 \\ 0 & 1 & 0 & \dots & -t_2 \\ \vdots & & & & \vdots \\ 0 & \dots & 0 & 1 & -t_n \end{pmatrix} \begin{pmatrix} W_{t_1} \\ W_{t_2} \\ \vdots \\ W_{t_n} \\ W_1 \end{pmatrix}, \quad t_1, \dots, t_n \in [0,1],$$

Spočítáme některé charakteristiky:

$$B_0^0 = B_1^0 = 0,$$

$$EB_t^0 = E(W_t - tW_1) = 0 - 0 = 0,$$

$$0 \leq t \leq s \leq 1: \ cov(B_s^0, B_t^0) = EB_s^0 B_t^0 = E(W_s - sW_1)(W_t - tW_1) =$$
$$= EW_sW_t - sEW_1W_t - tEW_sW_1 + stEW_1^2 = t - st - ts + st = t - st,$$

tedy $cov(B_s^0, B_t^0) = s \wedge t - st$.

Definice 3.2

Proces $B^0 = \{B_t^0, t \in [0,1]\}$ nazveme *Brownovým mostem*, pokud je centrovaný, gaussovský, má spojité trajektorie a platí pro něj, že

- $B_0^0 = B_1^0 = 0$,
- $cov(B_s^0, B_t^0) = s \wedge t - st$.

Tvrzení 3.2

Nechť $(\mathcal{C}([0,1]), \mathcal{B}(\mathcal{C}([0,1]))$ je prostor spojitých funkcí. Označme π_t projekce na tomto prostoru, $f(t) = \pi_t(f)$, $f \in \mathcal{C}([0,1])$. Potom platí, že

$$\mathcal{B}(\mathcal{C}([0,1]) = \sigma\{\pi_t, t \in [0,1]\}.$$

Důkaz

Protože projekce π_t jsou spojité, pak $\sigma\{\pi_t, t \in [0,1]\} \subseteq \mathcal{B}(\mathcal{C}([0,1]))$. Musíme dokázat i opačnou inkluzi. Protože $\mathbb{Q}$ je hustá v $\mathbb{R}$, platí pro metriku

$$\rho(f,g) = \sup_{s \in [0,1]} |f(s) - g(s)| = \sup_{s \in [0,1] \cap \mathbb{Q}} |f(s) - g(s)|.$$

Definujeme otevřenou kouli

$$\begin{aligned} B(f,\epsilon) &= \{g \in \mathcal{C}([0,1]), \rho(f,g) < \epsilon\} = \\ &= \{g \in \mathcal{C}([0,1]), \forall s \in [0,1] \cap \mathbb{Q} : |f(s) - \pi_s(g)| < \epsilon\}. \end{aligned}$$

Navíc

$$\{g \in \mathcal{C}([0,1]), |f(s) - \pi_s(g)| < \epsilon\} \in \sigma\{\pi_s\} \subseteq \sigma\{\pi_t, t \in [0,1]\}.$$

Tedy $\mathcal{B}(\mathcal{C}([0,1])) \subseteq \sigma\{\pi_t, t \in [0,1]\}$, celkem

$$\mathcal{B}(\mathcal{C}([0,1]) = \sigma\{\pi_t, t \in [0,1]\}.$$

□

Poznámka

Buďte $X = \{X_t, t \in [0,1]\}$ a $Y = \{Y_t, t \in [0,1]\}$ spojité procesy. Budeme psát $X \sim Y$, pokud mají stejné rozdělení na $(\mathbb{R}, (\mathcal{B}(\mathbb{R}))^{[0,1]}$, resp. na $(\mathbb{C}([0,1]), \mathcal{B}(\mathbb{C}([0,1])))$, což v obou případech vychází nastejno s tím, že procesy X a Y mají stejná konečně rozměrná rozdělení, tedy, že

$$(X_{t_1}, \ldots, X_{t_n}) \sim (Y_{t_1}, \ldots, Y_{t_n}), \quad t \in [0,1]^n, \quad n \in \mathbb{N}. \tag{5}$$

Důkaz Pokud $P(X \in B) = P(Y \in B)$ pro každou $B \in \mathcal{B}(\mathbb{R})^{[0,1]}$ nebo pro každou $B \in \mathcal{B}(\mathbb{C}([0,1]))$, pak volbou

$$B = [(\pi_{t_1}, \dots, \pi_{t_n}) \in C], \quad C \in \mathcal{B}(\mathbb{R}^n), \tag{6}$$

okamžitě dostaneme rovnost

$$P((X_{t_1}, \dots, X_{t_n}) \in C) = P((Y_{t_1}, \dots, X_{t_n}) \in C).$$

Platí-li naopak (5), pak míry P_X a P_Y se rovnají na množinách typu (6) v $\mathbb{R}^{[0,1]}$ i v $\mathbb{C}([0,1])$. Systém takovýchto množin tvoří algebru generující. Ať již σ-algebru $\mathcal{B}(\mathbb{R})^{[0,1]}$ nebo $\mathcal{B}(\mathbb{C}([0,1]))$, což nám spolu s Dynkinovým argumentem dává rovnost $P_X = P_Y$ na celém σ-obalu.

□

Příklad

Nechť Z je náhodná veličina s normovaným normálním rozdělením nezávislá s B^0. Definujeme proces $V = \{V_t, t \in [0,1]\}$ následujícím způsobem

$$V_t = B_t^0 + tZ.$$

Vidíme, že proces V má spojité trajektorie. Je gaussovský, protože

$$\begin{pmatrix} V_{t_1} \\ V_{t_2} \\ \vdots \\ V_{t_n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & t_1 \\ 0 & 1 & 0 & \dots & t_2 \\ \vdots & & & & \vdots \\ 0 & \dots & 0 & 1 & t_n \end{pmatrix} \begin{pmatrix} B_{t_1}^0 \\ B_{t_2}^0 \\ \vdots \\ B_{t_n}^0 \\ Z \end{pmatrix}, \quad t_1, \dots, t_n \in [0,1],$$

a platí pro něj, že

$$EV_t = 0, \quad V_0 = 0,$$

$$\begin{aligned} cov(V_s, V_t) = cov(B_s^0 + sZ, B_t^0 + tZ) = cov(B_s^0, B_t^0) + s\ cov(Z, B_t^0) + \\ + t\ cov(B_s^0, Z) + st = s \wedge t - st + st = s \wedge t. \end{aligned}$$

Tedy vidíme, že $\{V_t, t \in [0,1]\} \sim \{W_t, t \in [0,1]\}$.

Poznámka

Obecněji lze pomocí Brownova mostu definovat Wienerův proces na $[0,T]$, opět $Z \sim N(0,1)$, nezávislá s B^0:

$$W_t = B_{\frac{t}{T}}^0 + \frac{t}{\sqrt{T}} Z, \quad t \in [0,T].$$

Tvrzení 3.3

W_1 a B^0 jsou nezávislé.

Důkaz

Spočítáme kovarianci:

$$cov(W_1, B_t^0) = cov(W_1, W_t - tW_1) = cov(W_1, W_t) - t\, cov(W_1, W_1) = 1 \wedge t - t = 0,$$

navíc náhodný vektor $(W_1, B_{t_1}^0, \ldots, B_{t_n}^0)$ je gaussovský, tedy veličina W_1 je nezávislá s $(B_{t_1}^0, \ldots, B_{t_n}^0)$, pro každé $t_1, \ldots, t_n \in [0,1]$. Tedy platí, že

$$W_1 \perp\!\!\!\perp \sigma\Big(\bigcup_{t \in [0,1]^n} \sigma(B_{t_1}^0, \ldots, B_{t_n}^0)\Big) \quad \Rightarrow \quad W_1 \perp\!\!\!\perp \{B_t^0, 0 \le t \le 1\} \quad \Rightarrow \quad W_1 \perp\!\!\!\perp B^0.$$

$\square$

Tvrzení 3.4

Buď $g \in \mathbb{C}_b(\mathbb{R})$ a nechť $X \in L(\Omega, \mathcal{A}, P)$ a $x_0 \in \mathbb{R}$ jsou takové, že $P(|X - x_0| < \epsilon) > 0$ kdykoliv $\epsilon > 0$. Pak

$$E[g(X)||X - x_0| < \epsilon] \xrightarrow{\epsilon \to 0} g(x_0).$$

Důkaz

Buď $\epsilon > 0$ dané, najdeme $\delta > 0$ takové, že pro $|x - x_0| \le \delta$ je $|g(x) - g(x_0| \le \epsilon$. Pak zřejmě

$$|E\left[g(X)||X - x_0| \le \delta\right] - g(x_0)| \le E\left[|g(X) - g(x_0)|||X - x_0| \le \delta\right] \le \epsilon.$$

$\square$

Tvrzení 3.5

Pro podmíněná rozdělení a $\epsilon > 0$ platí:

$$P_{\{W_t, t \in [0,1]\}||W_1| < \epsilon} \quad \xrightarrow{w} \quad P_{B_t^0, t \in [0,1]}.$$

Důkaz

Protože náhodná veličina W_1 má normované normální rozdělení a je nezávislá s B^0, můžeme pomocí ní vyjádřit Wienerův proces jako

$$W_t = B_t^0 + tW_1, \quad t \in [0,1].$$

Buď nyní $f \in \mathbb{C}_b(\mathbb{C}([0,1]))$, pak z Lebesgueovy věty o majorantě víme, že

$$g(w) := Ef(B_t^0 + wt, t \in [0,1])$$

je spojitá omezená funkce a dle Tvrzení 3.4 platí

$$E[g(W_1)||W_1| < \epsilon] \to g(0) = Ef(B_t^0, t \in [0,1]).$$

Protože W_1 je nezávislá s Brownovým mostem $\{B_t^0, t \in [0,1]\}$, lze levou stranu upravit následujícím způsobem

$$\frac{1}{P(|W_1| < \epsilon} \int_\epsilon^\epsilon g(w)\, dP_W(w) = \frac{1}{P(|W_1| < \epsilon} \int_\epsilon^\epsilon Ef(B_t^0 + wt, t \in [0,1])\, dP_W(w) =$$

$$
\begin{aligned}
&= \frac{1}{P(|W_1| < \epsilon} \int_{\epsilon}^{\epsilon} E\left[f(B_t^0 + W_1 t, t \in [0,1])|W_1 = w\right] \; dP_W(w) = \\
&= E\left[E\left[f(B_t^0 + W_1 t, t \in [0,1])|W_1\right] ||W_1| < \epsilon\right] = \\
&= E\left[f(B_t^0 + W_1 t, t \in [0,1])||W_1| < \epsilon\right].
\end{aligned}
$$

Celkem tedy máme, že pro $f \in \mathbb{C}_b(\mathbb{C}([0,1]))$ platí

$$
\begin{aligned}
\int f \; dP_{W_t, t\in[0,1]||W_1|<\epsilon} &= E\left[f(W_t, t \in [0,1])||W_1| < \epsilon\right] \xrightarrow{\epsilon \to 0} g(0) = \\
&= Ef(B_t^0, t \in [0,1]) = \int f \; dP_{B_t^0, t\in[0,1]},
\end{aligned}
$$

což odpovídá slabé konvergenci

$$
P_{W_t, t\in[0,1]||W_1|<\epsilon} \quad \xrightarrow{w} \quad P_{B_t^0, t\in[0,1]}.
$$

□

Definice 3.3

Pro nezávislá, stejně rozdělená pozorování $X_1, X_2, \ldots, X_n$ s distribuční funkcí F definujeme *empirickou distribuční funkci* vztahem

$$
\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{(X_i \le x)}.
$$

Poznámka

Ze statistiky víme, že podle centrální limitní věty platí

$$
\sqrt{n}\left(\hat{F}_n(x) - F(x)\right) \xrightarrow{d} N(0, F(x)(1 - F(x))).
$$

Tvrzení 3.6

Nechť $X_1, X_2, \ldots, X_n$ jsou nezávislá, stejně rozdělená pozorování s rovnoměrným rozdělením na $(0,1)$ a empirickou distribuční funkcí $\hat{F}_n$. Nechť F je distribuční funkce rovnoměrného rozdělení $R(0,1)$. Nechť B^0 je Brownův most. Definujme proces

$$
\hat{H}_n = \{\hat{H}_n(t), t \in [0,1]\}, \quad \hat{H}_n(t) = \sqrt{n}(\hat{F}_n(t) - F(t)).
$$

Pak platí

$$
\hat{H}_n \quad \xrightarrow{d} \quad B^0.
$$

Důkaz

Pro funkci F platí, že $F(t) = P(X_i \le t) = t$, pak také $F(t)(1-F(t)) = t(1-t)$. Provedeme následující úpravy

$$
\hat{H}_n(t) = \sqrt{n}(\hat{F}_n(t) - F(t)) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} 1_{(X_i \le t)} - t\right),
$$

Z Poznámky víme, že

$$
\hat{H}_n(t) \xrightarrow{d} N(0, t(1-t)),
$$

Nyní spočítáme charakteristiky veličiny $\hat{H}_n(t)$:

$$E\hat{H}_n(t) = 0,$$

$$\begin{aligned}
s \le t: \quad & cov(\hat{H}_n(s), \hat{H}_n(t)) = \\
&= cov\left(\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} 1_{(X_i \le s)} - s\right), \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} 1_{(X_i \le t)} - t\right)\right) = \\
&= cov\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} 1_{(X_i \le s)}, \frac{1}{\sqrt{n}}\sum_{i=1}^{n} 1_{(X_i \le t)}\right) = \frac{1}{n}\sum_{i=1}^{n} cov\left(1_{(X_i \le s)}, 1_{(X_i \le t)}\right) = \\
&= \frac{1}{n}\sum_{i=1}^{n}(P(X_i \le s, X_i \le t) - P(X_i \le s)P(X_i \le t)) = s(1-t),
\end{aligned}$$

tedy $cov(\hat{H}_n(s), \hat{H}_n(t)) = s \wedge t - st$. Takže pro konečně rozměrná rozdělení platí, že

$$\{\hat{H}_n(t), t \in [0,1]\} \xrightarrow{d} \{B_t^0, t \in [0,1]\}.$$

Víme, že existují $\tilde{H}_n$ takové, že

$$\sup_{t \in [0,1]} |\tilde{H}_n(t)| \sim \sup_{t \in [0,1]} |\hat{H}_n(t)|,$$

$$\sup_{t \in [0,1]} |\tilde{H}_n(t)| \xrightarrow{s.j.} \sup_{t \in [0,1]} |\tilde{B}^0(t)|,$$

kde $\tilde{B}^0$ je spojitý reálný proces na $[0,1]$. Pro konečně rozměrná rozdělení tedy je $\tilde{B}^0 \sim B^0$. Protože rozdělení spojitého procesu na $[0,1]$ je určeno jednoznačně systémem konečně rozměrných rozdělení (Tvrzení 3.2), máme, že $\tilde{B}^0 \sim B^0$. BÚNO předpokládejme, že $\tilde{B}^0$ je Brownův most. Platí, že

$$\sup_{t \in [0,1]} |\hat{H}_n(t)| \sim \sup_{t \in [0,1]} |\tilde{H}_n(t)| \xrightarrow{d} \sup_{t \in [0,1]} |\tilde{B}^0(t)| \sim \sup_{t \in [0,1]} |B^0(t)|.$$

Takže platí

$$\hat{H}_n \quad \xrightarrow{d} \quad B^0.$$

□

Poznámka

Pokud pozorování $X_1, X_2, \ldots, X_n$ nejsou rozdělená rovnoměrně na $(0,1)$, ale mají spojitou distribuční funkci G, pak definujeme veličiny $U_1, U_2, \ldots, U_n$ vztahem

$$U_i = G(X_i) \sim R(0,1), \quad i = 1, \ldots, n.$$

Na tyto veličiny už můžeme aplikovat předchozí Tvrzení 3.4. Pro empirickou distribuční funkci pak platí

$$\hat{G}_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{(X_i \le t)} = \frac{1}{n}\sum_{i=1}^{n} 1_{(G^{-1}(U_i) \le t)} = \frac{1}{n}\sum_{i=1}^{n} 1_{(U_i \le G(t))} = \hat{F}_n(G(t)).$$

Závěr

Cílem této bakalářské práce bylo popsat využití Radon-Nikodymovy derivace pravděpodobnosti. Viděli jsme, že pomocí tohoto pojmu lze zavést podmíněného rozdělení, včetně případu spojitého podmiňování jevem nulové praděpodobnosti. Navíc byl v textu nastíněn netradiční přístup k definici podmíněné střední hodnoty, jakožto integrálu z podmíněného rozdělení. Tento přístup je analogický nepodmíněnému případu, což může vést k jeho snadnějšímu pochopení.

Podmíněné rozdělení nalézá své uplatnění také při zkoumání Wienerova procesu a Brownova mostu, využít tyto znalosti je pak možné například ve statistickém Kolmorogově-Smirnovově testu.

Seznam použité literatury

[1] Billingsley, P.: *Convergence of Probability Measures*, John Wiley & Sons, Inc., New York, 1968

[2] Cohn, D.L.: *Measure Theory*, Birkhäuser, Boston, 1980

[3] Dudley, R.M.: *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2004

[4] Kadlec, K.: *Modifikace stochastických objektů*, Diplomová práce, Praha, 2012

[5] Kallenberg, O.: *Foundations of Modern Probability*, Springer, New York, 1997

[6] Rudin, W.: *Real and Complex Analysis*, McGraw-Hill Book Co., Singapore, 1970