Teorie pravděpodobnosti

Studijní text ke kurzu MA750

Ondřej Pokora

Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita

21. 10. 2019

Konstrukce pravděpodobnostního prostoru

Definice (Konečná aditivita)

Nechť A, B jsou disjunktní jevy. Vlastnost

$$P(A \cup B) = P(A) + P(B)$$

se nazývá **(konečná) aditivita**.
Indukcí lze rozšířit na konečný počet disjunktních jevů $A_1, \ldots, A_n$,

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k).$$

Definice (Spočetná aditivita, σ-aditivita)

Nechť $A_1, A_2, A_3, \ldots$ je spočetná (konečná či nekonečná) posloupnost vzájemně disjunktních jevů. Vlastnost

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

se nazývá **spočetná aditivita (σ-aditivita)**.

Příklad

Pro podmnožinu $A \subseteq [0;1]$ a číslo $r \in [0;1]$ definujme tzv. *r-posun* množiny A:

$$A \oplus r = \{a + r;\ a \in A, a + r \leq 1\} \cup \{a + r - 1;\ a \in A, a + r - 1 > 0\}.$$

Pro náhodnou veličinu $X \sim Rs([0;1])$ by mělo platit

$$P(A \oplus r) = P(A), \qquad A \subseteq [0;1], r \in [0;1].$$

Tvrzení

Neexistuje pravděpodobnost P definovaná pro všechny podmnožiny $A \subseteq [0;1]$, splňující vlastnosti:

- $P([a,b]) = P((a,b)) = P((a,b]) = P([a,b)) = b - a, \quad 0 \leq a \leq b \leq 1,$
- $P(A \oplus r) = P(A), \qquad r \in [0;1],$
- *P je spočetně aditivní (σ-aditivní).*

Tvrzení říká, že chceme-li, aby pravděpodobnost P na $[0;1]$ měla *rozumné* vlastnosti, nelze ji definovat pro všechny podmnožiny $A \subseteq [0;1]$. Musíme definici P omezit jen na tzv. **měřitelné** množiny množiny A.

Tvrzení se dokazuje sporem. Předpokládejme tedy, že $P(A)$ je definována pro všechny $A \subseteq [0;1]$. Pro $x, y \in [0;1]$ definujeme relaci ekvivalence $x \sim y \Leftrightarrow y - x \in \mathbb{Q}$. Ta interval rozdělí disjunktních do tříd ekvivalence. Zvolíme v každé třídě jednoho zástupce (přičemž místo příp. 0 zvolíme jako zástupce např. $\frac{1}{2}$) a z nich vytvoříme množinu H. Nyní platí

$$(0;1] = \bigcup_{r \in \mathbb{Q} \cap [0;1)} (H \oplus r),$$

přičemž množiny $(H \oplus r)$ jsou vzájemně disjunktní.
Ze σ-aditivity a r-posunu nyní obdržíme

$$P((0;1]) = \sum_{r \in \mathbb{Q} \cap [0;1)} P(H \oplus r) = \sum_{r \in \mathbb{Q} \cap [0;1)} P(H).$$

Spočetná suma stejných pravděpodobností může být rovna pouze 0, anebo $\pm\infty$. To je však ve sporu s tím, že chceme $P((0;1]) = 1 - 0 = 1$.

Definice

Množina $\mathcal{A}$ podmnožin Ω je **algebra** *(algebra, field)*, pokud:

- $\mathcal{A}$ je neprázdná,
- $\mathcal{A}$ je uzavřená na doplňky,
- $\mathcal{A}$ je uzavřená na konečná sjednocení.

Množina $\mathcal{A}$ podmnožin Ω je **σ-algebra** *(σ-algebra, σ-field)*, pokud:

- $\mathcal{A}$ je neprázdná,
- $\mathcal{A}$ je uzavřená na doplňky,
- $\mathcal{A}$ je uzavřená na spočetná sjednocení.

Množina $\mathcal{A}$ podmnožin Ω je **semialgebra**, pokud:

- $\emptyset, \Omega \in \mathcal{A}$,
- $\mathcal{A}$ je uzavřená na konečné průniky,
- doplněk každého prvku z $\mathcal{A}$ je rovný konečnému sjednocení disjunktních prvků z $\mathcal{A}$.

Tvrzení

- $\emptyset, \Omega \in \mathcal{A}$,
- *algebra $\mathcal{A}$ je uzavřená na průniky, σ-algebra $\mathcal{A}$ i na spočetné průniky.*

Definice (Pravděpodobnostní prostor)

Pravděpodobnostní prostor *(probability space, probability triple, probability measurable space)* je trojice $(\Omega, \mathcal{A}, P)$, kde

- $\Omega \neq \emptyset$ je **základní prostor** *(sample space)*,
- $\mathcal{A}$ je σ-algebra podmnožin Ω,
- **pravděpodobnost (pravděpodobnostní míra)** *(probability measure)* P je zobrazení $P : \mathcal{A} \to [0;1]$ s vlastnostmi:
  - P je spočetně aditivní,
  - $P(\Omega) = 1$.

Dvojice $(\Omega, \mathcal{A})$ se nazývá **jevové pole (měřitelný prostor)** *(measurable space)*. Prvky $A \in \mathcal{A}$ nazýváme **jevy** *(events)* nebo **měřitelné množiny** *(measurable sets)*. Jsou to takové podmnožiny $A \subseteq \Omega$, pro které je pravděpodobnost $P(A)$ korektně definována. Na příkladu jsme viděli, že $\mathcal{A}$ obecně nemusí obsahovat všechny podmnožiny Ω.

Tvrzení (Vlastnosti pravděpodobnosti)

Pro $A, B, A_1, A_2, \dots \in \mathcal{A}$ platí:

- $P(\emptyset) = 0$,
- $P(\overline{A}) = 1 - P(A)$,
- $P(A) \leq P(B)$ *pro* $A \subseteq B$, *(monotonie pravděpodobnosti)*
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, *(princip inkluze a exkluze)*
- $P\left(\bigcup_{k=1}^{\infty} A_i\right) \leq \sum_{k=1}^{\infty} P(A_i)$. *($\sigma$-subaditivita)*

Tvrzení

Nechť Ω je konečná nebo spočetná, neprázdná množina. Nechť $p : \Omega \to [0;1]$ je libovolné zobrazení splňující $\sum_{\omega \in \Omega} p(\omega) = 1$. Potom $(\Omega, \mathcal{A}, P)$, kde $\mathcal{A} = 2^{\Omega}$ a $P(A) = \sum_{\omega \in A} p(\omega)$, je pravděpodobnostní prostor.

Tvrzení říká, že konstrukce pravděpodobnostního prostoru je pro nejvýše spočetný základní prostor Ω přímočará. Bez obav lze jako σ-algebru jevů zvolit množinu všech podmnožin a pravděpodobnost jevů definovat tak, jak ji známe z příkladů náhodných veličin diskrétního typu.

Příklad (Klasická pravděpodobnost)

Uvažujme konečný základní prostor $\Omega \neq \emptyset$, $\mathcal{A} = 2^{\Omega}$ a pravděpodobnost $P(A) = \frac{|A|}{|\Omega|}$, kde $|\cdot|$ označuje počet prvků množiny. Potom $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor. Nazývá se **rovnoměrné rozdělení** na Ω, $Ro(\Omega)$. Jednoduše ověříme, že $p(\omega) = P(\{\omega\}) = \frac{1}{n}$, kde $n = |\Omega|$, jedná se o velmi dobře známý případ, jde o **klasickou pravděpodobnost**.

V případě nespočetného základního prostoru Ω je však situace zcela odlišná. Představme si $\Omega = [0;1]$. Jak zvolíme σ-algebru $\mathcal{A}$ a na ní pravděpodobnost P? Již víme, že ani pro tak jednoduchý základní prostor nelze vzít $\mathcal{A} = 2^{\Omega}$, protože pro některé podmnožiny $A \subseteq [0;1]$ nelze pravděpodobnost $P(A)$ vůbec definovat. Na druhou stranu, z praktických důvodů, určitě chceme, aby $\mathcal{A}$ obsahovala alespoň všechny intervaly $I \subseteq [0;1]$ a jednobodové podmnožiny:
$\mathcal{J} = \{$všechny intervaly tvarů $[a,b], (a,b), [a,b), (a,b), \{a\}\}$, $\mathcal{J} \subseteq \mathcal{A}$.
Z **rovnoměrného spojitého** rozdělení $Rs([0;1])$ jsme totiž zvyklí pravděpodobnost takových intervalů $I \in J$ počítat jako jejich délku, $P(I) = b - a$, $P(\{a\}) = 0$. Ověřte, že $\mathcal{J}$ je semialgebra. Ověřte, že když do $\mathcal{J}$ přidáme všechna konečná sjednocení prvků z $\mathcal{J}$, obdržíme algebru. Ale ani poté, co do $\mathcal{J}$ přidáme všechna konečná a spočetná sjednocení, nestane se σ-algebrou. Zjišťujeme, že zkonstruovat korektní σ-algebru $\mathcal{A}$ i na jednoduchém nespočetném základním prostoru $\Omega = [0;1]$ není triviálním úkolem.

Tvrzení (Věta o rozšíření)

Nechť $\mathcal{J}$ je semialgebra podmnožin množiny Ω a nechť zobrazení $P : \mathcal{J} \to [0;1]$ má následující vlastnosti:

- $P(\emptyset) = 0, \quad P(\Omega) = 1,$
- *pro vzájemně disjunktní* $A_1, A_2, \ldots, A_n \in \mathcal{J}$, $\bigcup_{i=1}^{n} A_i \in \mathcal{J}$:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i), \qquad \text{(konečná superaditivita)}$$

- *pro* $A_1, A_2, \dots \in \mathcal{J}$, $A \subseteq \bigcup_{i=1}^{\infty} A_i$:

$$P(A) \leq \sum_{i=1}^{\infty} P(A_i). \qquad \text{(σ-monotonie)}$$

Potom existuje σ-algebra $\mathcal{A}$, $\mathcal{J} \subseteq \mathcal{A}$, a spočetně aditivní pravděpodobnostní míra P^* na $\mathcal{A}$, tak, že

$$P^*(A) = P(A), \qquad A \in \mathcal{J},$$

a $(\Omega, \mathcal{A}, P^*)$ je pravděpodobnostní prostor.

Věta o rozšíření dává návod na konstrukci pravděpodobnostního prostoru, když máme definovanou pravděpodobnost P na semialgebře $\mathcal{J}$. Musíme přitom vyřešit dvě otázky: jak definovat $P^*(A)$ pro množiny $A \in \mathcal{A} \setminus \mathcal{J}$ a jak zkonstruovat σ-algebru $\mathcal{A}$, jejíž existence je Větou zaručena.

Ověřování podmínek ve výše uvedené Větě může být velmi obtížné. Existují i alternativní ekvivalentní formulace těchto předpokladů, které nyní uvedeme.

Ekvivalentní předpoklady Věty o rozšíření

- $P(\emptyset) = 0, \quad P(\Omega) = 1$,
- $P(A) \leq P(B) \quad$ pro $A, B \in \mathcal{J},\ A \subseteq B$,
- P je konečně superaditivní,
- P je σ-subaditivní.

Ekvivalentní předpoklady Věty o rozšíření pro pravděpodobnost

- $P(\Omega) = 1$,
- P je σ-aditivní.

Míra, vnější míra

Definice (Míra)

Nechť $(\Omega, \mathcal{A})$ je měřitelný prostor. **Míra** *(measure)* je zobrazení

$$\mu : \mathcal{A} \to [0; \infty]$$

s vlastnostmi:

- $\mu(A) \geq 0 \quad$ pro $A \in \mathcal{A}$,
- $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k) \quad$ pro disjunktní $A_1, A_2, \dots \in \mathcal{A}$,
- $\mu(\emptyset) = 0$.

Míra μ je:

- **nekonečná**, pokud $\mu(\Omega) = \infty$,
- **σ-konečná**, pokud $\mu(\Omega) = \infty$, ale přitom existuje úplný systém jevů $\{A_1, A_2, \dots\}$ na Ω takový, že $\mu(A_k) < \infty$, $k = 1, 2, \dots$,
- **konečná**, pokud $\mu(\Omega) < \infty$,
- **pravděpodobnostní**, pokud $\mu(\Omega) = 1$.

Trojice $(\Omega, \mathcal{A}, \mu)$ je **prostor s mírou**.

Definice (Vnější míra)

Vnější míra *(outer measure)* je zobrazení

$$\mu^\circ : 2^\Omega \to [0; \infty]$$

s vlastnostmi:

- $\mu^\circ(\emptyset) = 0.$
- $\mu^\circ(A) \leq \mu^\circ(B) \quad$ pro $A \subseteq B$,
- $\mu^\circ\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu^\circ(A_k) \quad$ pro $A_1, A_2, \dots \subseteq \Omega.$

Speciální vnější míru lze definovat pomocí míry na algebře $\mathcal{A} \subseteq 2^\Omega$:

$$\mu^*(A) = \inf \sum_{k=1}^{\infty} \mu(A_k),$$

kde infimum je počítáno přes všechna spočetná pokrytí množiny A množinami z algebry $\mathcal{A}$, tj. přes všechny $A_1, A_2, \dots \in \mathcal{A}$, $A \subseteq \bigcup_{k=1}^{\infty} A_i$.

Tvrzení (Vlastnosti míry)

Nechť $(\Omega, \mathcal{A}, \mu)$ je prostor s mírou. Potom platí:

- μ *je konečně aditivní,*
- $\mu(A) \leq \mu(B) \quad$ pro $A \subseteq B$,
- μ *je σ-subaditivní, tj.* $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{n} \mu(A_k)$.

Tvrzení (Vlastnosti vnější míry)

- *Vnější míra μ° je nezáporná.*
- *Vnější míra μ° je konečně subaditivní.*
- *Míra je restrikce vnější míry μ° na σ-algebru $\mathcal{A} \subseteq 2^\Omega$.*
- *Zobrazení μ^* je vnější mírou (tzn. μ^* je speciální případ μ°).*
- *μ^* je rozšíření míry μ ze σ-algebry $\mathcal{A}$ na všechny podmnožiny, 2^Ω.*
- *Je-li μ σ-konečná na $\mathcal{A}$, potom je μ^* σ-konečná na 2^Ω.*
- *Je-li μ konečná na $\mathcal{A}$, potom je μ^* konečná na 2^Ω.*

Porovnejte definici a vlastnosti míry s definicí a vlastnostmi pravděpodobnosti. Pravděpodobnost je vlastně speciální mírou, $P = \mu$, když $\mu(\Omega) = 1$. Naopak, na míru lze nahlížet jako na zobecnění pravděpodobnosti, kdy jevy (měřitelné množiny) ohodnocujeme nezáporným číslem (včetně nekonečna), $\mu(A) \in [0; \infty]$.

Vnější míra je důležitý koncept pro konstrukci míry (pravděpodobnosti). Na rozdíl od míry (pravděpodobnosti) je totiž definována pro všechny podmnožiny $A \subseteq \Omega$ a nevyžaduje σ-algebru $\mathcal{A}$.

Tvrzení (Carathéodoryho věta o rozšíření)

Nechť μ je míra na algebře $\mathcal{J}$ podmnožin množiny Ω. Potom platí:

- *Míru μ lze rozšířit na σ-algebru $\mathcal{A}$ generovanou algebrou $\mathcal{J}$.*
- *Je-li μ σ-konečná na $\mathcal{J}$, její rozšíření na $\mathcal{A}$ je jednoznačné a σ-konečné.*
- *Je-li μ konečná na $\mathcal{J}$, její rozšíření na $\mathcal{A}$ je jednoznačné a konečné.*

V předchozí formulaci Věty o rozšíření sice stačí, aby $\mathcal{J}$ byla jen semialgebra, navíc je ale požadováno splnění dalších vlastností pravděpodobnosti P.

Pokud míra μ na $\mathcal{J}$ není konečná ani σ-konečná, její rozšíření na $\mathcal{A}$ nemusí být jednoznačné, tzn. může existovat více (až nekonečně mnoho) rozdílných rozšíření.

Pokud pracujeme s pravděpodobnostní mírou, $\mu = P$, Carathéodoryho věta říká, že rozšíření P z algebry $\mathcal{J}$ na σ-algebru $\mathcal{A}$ existuje a je jednoznačné. Nutnou podmínkou přitom je, že míra μ, resp. pravděpodobnost P, je korektně definována na algebře $\mathcal{J}$.

Definice (Množiny měřitelné vzhledem ke vnější míře)

Nechť μ° je vnější míra. Množina $A \in 2^\Omega$ je tzv. **μ°-měřitelná**, pokud pro všechny *testovací* množiny $E \in 2^\Omega$ platí

$$\mu^\circ(E) = \mu^\circ(A \cap E) + \mu^\circ(\overline{A} \cap E),$$

tzn. pokud je μ° aditivní na množinách $A \cap E$ a $\overline{A} \cap E$.

μ°-měřitelné množiny existují, např. si ověřte, že $\emptyset$, Ω jsou vždy μ°-měřitelné. Dále, pokud je nějaká množina A μ°-měřitelná, pak je i její doplněk $\overline{A}$ μ°-měřitelný.

Následující věta je důležitá. Dává totiž návod, jak na Ω zkonstruovat σ-algebru a na ní míru. Vychází přitom jen z existence vnější míry.

Tvrzení

Systém všech μ°-měřitelných množin,

$$\mathcal{A} = \left\{A \in 2^\Omega : \mu^\circ(E) = \mu^\circ(A \cap E) + \mu^\circ(\overline{A} \cap E) \text{ pro všechny } E \in 2^\Omega\right\},$$

je vždy σ-algebrou na Ω a vnější míra μ° je dokonce mírou na této σ-algebře $\mathcal{A}$.

Nyní se vraťme k Větě o rozšíření. Zobrazení P^* je vlastně vnější mírou počítanou pomocí pravděpodobnosti P na semialgebře $\mathcal{J}$:

$$P^*(A) = \inf \sum_{k=1}^{\infty} P(A_k),$$

kde infimum je počítáno přes všechna spočetná pokrytí množiny A množinami z $\mathcal{J}$, tj. přes všechny $A_1, A_2, \dots \in \mathcal{J}$, $A \subseteq \bigcup_{k=1}^\infty A_k$.

Na základě předchozích tvrzení si shrňme důležité poznatky:

1. $P^*(\varnothing) = 0$,
2. $P^*(A) \leq P^*(B)$ pro $A \subset B$,
3. P^* je rozšíření P z $\mathcal{J}$ na 2^Ω, tzn. pro $A \in \mathcal{J}$ platí $P^*(A) = P(A)$,
4. P^* je σ-subaditivní,
5. systém $\mathcal{A}$ všech P^*-měřitelných množin,

$$\mathcal{A} = \left\{A \in 2^\Omega : P^*(E) = P^*(A \cap E) + P^*(\overline{A} \cap E) \text{ pro všechny } E \in 2^\Omega\right\},$$

   je σ-algebrou na Ω,
6. $\mathcal{J} \subseteq \mathcal{A}$,
7. P^* je σ-aditivní na $\mathcal{A}$.

Tyto vlastnosti říkají, že $\mathcal{A}$ je σ-algebra obsahující semialgebru $\mathcal{J}$, P^* je pravděpodobnostní míra na $\mathcal{A}$ a je rozšířením pravděpodobnosti P z $\mathcal{J}$. Zkonstruovaný pravděpodobnostní prostor $(\Omega, \mathcal{A}, P^*)$ je navíc **úplný** *(complete)*, tzn. pokud pro $A \in \mathcal{A}$ platí $P^*(A) = 0$, potom pro libovolnou $B \subseteq A$ platí: $B \in \mathcal{A}$ a $P^*(B) = 0$.

Definice (Součinový prostor, součinová míra)

Nechť $(\Omega_1, \mathcal{A}_1, \mu_1)$, $(\Omega_2, \mathcal{A}_2, \mu_2)$ jsou dva prostory s mírou.
Součinový prostor *(product space)* je kartézský součin
$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$.
Na systému $\mathcal{J} = \{A_1 \times A_2 : A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2\}$ definujeme **součinovou míru** *(product measure)* $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$.
Množina $A_1 \times A_2 \in \mathcal{J}$ je tzv. **měřitelný obdélník** *(measurable rectangle)*, A_1 a A_2 jsou jeho **strany**.

Ověřte, že $\mathcal{J}$ je semialgebra.

V případě, že oba prostory jsou pravděpodobnostní, tzn. $\mu_1 = P_1$, $\mu_2 = P_2$, definujeme analogicky $P(A_1 \times A_2) = P_1(A_1)P_2(A_2)$.

Ověřte, že platí $P(\varnothing \times \varnothing) = 0$, $P(\Omega_1 \times \Omega_2) = 1$.

Příklady konstrukcí pravděpodobnostních prostorů

Vraťme se k nespočetnému základnímu prostoru $\Omega = [0;1]$. Zvolme systém $\mathcal{J} = \{$všechny intervaly tvarů $[a, b], (a, b), [a, b), (a, b), \{a\}\}$. Systém $\mathcal{J}$ je semialgebra na Ω. Pro intervaly $I \in \mathcal{J}$ definujeme $P(I)$ jako délku daného intervalu, např. $P([a, b]) = b - a$, $P(\{a\}) = 0$, $P(\emptyset) = 0$, $P([0;1]) = 1$. Zobrazení P má na $\mathcal{J}$ vlastnosti pravděpodobnosti.

Věta o rozšíření zaručuje existenci jednoznačného rozšíření P na σ-algebru $\mathcal{A}$, $\mathcal{J} \subseteq \mathcal{A}$, tak, že $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor. Tento pravděpodobnostní prostor se nazývá **Lebesgueova míra** *(Lebesgue measure)* na $[0;1]$. Zkonstruovaná pravděpodobnostní míra se někdy označuje $\lambda = P$.

$\mathcal{B}_{[0;1]} = \sigma(\mathcal{J})$ je tzv. **borelovská σ-algebra** na intervalu $[0;1]$, jedná se o σ-algebru generovanou systémem $\mathcal{J}$. tzn. že je to nejmenší σ-algebra obsahující systém $\mathcal{J}$. Pro σ-algebru $\mathcal{A}$ z Věty o rozšíření musí tedy platit $\mathcal{B}_{[0;1]} \subseteq \mathcal{A}$.

Lze ukázat, že Lebesgueova míra na $\mathcal{B}_{[0;1]}$ však není úplná, zatímco na $\mathcal{A}$ úplná je.

Již víme, že $\lambda([a, b]) = b - a$, $\lambda(\{a\}) = 0$. Ze σ-aditivity míry dokonce víme, že $\lambda(A) = 0$ pro libovolnou spočetnou množinu $A \subseteq [0;1]$, protože takovou A lze zapsat jako spočetné sjednocení disjunktních jednoprvkových množin, které mají nulovou míru. Např. tedy také víme, že $\lambda(\mathbb{Q}) = 0$.

Připomeňme, že s Lebesgueovou mírou vlastně pracujeme v případě náhodné veličiny X s rozdělením $X \sim Rs([0;1])$. Vybíráme-li tedy náhodně reálné číslo X z intervalu $[0;1]$, víme, že $P(X \in \mathbb{Q}) = \lambda(\mathbb{Q}) = 0$.

Existují dokonce nespočetné množiny, které mají nulovou Lebesgueovu míru. Tzv. Cantorova množina K je definována následujícím postupem. Interval $[0;1]$ rozdělíme na třetiny a prostřední část ve tvaru otevřeného intervalu odebereme. Na zbylé intervaly použijeme stejný postup, ten dále induktivně opakujeme, tzn. v k-tém kroku odstraníme 2^{k-1} prostředních třetin (odstraňujeme je ve formě otevřených intervalů) zbylých intervalů, které v aktuálním kroku mají délku $\frac{1}{3^k}$. Cantorova množina K je tvořena těmi body, které zůstanou neodstraněné.

Počítejme Lebesgueovu míru doplňku Cantorovy množiny,

$$\lambda(\overline{K}) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{9} + 4 \cdot \frac{1}{27} + \cdots = \sum_{k=1}^{\infty} 2^{k-1} \cdot \frac{1}{3^k} = 1.$$

Lebesgueova míra na $[0;1]$ je pravděpodobnostní, tedy platí

$$\lambda(K) = 1 - \lambda(\overline{K}) = 0,$$

Lebesgueova míra Cantorovy množiny je nulová.

Cantorova množina je zároveň nespočetná. Totiž, pro $x \in K$ označme $d_k(x) = 0$, resp. $d_k(x) = 1$, pokud v k-tém kroku konstrukce K byl bod x nalevo, resp. napravo, od nejbližšího odstraňovaného otevřeného intervalu. Definujeme funkci $g : K \to [0;1]$ předpisem $g(x) = \sum_{k=1}^{\infty} \frac{d_k(x)}{2^k}$. Protože $g(K) = [0;1]$ a interval $[0;1]$ je nespočetná množina, musí být i Cantorova množina K nespočetná.

Borelovská σ-algebra na $\mathbb{R}$ je σ-algebra $\mathcal{B} = \sigma(\mathcal{J})$ generovaná systémem $\mathcal{J}$ všech intervalů na $\mathbb{R}$.

Lze ukázat, že $\mathcal{B} = \sigma(\mathcal{M})$, kde $\mathcal{M} = \{(-\infty, b); b \in \mathbb{R}\}$.

Házíme n-krát mincí. Jako pravděpodobnostní prostor přímočaře volíme

$$\Omega = \{(b_1, \ldots, b_n); b_k \in \{0;1\}, k = 1, \ldots, n\}, \quad \mathcal{A} = 2^\Omega, \quad P(A) = \frac{|A|}{2^n}.$$

Pokračujme nyní nekonečným házením mincí. Základní prostor definujeme jako množinu všech posloupností čísel 0 (rub) a 1 (líc),

$$\Omega = \{(b_1, b_2, \dots); b_k \in \{0;1\}, k \in \mathbb{N}\}.$$

Jak zkonstruujeme σ-algebru $\mathcal{A}$ a na ní pravděpodobnost P, abychom dostali pravděpodobnostní prostor?

Pro $n \in \mathbb{N}$ a $a_1, a_2, \dots, a_n \in \{0;1\}$ definujeme jevy

$$A_{a_1 a_2 \dots a_n} = \{(b_1, b_2, \dots) \in \Omega : a_k = b_k \text{ pro } k = 1, \dots, n\}.$$

Jakým situacím tyto jevy (při konkrétní volbě 0 a 1 za a_k) odpovídají? Intuitivně na základě znalosti z konečného házení mincí chceme, aby $P(A_{a_1 a_2 \dots a_n}) = \frac{1}{2^n}$. Pomocí těchto jevů vytvoříme systém $\mathcal{J}$,

$$\mathcal{J} = \{A_{a_1 a_2 \dots a_n}; n \in \mathbb{N}, a_k \in \{0;1\}, k = 1, \dots, n\} \cup \{\emptyset, \Omega\}.$$

Ověřte, že $\mathcal{J}$ je semialgebra a že P je na $\mathcal{J}$ (konečně) aditivní. Lze ukázat (obtížný důkaz), že P je na $\mathcal{J}$ dokonce σ-aditivní.

Použijeme alternativní formulaci Věty o rozšíření, která tvrdí, že P lze rozšířit na σ-algebru $\mathcal{A}$ obsahující $\mathcal{J}$. Tato σ-algebra $\mathcal{A}$ však bude velmi složitá, neočekávejme žádný její explicitní popis. Důležité je vědět, že vůbec existuje.

Pro $n \in \mathbb{N}$ definujeme jevy $H_n = \{(b_1, b_2, \dots) \in \Omega : b_n = 1\}$. Interpretujte je a pomocí vhodných jevů $A_{\dots}$ a σ-aditivity P ověřte, že $P(H_n) = \frac{1}{2}$.

Zajímavý postřeh je, že každé číslo $x \in [0;1]$ lze jednoznačně zapsat ve tvaru binárního rozvoje $x = \sum_{k=1}^{\infty} \frac{b_k}{2^k}$, kde $b_k \in \{0;1\}$, $k = 1, 2, \dots$. Každé číslo $x \in [0;1]$ lze tedy jednoznačně ztotožnit s jednou z posloupností $(b_1, b_2, \dots) \in \Omega$. Pravděpodobnostní prostor při nekonečném házení mincí tedy *v postatě odpovídá* Lebesgueově míře na $[0;1]$.

Příklad

Uvažujme algebry $\mathcal{J}_1 = \{\emptyset, \Omega, A, \overline{A}\}$, $\mathcal{J}_2 = \{\emptyset, \Omega, B, \overline{B}\}$ pro $A \neq B$, $A \cap B \neq \emptyset$ a zaveďme $\mathcal{J}$ jako algebru generovanou sjednocením $\mathcal{J}_1 \cup \mathcal{J}_2$. Definujeme pravděpodobnostní míry μ_1 a μ_2:

$$\mu_k(\emptyset) = 0,\ \ \mu_k(\Omega) = 1,\ \ \mu_k(A) = 0{,}40,\ \ \mu_k(B) = 0{,}35, \qquad k = 1, 2;$$

$$\mu_1(A \cup B) = 0{,}50,\ \ \mu_2(A \cup B) = 0{,}60.$$

Nalezněte všechny prvky $\mathcal{J}$ a rozšiřte μ_1 a μ_2 na celé $\mathcal{J}$. Všimněte si, že rozšířeni míry z $\mathcal{J}_1 \cup \mathcal{J}_2$ na $\mathcal{J}$ není jednoznačné. Je to důsledkem toho, že původní systém $\mathcal{J}_1 \cup \mathcal{J}_2$, z něhož jsme generovali $\mathcal{J}$ není algebrou.

Příklad

Uvažujte konečný, resp. spočetný, základní prostor Ω. Pro $A \subset \Omega$ definujme zobrazení μ, $\mu(A) = |A|$, jako počet prvků A (příp. ∞).
Ověřte, že μ je míra na 2^{Ω}, a že je konečná, resp. σ-konečná. Nazývá se **čítací míra** *(counting measure)*.

Příklad

Uvažujte $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$
a $\mathcal{J} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_3, \omega_4\}\}$. Definujeme:

A	$\emptyset$	Ω	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_3\}$	$\{\omega_4\}$	$\{\omega_1, \omega_2\}$	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_4\}$	$\{\omega_3, \omega_4\}$
$\mu(A)$	0	6					3	3	3	3
$\mu_1(A)$			1	2	2	1				
$\mu_2(A)$			2	1	1	2				

1. Ověřte, že $\mathcal{J}$ není algebra.

2. Ověřte, že μ je míra na $\mathcal{J}$.

3. Dodefinujte μ_1 a μ_2 tak, aby obě byly míry na 2^Ω.

4. Ověřte, že μ_1 a μ_2 jsou dvě rozšíření μ z $\mathcal{J}$ na 2^Ω.

5. Spočítejte vnější míru μ^* na 2^Ω pomocí μ.

6. Všimněte si, že $\mu_1 \neq \mu_2 \neq \mu^*$. Důvodem je to, že $\mathcal{J}$ není algebra.

Příklad

Uvažujte $\Omega = \mathbb{N}_0$, $A = \{\text{lichá přirozená čísla}\}$, $\mathcal{J} = \{\emptyset, \Omega, A, \overline{A}\}$. Nechť μ je čítací míra na $\mathcal{J}$ a definujme $\mu_1(A) = |A|$, $\mu_2(A) = 2|A|$ pro $A \in 2^\Omega$.

1. Ověřte, že $\mathcal{J}$ je algebra.
2. Ověřte, že μ není σ-konečná míra.
3. Ověřte, že μ_1 a μ_2 jsou rozšíření μ z $\mathcal{J}$ na 2^Ω.
4. Ověřte, že μ_1 a μ_2 jsou σ-konečné míry.
5. Spočítejte vnější míru μ^*.
6. Ověřte, že σ-algebra μ^*-měřitelných množin je rovna 2^Ω, tj. že všechny podmnožiny Ω jsou zde μ^*-měřitelné.

Náhodné veličiny

Definice (Náhodná veličina)

Uvažujme pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$. **Náhodná veličina** *(random variable)* je zobrazení $X : \Omega \to \mathbb{R}$ splňující podmínku

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A} \qquad \text{pro všechna } x \in \mathbb{R}.$$

Uvedená podmínka je tzv. **měřitelnost**. Lze ji zapsat také ve tvaru $\{X \leq x\} \in \mathcal{A}$ nebo $X^{-1}((-\infty, x]) \in \mathcal{A}$. Protože doplňky, sjednocení i průniky jsou v úplných vzorech zachovávány, je podmínka ekvivalentní

$$X^{-1}(B) = \{X \in B\} = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A} \text{ pro všechny } B \in \mathcal{B}.$$

Podmínka vlastně zajišťuje, že pro libovolnou borelovskou množinu B je vůbec definována pravděpodobnost $P(X \in B)$.

Definice (Indikátorová funkce)

Indikátorová funkce množiny A je zobrazení $\mathbf{1}_A : \Omega \to \{0;1\}$ definované

$$\mathbf{1}_A(\omega) = \begin{cases} 0, & \omega \notin A; \\ 1, & \omega \in A. \end{cases}$$

Příklad

Ne každé zobrazení $X : \Omega \to \mathbb{R}$ je náhodnou veličinou. Např. pro Lebesgueovu míru $(\Omega = [0;1], \mathcal{A}, \lambda)$ uvažujme neměřitelnou množinu H, $H \subseteq \Omega$, $H \notin \mathcal{A}$, a zobrazení $X = \mathbf{1}_{\overline{H}}$.
Potom $X^{-1}\left((-\infty, \frac{1}{2})\right) = X^{-1}(\{0\}) = \{\omega \in \Omega : \omega \in H\} = H \notin \mathcal{A}$, tzn. zobrazení X není náhodnou veličinou.

Definice (borelovsky měřitelná funkce)

Funkce $h : \mathbb{R} \to \mathbb{R}$ je **borelovsky měřitelná**, pokud $h^{-1}(B) \in \mathcal{B}$ pro všechny $B \in \mathcal{B}$, tzn. je vlastně náhodnou veličinou na měřitelném prostoru $(\Omega = \mathbb{R}, \mathcal{A} = \mathcal{B})$.

Tvrzení (Jak tvořit náhodné veličiny)

- *Indikátorová funkce, $X = \mathbf{1}_A$, jevu $A \in \mathcal{A}$ je náhodnou veličinou.*
- *Pokud X, Y jsou náhodné veličiny a $c \in \mathbb{R}$, potom také $X + c$, $c\,X$, X^2, $X + Y$, $X\,Y$ jsou náhodné veličiny.*
- *Pokud $X_1, X_2, \ldots$ je posloupnost náhodných veličin taková, že limita číselné posloupnosti $\lim_{n\to\infty} X_n(\omega)$ existuje pro všechna $\omega \in \Omega$, potom $X(\omega) = \lim_{n\to\infty} X_n(\omega)$ je náhodnou veličinou.*
- *Pokud X je náhodná veličina a $h : \mathbb{R} \to \mathbb{R}$ je borelovsky měřitelná funkce, potom zobrazení $Y = h \circ X = h(X)$, kde $Y(\omega) = f(X)(\omega) = h(X(\omega))$ pro všechna $\omega \in \Omega$, je náhodnou veličinou (na stejném pravděpodobnostním prostoru).*

Každá spojitá nebo po částech spojitá funkce je borelovsky měřitelná. Tvrzení říká, že všechny *rozumné* způsoby transformací náhodných veličin jsou správné. Speciálně si připomeňme funkce $h(x) = x^k$, vedoucí k náhodným veličinám $Y = h(X) = X^k$, $k \in \mathbb{N}$.

Definice (Nezávislost náhodných jevů)

Jevy $A, B \in \mathcal{A}$ jsou **(stochasticky) nezávislé** *(independent events)*, pokud $P(A \cap B) = P(A)P(B)$.

Systém jevů $\{A_i\}_{i \in I}$ je nezávislý, pokud pro každé $n \in \mathbb{N}$ a každou volbu indexů $i_1, \ldots, i_n$ z indexové množiny I platí $P\left(\bigcap_{k=1}^{n} A_{i_k}\right) = \prod_{k=1}^{n} P(A_{i_k})$.

Definice (Nezávislost náhodných veličin)

Náhodné veličiny X, Y jsou **nezávislé** *(independent)*, pokud jsou pro každé $B_1, B_2 \in \mathcal{B}$ nezávislé jevy $X^{-1}(B_1)$ a $Y^{-1}(B_2)$, tzn. $P(X \in B_1, Y \in B_2) = P(X \in B_1)P(Y \in B_2)$.

Systém náhodných veličin $\{X_i\}_{i \in I}$ je nezávislý, pokud pro každé $n \in \mathbb{N}$ a každou volbu indexů $i_1, \ldots, i_n$ z indexové množiny I platí $P\left(\bigcap_{k=1}^{n} X_{i_k} \in B_k\right) = \prod_{k=1}^{n} P(X_{i_k} \in B_k)$ pro každé $B_1, \ldots, B_n \in \mathcal{B}$.

Připomeňme dvě tvrzení o nezávislosti, která jsme zvyklí běžně používat.

Tvrzení

Nechť X *a* Y *jsou nezávislé náhodné veličiny a* $g, h : \mathbb{R} \to \mathbb{R}$ *borelovsky měřitelné funkce. Potom náhodné veličiny* $g \circ X = g(X)$ *a* $h \circ Y = h(Y)$ *jsou opět nezávislé.*

Tvrzení

Nechť X *a* Y *jsou náhodné veličiny na stejném pravděpodobnostním prostoru. Potom* X *a* Y *jsou nezávislé* $\Leftrightarrow P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ *pro všechna* $x, y \in \mathbb{R}$.

Spojitost pravděpodobnosti a limitní jevy

Definice

Pro posloupnost jevů $A_1, A_2, \ldots, A \in \mathcal{A}$ v pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ zavádíme značení:

1. $A_n \nearrow A$ označuje $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots, \ \bigcup_{n=1}^{\infty} A_n = A$,

2. $A_n \searrow A$ označuje $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots, \ \bigcap_{n=1}^{\infty} A_n = A$.

Následující věta říká, že pokud *v jistém smyslu konverguje* posloupnost jevů, konverguje i posloupnost pravděpodobností.

Tvrzení (O spojitosti pravděpodobnosti)

Pokud $A_n \nearrow A$*, nebo* $A_n \searrow A$*, potom* $\lim_{n\to\infty} P(A_n) = P(A)$.

Podstatným předpokladem přitom je vnoření jevů v posloupnosti $A_n \nearrow A$, nebo $A_n \searrow A$. I pokud posloupnost jevů konverguje, ale jevy nejsou správně vnořené, rovnost $\lim_{n\to\infty} P(A_n) = P(A)$ nemusí platit, dokonce limita pravděpodobností ani nemusí existovat.

Příklad

Uvažujme posloupnost jevů $A_n = \Omega$ pro lichá n, $A_n = \emptyset$ pro sudá n. Dostáváme $A = \bigcup_{n=1}^{\infty} A_n = \Omega$ a $P(A) = P(\Omega) = 1$, resp. $B = \bigcap_{n=1}^{\infty} A_n = \emptyset$ a $P(B) = P(\emptyset) = 0$. Ale limita pravděpodobností $\lim\limits_{n\to\infty} P(A_n)$ neexistuje, neboť posloupnost pravděpodobností $\{P(A_n)\}_{n=1}^{\infty}$ osciluje, $\liminf\limits_{n\to\infty} P(A_n) = 0$, $\limsup\limits_{n\to\infty} P(A_n) = 1$.

Definice (Limitní jevy)

Pro jevy $A_1, A_2, \dots \in \mathcal{A}$ v pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ definujeme limitní jevy

1. A_n **nekonečně často** *(infinitely often, i. o.)*

$$\limsup_n A_n = \{\omega \in \Omega : \text{v nekonečně mnoha jevech } A_n\} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k,$$

2. A_n **skoro vždy** *(almost always, a. a.)*

$$\liminf_n A_n = \{\omega \in \Omega : \text{ve všech kromě konečně mnoha } A_n\} = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Příklad (Ověřte, že platí:)

1. $\limsup_n A_n,\ \liminf_n A_n \in \mathcal{A}$,

2. $\overline{\limsup_n A_n} = \liminf_n \overline{A_n}$ a tedy $P(\limsup_n A_n) = 1 - P(\liminf_n \overline{A_n})$,

3. $\overline{\liminf_n A_n} = \limsup_n \overline{A_n}$ a tedy $P(\liminf_n A_n) = 1 - P(\limsup_n \overline{A_n})$.

Tvrzení

$$P(\liminf_n A_n) \leq \liminf_{n\to\infty} P(A_n) \leq \limsup_{n\to\infty} P(A_n) \leq P(\limsup_n A_n).$$

Příklad

Pravděpodobnostní prostor $(\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}, \mathcal{A} = 2^\Omega, P)$ $P(\{\omega_1\}) = \frac{1}{6}$, $P(\{\omega_2\}) = \frac{1}{3}$, $P(\{\omega_3\}) = \frac{1}{5}$, $P(\{\omega_4\}) = \frac{3}{10}$. Zvolme posloupnost jevů $A_n = \{\omega_1, \omega_2\}$ pro lichá n, $A_n = \{\omega_2, \omega_3\}$ pro sudá n.

1. Ověřte, že $\liminf_n A_n = \{\omega_2\}$.
2. Ověřte, že $\limsup_n A_n = \{\omega_1, \omega_2, \omega_3\}$.
3. Ověřte, že $P(\liminf_n A_n) = \frac{1}{3}$.
4. Ověřte, že $\liminf_{n\to\infty} P(A_n) = \frac{1}{2}$.
5. Ověřte, že $\limsup_{n\to\infty} P(A_n) = \frac{8}{15}$.
6. Ověřte, že $P(\liminf_n A_n) = \frac{7}{10}$.

Tvrzení (Borelovo-Cantelliho lemma)

Nechť $A_1, A_2, \dots \in \mathcal{A}$.

- *Pokud* $\sum_{n=1}^{\infty} P(A_n) < \infty$, *potom* $P(\limsup_n A_n) = 0$.
- *Pokud* $\sum_{n=1}^{\infty} P(A_n) = \infty$ *a jevy* $A_1, A_2, \dots$ *jsou nezávislé, potom* $P(\limsup_n A_n) = 1$.

Toto tvrzení je zajímavé tím, že říká, že pro nezávislé jevy je $P(\limsup_n A_n)$ rovna buď 0, anebo 1, nikdy žádnému jinému číslu. Používá se při výpočtech pravděpodobností jevů, které mají nastat nekonečně často nebo skoro vždy.

Příklad (Nekonečně mnoho líců/rubů)

Vraťme se k náhodnému pokusu s nekonečným házením mincí a opět pro $n \in \mathbb{N}$ definujeme jevy $H_n = \{(b_1, b_2, \dots) \in \Omega : b_n = 1\}$.

1. Interpretujte jevy $\limsup_n H_n$ a $\liminf_n H_n$.

2. Ověřte, že $\limsup_{n\to\infty} P(H_n) = \frac{1}{2}$.

3. Pomocí B-C lemmatu dokažte, že pravděpodobnost padnutí nekonečně mnoha líců je rovna 1 a pravděpodobnost padnutí konečně mnoha rubů je rovna 0.

4. Pomocí B-C lemmatu dokažte, že pravděpodobnost padnutí nekonečně mnoha rubů je rovna 1 a pravděpodobnost padnutí konečně mnoha líců je rovna 0.

5. Dávají výsledky smysl? Mohou být pravděpodobnosti padnutí nekonečně mnoha líců i nekonečně mnoha rubů obě rovny 1?

Střední hodnota

Definice (Jednoduchá náhodná veličina)

Náhodná veličina X na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je **jednoduchá** *(simple random variable)*, pokud její obor hodnot je konečná množina, $\{X(\omega); \omega \in \Omega\} = \{x_1, \dots, x_n\}$.
Jednoduchou náhodnou veličinu lze zapsat ve tvaru

$$X(\omega) = \sum_{k=1}^{n} x_k \mathbf{1}_{A_k}(\omega),$$

kde $A_k = X^{-1}(\{x_k\})$, $k = 1, \dots, n$, jsou úplné vzory obrazů x_k.

Definice (Střední hodnota jednoduché náhodné veličiny)

Pro jednoduchou náhodnou veličinu definujeme **střední hodnotu** *(expected value, expectation, mean)* předpisem

$$\mathrm{E}(X) = \sum_{k=1}^{n} x_k P(A_k) = \sum_{k=1}^{n} x_k P(X = x_k).$$

Uvedený předpis pro střední hodnotu je ne náhodou podobný vzorci pro výpočet střední hodnoty náhodné veličiny diskrétního typu. Zde je však omezujeme jen na konečné součty.

Všimněme si využití zápisu jednoduché náhodné veličiny

$$\mathrm{E}\left(\sum_{k=1}^{n} x_k \mathbf{1}_{A_k}\right) = \sum_{k=1}^{n} x_k P(A_k).$$

Příklad

Na pravděpodob. prostoru $(\Omega = [0;1], \mathcal{A}, \lambda)$ definujme náhodné veličiny

$$X(\omega) = \begin{cases} 5, & \omega > \frac{1}{3} \\ 3, & \omega \le \frac{1}{3} \end{cases}, \qquad Y(\omega) = \begin{cases} 2, & \omega \in \mathbb{Q} \\ 4, & \omega = \frac{\sqrt{2}}{2} \\ 6, & \omega \le \frac{1}{4}, \omega \notin \mathbb{Q} \\ 8, & \text{jinak} \end{cases}.$$

Ověřte, že $\mathrm{E}(X) = \frac{13}{3}$, $\mathrm{E}(Y) = \frac{15}{2}$.

Vlastnosti jednoduchých náhodných veličin a střední hodnoty:

1. $\mathrm{E}\,(\boldsymbol{1}_A) = P(A)$ pro $A \in \mathcal{A}$,
neboť $\mathrm{E}\,(\boldsymbol{1}_A) = \mathrm{E}\left(0 \cdot \boldsymbol{1}_A^{-1}(\{0\}) + 1 \cdot \boldsymbol{1}_A^{-1}(\{1\})\right) = 0 \cdot P(\overline{A}) + 1 \cdot P(A)$.

2. $\mathrm{E}\,(c) = c$ pro $c \in \mathbb{R}$,
neboť $\mathrm{E}\,(c) = \mathrm{E}\,(c\boldsymbol{1}_\Omega) = c\,P(\Omega) = c$.

3. Střední hodnota je lineární operátor, tzn.

$$\mathrm{E}\,(a\,X + b\,Y) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y) \quad \text{pro náhodné veličiny } X,\ Y \text{ a } a,\ b \in \mathbb{R}.$$

Totiž pokud $X = \sum_{k=1}^{n} x_k \boldsymbol{1}_{A_k}$, $Y = \sum_{l=1}^{m} y_l \boldsymbol{1}_{A_l}$, potom $\mathrm{E}\,(a\,X + b\,Y) =$
$= \mathrm{E}\,(\sum_{k=1}^{n} \sum_{l=1}^{m} (a\,X + b\,Y)\boldsymbol{1}_{A_k \cap B_l}) = \sum_{k=1}^{n} \sum_{l=1}^{m} (a\,x_k + b\,y_l) P(A_k \cap B_l) =$
$= a \sum_{k=1}^{n} x_k\, P(A_k) + b \sum_{l=1}^{m} y_l\, P(B_l) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y)$.

4. $\mathrm{E}(X) \le \mathrm{E}(Y)$ pokud $X \le Y$.
V případě $X(\omega) \le Y(\omega)$ pro všechna $\omega \in \Omega$ totiž máme $Y - X \ge 0$, což dle definice znamená $\mathrm{E}(Y - X) \ge 0$ a z linearity obdržíme $\mathrm{E}(Y) - \mathrm{E}(X) \ge 0$.

5. $|\mathrm{E}(X)| \leq \mathrm{E}(|X|)$,
protože $-|X| \leq X \leq |X|$.

6. $\mathrm{E}(X\,Y) = \mathrm{E}(X)\,\mathrm{E}(Y)$, pokud jsou X, Y nezávislé.
Za nezávislosti totiž $\mathrm{E}(X\,Y) = \sum_{k=1}^{n} \sum_{l=1}^{m} x_k\, y_l\, P(A_k \cap B_l) =$
$= \sum_{k=1}^{n} x_k\, P(A_k) \cdot \sum_{l=1}^{m} y_l\, P(B_l) = \mathrm{E}(X)\,\mathrm{E}(Y)$.

7. Pokud X je jednoduchá náhodná veličina a $h : \mathbb{R} \to \mathbb{R}$ funkce, potom

$$Y = h \circ X = h(X) = \sum_{k=1}^{n} h(x_k) \boldsymbol{1}_{A_k}$$

je opět jednoduchá náhodná veličina a platí

$$\mathrm{E}(Y) = \mathrm{E}[h(X)] = \sum_{k=1}^{n} h(x_k) P(A_k).$$

8. Při speciální volbě $h(x) = [x - \mathrm{E}(X)]^2$ dostáváme $Y = h(X) = [X - \mathrm{E}(X)]^2$ a definujeme **rozptyl** *(variance)* náhodné veličiny X jako

$$\mathrm{Var}(X) = \mathrm{E}(Y) = \mathrm{E}\left([X - \mathrm{E}(X)]^2\right).$$

9. $\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - [\mathrm{E}(X)]^2$ a $0 \le \mathrm{Var}(X) \le \mathrm{E}\left(X^2\right)$.

10. $\mathrm{Var}(a\,X + b) = a^2\,\mathrm{Var}(X)$ pro $a, b \in \mathbb{R}$.

11. $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$,
kde $\mathrm{Cov}(X, Y) = \mathrm{E}\left([X - E(X)][Y - E(Y)]\right) = \mathrm{E}(X\,Y) - \mathrm{E}(X)\,\mathrm{E}(Y)$ je tzv. **kovariance** *(covariance)*.

12. $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$, pokud jsou X, Y nezávislé, neboť v tom případě je $\mathrm{Cov}(X, Y) = 0$.

13. Pokud $\mathrm{Var}(X) > 0$, $\mathrm{Var}(Y) > 0$, definujeme tzv. **korelaci** *(correlation)* předpisem $\mathrm{Cor}(X, Y) = \dfrac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} \in [-1;1]$.

Vztahy platné pro součet, resp. lineární kombinaci, dvou (nezávislých) jednoduchých náhodných veličin je přitom indukcí možné zobecnit na součet, resp. lineární kombinaci, libovolného počtu (nezávislých) jednoduchých náhodných veličin.

Definice (Střední hodnota nezáporné náhodné veličiny)

Střední hodnotu nezáporné náhodné veličiny $X \geq 0$, tzn. $X(\omega) \geq 0$ pro všechna $\omega \in \Omega$, jako supremum

$$\mathrm{E}(X) = \sup\left\{\mathrm{E}(Y) : Y \text{ je jednoduchá náhodná veličina}, Y \leq X\right\}.$$

Tato definice funguje i pro jednoduché náhodné veličiny (stačí zvolit $Y = X$). Funguje však nyní i v případě $\mathrm{E}(X) = \infty$.

Příklad

Na pravděpodobnostním prostoru $(\Omega = [0;1], \mathcal{A}, \lambda)$ definujme náhodnou veličinu X: $X(\omega) = 2^k$ pro $\frac{1}{2^k} \leq \omega < \frac{1}{2^{k-1}}$, $k = 1, 2, \ldots$.

$$\mathrm{E}(X) = \sup\left\{\mathrm{E}(Y) : Y \text{ jednoduchá}, Y \leq X\right\} \geq \mathrm{E}(Y_n) =$$
$$= \sum_{k=1}^{n} 2^k \left(\frac{1}{2^{k-1}} - \frac{1}{2^k}\right) = n \qquad \text{pro každé } n \in \mathbb{N},$$

tedy $\mathrm{E}(X) = \infty$. Náhodné veličiny Y_n jsou přitom jednoduché.

Definice (Monotónní konvergence)

Posloupnost náhodných veličin $X_1, X_2, \ldots$ **konverguje monotónně** k náhodné veličině X, $X_n \nearrow X$,
pokud $X_1 \leq X_2 \leq \cdots$ a $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ pro všechna $\omega \in \Omega$.

Tvrzení (O monotónní konvergenci)

Nechť náhodné veličiny $X_1, X_2, \ldots$ konvergují monotónně k náhodné veličině X, $X_n \nearrow X$, a $\mathrm{E}(X_1) > -\infty$. Potom platí $\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X)$.

*Rovnost $\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X)$ platí i v případě, když $X_n \nearrow X$ **skoro jistě**, tzn. když $X_1 \leq X_2 \leq \cdots$ a $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ pro všechna $\omega \in B$, $B \subseteq \Omega$, $P(B) = 1$.*

Příklad

Na pravděpodobnostním prostoru $(\Omega = [0;1], \mathcal{A}, \lambda)$ definujme posloupnost náhodných veličin $X_n = n\,\boldsymbol{1}_{(0;\,1/n)}$, $n = 1, 2, \ldots$.
Potom $X_n \to 0$, neboť pro $\omega \in [0;1]$ máme $X_n(\omega) = 0$, když $n \geq \frac{1}{\omega}$. Přitom však $\mathrm{E}(X_n) = n\left(\frac{1}{n} - 0\right) + 0\left(1 - \frac{1}{n}\right) = 1$ pro všechna $n = 1, 2, \ldots$. Vidíme, že $\lim\limits_{n\to\infty} \mathrm{E}(X_n) = 1 \neq 0 = \mathrm{E}(0)$. Máme sice $X_n \to 0$, ale nikoliv $X_n \nearrow 0$.

Pro $n = 1, 2, \ldots$ zaveďme funkce

$$\Psi_n(x) = \min\left\{n, \frac{\lfloor x\,2^n \rfloor}{2^n}\right\}, \qquad \text{pro } x \geq 0.$$

Jedná se o směrem dolů na nejbližší násobek $\frac{1}{2^n}$ zaokrouhlené x, navíc useknuté v n, tj. $\Psi_n(x) = n$ pro $x \geq n$. Nakreslete grafy těchto funkcí pro $n = 1, 2, 3$. Ověřte, že obor hodnot funkce Ψ_n je konečný (funkce nabývá právě $1 + n\,2^n$ hodnot a že pro libovolné pevně zvolené $x \geq 0$ je $\Psi_n(x) \nearrow x$.

Tvrzení

Nechť X nezáporná náhodná veličina. Položme $X_n = \Psi_n(X)$. Potom každá náhodná veličina X_n, $n = 1, 2, \ldots$, je jednoduchá, a posloupnost monotónně konverguje k náhodné veličině X, $X_n \nearrow X$.

Ke každé nezáporné náhodné veličině X tedy existuje posloupnost jednoduchých náhodných veličin $X_n \nearrow X$; dokonce takovou posloupnost umíme zkonstruovat.

Všechny vlastnosti střední hodnoty jednoduchých náhodných veličin zůstávají v platnosti i pro nezáporné náhodné veličiny.

K důkazu linearity střední hodnoty nezáporné náhodné veličiny ale nyní potřebujeme umět konstruovat posloupnost monotónně konvergující k takové náhodné veličině. Předchozí věty nám pak např. zaručují, že lze důkaz linearity vést takto:

$$\mathrm{E}(aX + bY) = \lim_{n\to\infty} \mathrm{E}(aX_n + bY_n) = \lim_{n\to\infty} \left[a\mathrm{E}(X_n) + b\mathrm{E}(Y_n)\right] = a\mathrm{E}(X) + b\mathrm{E}(Y).$$

Analogicky se postupuje i u důkazů ostatních vlastností.

Navíc obdržíme dokonce spočetnou linearitu pro nezáporné náhodné veličiny $X_1, X_2, \ldots$:

$$\mathrm{E}\left(\sum_{k=1}^{\infty} X_k\right) = \sum_{k=1}^{\infty} \mathrm{E}(X_k).$$

Definice (Střední hodnota náhodné veličiny)

Střední hodnotu náhodné veličiny definujeme předpisem

$$\mathrm{E}(X) = \mathrm{E}(X^+) - \mathrm{E}(X^-),$$

kde $X^+ = \max\{X, 0\}$ a $X^- = \max\{-X, 0\}$.

- $\mathrm{E}(X^+) = \infty$, $\mathrm{E}(X^-) = \infty \Rightarrow \mathrm{E}(X)$ není definována,
- $\mathrm{E}(X^+) = \infty$, $\mathrm{E}(X^-) < \infty \Rightarrow \mathrm{E}(X) = \infty$,
- $\mathrm{E}(X^+) < \infty$, $\mathrm{E}(X^-) = \infty \Rightarrow \mathrm{E}(X) = -\infty$,
- $\mathrm{E}(X^+) < \infty$, $\mathrm{E}(X^-) < \infty \Rightarrow \mathrm{E}(X) \in \mathbb{R}$.

Náhodné veličiny X^+ a X^- jsou nezáporné. Pro obecnou náhodnou veličinu X tedy střední hodnotu definujeme jako rozdíl středních hodnot nezáporných veličin X^+ a X^- podle předchozích definic.

Definice (Momenty)

k-tý **moment** náhodné veličiny je $\mathrm{E}\left(X^k\right)$, $\qquad k \in \mathbb{N}$.
Říkáme, že k-tý moment je konečný nebo že X má konečný k-tý moment, pokud $\mathrm{E}\left(\left|X^k\right|\right) < \infty$.

Protože $|x|^{k-1} \leq \max\{|x|^k, 1\} \leq |x|^k + 1$ pro všechna $x \in \mathbb{R}$, platí, že pokud náhodná veličina X má konečný k-tý moment, potom má konečný i každý l-tý moment, $l = 1, 2, \ldots, k-1$.

Vlastnosti střední hodnoty se přenáší z vlastností střední hodnoty nezáporné náhodné veličiny. U součtu i součinu náhodných veličin však nyní musíme být *opatrní* a známé vztahy definujeme pouze pro náhodné veličiny s konečným prvním momentem.

Integrál podle míry

Definice (Integrál podle míry)

Pro náhodnou veličinu X na prostoru s mírou $(\Omega, \mathcal{A}, \mu)$ definujeme **integrál podle míry** μ, $\int_\Omega X \,\mathrm{d}\mu = \int X \,\mathrm{d}\mu$, následovně:

1. pro nezápornou jednoduchou náhodnou veličinu $X = \sum_{k=1}^{n} x_k \mathbf{1}_{A_k}$:
$$\int_\Omega X \,\mathrm{d}\mu = \sum_{k=1}^{n} x_k \,\mu(A_k).$$
2. pro nezápornou náhodnou veličinu:
$$\int_\Omega X \,\mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega X_n \,\mathrm{d}\mu$$
pro libovolnou posloupnost nezáporných náhodných veličin $X_n \nearrow X$.
3. pro obecnou náhodnou veličinu:
$$\int_\Omega X \,\mathrm{d}\mu = \int_\Omega X^+ \,\mathrm{d}\mu - \int_\Omega X^- \,\mathrm{d}\mu,$$
pokud alespoň jeden z integrálů vpravo existuje;
přitom $X^+ = \max\{X, 0\}$ a $X^- = \max\{-X, 0\}$.

Definice

Náhodná veličina je **integrovatelná**, pokud $\int_\Omega X^+ \, \mathrm{d}\mu < \infty$ a $\int_\Omega X^- \, \mathrm{d}\mu < \infty$, tj. pokud $\int_\Omega X \, \mathrm{d}\mu$ existuje a je konečný, $\left|\int_\Omega X \, \mathrm{d}\mu\right| < \infty$.

Pro $A \in \mathcal{A}$ definujeme integrál na měřitelné množině A předpisem

$$\int_A X \, \mathrm{d}\mu = \int_\Omega (X \mathbf{1}_A) \, \mathrm{d}\mu.$$

Pro pravděpodobnostní míru, $\mu = P$, máme $\int_\Omega X \, \mathrm{d}\mu = \int_\Omega X \, \mathrm{d}P = \mathrm{E}(X)$.

Vlastnosti integrálu:

1. Pro nezáporné náhodné veličiny X, Y platí

$$\int (X + Y)\,\mathrm{d}\mu = \int X\,\mathrm{d}\mu + \int X\,\mathrm{d}\mu.$$

2. Pokud existují integrály $\int X\,\mathrm{d}\mu$, $\int Y\,\mathrm{d}\mu$ a $\int X\,\mathrm{d}\mu + \int Y\,\mathrm{d}\mu$, potom existuje integrál $\int (X + Y)\,\mathrm{d}\mu$ a platí

$$\int (X + Y)\,\mathrm{d}\mu = \int X\,\mathrm{d}\mu + \int X\,\mathrm{d}\mu.$$

3. Nechť $A, B \in \mathcal{A}$, $A \cap B = \varnothing$. Pokud existuje integrál $\int_{A\cup B} X\,\mathrm{d}\mu$, existují i integrály $\int_A X\,\mathrm{d}\mu$ a $\int_B X\,\mathrm{d}\mu$ a platí

$$\int_{A\cup B} X\,\mathrm{d}\mu = \int_A X\,\mathrm{d}\mu + \int_B X\,\mathrm{d}\mu.$$

4. Nechť $A, B \in \mathcal{A}$, $A \cap B = \emptyset$. Pokud existují integrály $\int_A X \, \mathrm{d}\mu$, $\int_B X \, \mathrm{d}\mu$ a $\int_A X \, \mathrm{d}\mu + \int_B X \, \mathrm{d}\mu$, potom existuje integrál $\int_{A \cup B} X \, \mathrm{d}\mu$ a platí

$$\int_{A \cup B} X \, \mathrm{d}\mu = \int_A X \, \mathrm{d}\mu + \int_B X \, \mathrm{d}\mu.$$

5. Pokud existuje integrál $\int X \, \mathrm{d}\mu$, potom existují integrály $\int_A X \, \mathrm{d}\mu$, $\int_{\overline{A}} X \, \mathrm{d}\mu$ pro libovolnou $A \in \mathcal{A}$ a platí

$$\int X \, \mathrm{d}\mu = \int_A X \, \mathrm{d}\mu + \int_{\overline{A}} X \, \mathrm{d}\mu.$$

Pokud je X integrovatelná (na Ω), potom je integrovatelné i na libovolné $A \in \mathcal{A}$.

6. Pokud existuje integrál $\int X \, \mathrm{d}\mu$ a $c \in \mathbb{R}$, potom existuje integrál $\int_A c X \, \mathrm{d}\mu$ a platí

$$\int c X \, \mathrm{d}\mu = c \int X \, \mathrm{d}\mu.$$

7. Pokud je X nezáporná náhodná veličina, potom $\int X \, \mathrm{d}\mu \geq 0$.

8. Pokud $X \leq Y$ a existují integrály $\int X \, \mathrm{d}\mu$ a $\int Y \, \mathrm{d}\mu$, potom $\int X \, \mathrm{d}\mu \leq \int Y \, \mathrm{d}\mu$.

9. Pokud $X \leq Y$ a $\int X^- \, \mathrm{d}\mu < \infty$, potom existuje integrál $\int Y \, \mathrm{d}\mu$ a $\int X \, \mathrm{d}\mu \leq \int Y \, \mathrm{d}\mu$.

10. Pokud $X \leq Y$ a $\int Y^+ \, \mathrm{d}\mu < \infty$, potom existuje integrál $\int X \, \mathrm{d}\mu$ a $\int X \, \mathrm{d}\mu \leq \int Y \, \mathrm{d}\mu$.

11. Pokud existuje integrál $\int X \, \mathrm{d}\mu$, potom $\left|\int X \, \mathrm{d}\mu\right| \leq \int |X| \, \mathrm{d}\mu$.

12. Pokud $X = Y$ skoro všude a existuje integrál $\int X \, \mathrm{d}\mu$, potom existuje integrál $\int X \, \mathrm{d}\mu$ a platí $\int X \, \mathrm{d}\mu = \int Y \, \mathrm{d}\mu$.

13. X je integrovatelná $\Leftrightarrow$ $|X|$ je integrovatelná.

14. Pokud $|X| \leq Y$ a existuje integrál $\int Y \, \mathrm{d}\mu$, potom je X integrovatelná.

Příklad

V pravděpodobnostním prostoru $(\Omega = [0;1], \mathcal{A}, \lambda)$ s Lebesgueovou mírou uvažujme náhodnou veličinu $X(\omega) = \mathbf{1}_{\overline{\mathbb{Q}}}(\omega)$. Tato náhodná veličina (měřitelná funkce) není Riemannovsky integrovatelná. Je však Lebesgueovsky integrovatelná a $\mathrm{E}(X) = \int X \,\mathrm{d}P = 1$.

Příklad

Uvažujme prostor s mírou $(\Omega = \{-2; -1; 3; -4\}, \mathcal{A} = 2^{\Omega}, \mu)$, kde $\mu(\{-2\}) = 2$, $\mu(\{-1\}) = 1$, $\mu(\{3\}) = 3$, $\mu(\{7\}) = 7$, a náhodnou veličinu X, $X(-2) = X(-1) = -1$, $X(3) = X(7) = 1$.
Spočítejte integrál $\int_A X \,\mathrm{d}\mu$ pro $A = \{-2; 3; 7\}$. [8]

Příklad

Uvažujme prostor s Lebesgueovou mírou $(\Omega = (-5;5), \mathcal{A} = \mathcal{B}_{(-5;5)}, \lambda)$ na intervalu $(-5;5)$ a náhodnou veličinu X, $X(\omega) = \begin{cases} \frac{1}{2}, & \omega \in (-5;2) \\ \frac{1}{3}, & \omega = 2 \\ 1, & \omega \in (2;3] \\ 0, & \omega \in (3;5) \end{cases}$.

Spočítejte integrál $\int_A X \,\mathrm{d}\lambda$ pro $A = [-1;4]$. $\left[\frac{5}{2}\right]$

Nerovnosti pro náhodné veličiny

Tvrzení (Markovova nerovnost)

Nechť X *je nezáporná náhodná veličina na pravděpodobnostním prostoru* $(\Omega, \mathcal{A}, P)$*. Potom pro všechna* $a > 0$ *platí*

$$P(X \geq a) \leq \frac{\mathrm{E}(X)}{a}.$$

Tvrzení (Čebyševova nerovnost)

Nechť X *je náhodná veličina s konečnou střední hodnotou na pravděpodobnostním prostoru* $(\Omega, \mathcal{A}, P)$*. Potom pro všechna* $a > 0$ *platí*

$$P\big(|X - \mathrm{E}(X)| \geq a\big) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

Tvrzení (Cauchyova-Schwarzova nerovnost)

Nechť X, Y je náhodné veličiny s konečnými druhými momenty na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. Potom platí

$$\left[\mathrm{E}\left(|X\,Y|\right)\right]^2 \leq \mathrm{E}(X^2)\,\mathrm{E}(Y^2).$$

Tvrzení (Jensenova nerovnost)

Nechť X je náhodná veličina s konečnou střední hodnotou na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. Nechť $h : \mathbb{R} \to \mathbb{R}$ je libovolná konvexní funkce. Potom platí

$$\mathrm{E}\big(h(X)\big) \geq h\big(\mathrm{E}(X)\big).$$

Příklad

Pro náhodnou veličinu $X \sim \mathrm{Rs}([0;4])$ spočítejte pro $a = 0;1;2;3;4$ pravděpodobnosti $P(X \geq a)$ a zároveň je odhadujte pomocí Markovovy nerovnosti. Dále spočítejte pravděpodobnosti $P(|X-2| \geq a-2)$ a zároveň je odhadujte pomocí Čebyševovy nerovnosti.

Příklad

Nechť X je náhodná veličina s konečnou střední hodnotou a rozptylem σ^2 na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. Dokažte, že potom pro všechna $c > 0$ platí $P\big(|X - \mathrm{E}(X)| \geq c\,\sigma\big) \leq \frac{1}{c^2}$.

Příklad

Ověřte, že pro náhodnou veličinu X, $P(X = a) = P(X = -a) = \frac{1}{2}$ pro $a > 0$, dává Čebyševova nerovnost přesný odhad pravděpodobnosti, tj. že $1 = P\big(|X - \mathrm{E}(X)| \geq a\big) \leq 1$.

Příklad

Nechť $\mathrm{E}\left(2^X\right) = 4$. Dokažte, že $P(X \geq 3) \leq \frac{1}{2}$.

Konvergence náhodných veličin

Definice (Bodová konvergence)

Posloupnost náhodných veličin $X_1, X_2, \ldots$ na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ **konverguje bodově** *(pointwise convergence)* k náhodné veličině X, $X_n \longrightarrow X$, $\lim\limits_{n\to\infty} X_n = X$, pokud

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \qquad \text{pro každé } \omega \in \Omega.$$

Definice (Konvergence skoro jistě)

Posloupnost náhodných veličin $X_1, X_2, \ldots$ na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ **konverguje skoro jistě (s. j.)** *(almost surely, a. s.)* k náhodné veličině X, $X_n \xrightarrow{\text{s.j.}} X$, pokud

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1,$$

$$\text{t. j. pokud} \qquad P\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

Definice (Konvergence podle pravděpodobnosti)

Posloupnost náhodných veličin $X_1, X_2, \ldots$ na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ **konverguje podle pravděpodobnosti** *(in probability)* k náhodné veličině X, $X_n \xrightarrow{P} X$, pokud

$$\forall \varepsilon > 0 : \qquad \lim_{n\to\infty} P\Big(|X_n - X| \geq \varepsilon\Big) = 0,$$

t. j. pokud $\quad \forall \varepsilon > 0 : \quad P\Big(\{\omega \in \Omega : |X_n - X| \geq \varepsilon\}\Big) \longrightarrow 0.$

Definice (L^p konvergence, $p = 1, 2, \dots$)

Posloupnost náhodných veličin $X_1, X_2, \dots$ na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ **konverguje v p-tém momentu** *(in the p-th mean)* (v L^p) k náhodné veličině X, $X_n \xrightarrow{L^p} X$, pokud

$$\lim_{n\to\infty} \mathrm{E}\left(|X_n - X|^p\right) \longrightarrow 0.$$

Pro $p = 1$ jde o tzv. **limitu podle středu** *(limit in the mean)*:

$$\lim_{n\to\infty} \mathrm{E}\left(|X_n - X|\right) \longrightarrow 0.$$

Pro $p = 2$ jde o tzv. **limitu podle kvadratického středu** *(limit in the mean-square)*:

$$\lim_{n\to\infty} \mathrm{E}\left(|X_n - X|^2\right) \longrightarrow 0.$$

Definice (Konvergence skoro všude)

Posloupnost náhodných veličin $X_1, X_2, \dots$ na prostoru s mírou $(\Omega, \mathcal{A}, \mu)$ **konvergujе skoro všude (s. v.)** *(almost everywhere, a. e.)* k náhodné veličině X, $X_n \xrightarrow{\text{s.v.}} X$, pokud

$$\mu\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\right\}\right) = 0,$$

t. j. pokud $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ pro všechna $\omega \in \Omega \setminus N,\ \mu(N) = 0.$

Množina N, $\mu(N) = 0$, je tzv. **μ-nulová** množina.

Pokud je míra pravděpodobnostní, $\mu = P$, potom je **s.v.** konvergence vlastně konvergencí **s.j.**.

Definice (Konvergence podle míry)

Posloupnost náhodných veličin $X_1, X_2, \ldots$ na prostoru s mírou $(\Omega, \mathcal{A}, \mu)$ **konvergiuje podle míry** *(in measure)* k náhodné veličině X, $X_n \xrightarrow{\mu} X$, pokud

$$\forall \varepsilon > 0 : \qquad \lim_{n\to\infty} \mu\Big(|X_n - X| \geq \varepsilon\Big) = 0,$$

$$\text{t. j. pokud} \quad \forall \varepsilon > 0 : \quad \mu\Big(\{\omega \in \Omega : |X_n - X| \geq \varepsilon\}\Big) \longrightarrow 0.$$

Pokud je míra pravděpodobnostní, $\mu = P$, potom je konvergence podle míry vlastně konvergencí podle pravděpodobnosti.

Vlastnosti:

1. Pokud $X_n \longrightarrow X$, potom $X_n \xrightarrow{\text{s.j.}} X$.
2. Pokud pro každé $\varepsilon > 0$ je $P\left(|X_n - X| \geq \varepsilon \text{ nekonečně často}\right) = 0$, potom $X_n \xrightarrow{\text{s.j.}} X$.
3. Pokud pro každé $\varepsilon > 0$ je $\sum_{n=1}^{\infty} P\left(|X_n - X| \geq \varepsilon\right) < \infty$, potom $X_n \xrightarrow{\text{s.j.}} X$.
4. Pokud je míra μ konečná a $X_n \xrightarrow{\text{s.v.}} X$, potom $X_n \xrightarrow{\mu} X$; speciálně: pokud $X_n \xrightarrow{\text{s.j.}} X$, potom $X_n \xrightarrow{P} X$.
5. Pokud $X_n \xrightarrow{\mu} X$ a $X_n \xrightarrow{\mu} Y$, potom $\mu(X \neq Y) = 0$.
6. Pokud $X_n \xrightarrow{\mu} X$, potom existuje podposloupnost $\{X_{n_k}\} \subseteq \{X_n\}$ a náhodná veličina Y tak, že $X_{n_k} \xrightarrow{\text{s.v.}} Y$, a $\mu(X \neq Y) = 0$.
7. Pokud $X_n \xrightarrow{\text{s.v.}} X$ a $X_n \xrightarrow{\text{s.v.}} Y$, potom $\mu(X \neq Y) = 0$.
8. Pokud $X_n \xrightarrow{\text{s.j.}} X$ a $X_n \xrightarrow{\text{s.j.}} Y$, potom $P(X = Y) = 1$.
9. Pokud $X_n \xrightarrow{\mu} X$ a $X_n \xrightarrow{\mu} Y$, potom $\mu(X \neq Y) = 0$.
10. Pokud $X_n \xrightarrow{L^2} X$, potom $X_n \xrightarrow{P} X$.
11. Pokud $X_n \xrightarrow{\text{s.j.}} X$ a $X_n \xrightarrow{L^1} Y$, potom $P(X = Y) = 1$.

Příklad

Uvažujte posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$,
$P(X_n = 1) = \frac{1}{n}, P(X_n = 0) = 1 - \frac{1}{n}$.
Ověřte, že $X_n \xrightarrow{P} 0$, ale $X_n \overset{\text{s.j.}}{\not\longrightarrow} 0$. Využijte přitom Borelovo-Cantelliho lemma k výpočtu $P(X_n = 1 \text{ nekonečně často}) = 1$.

Příklad

Na pravděpodobnostním prostoru $(\Omega = [0;1], \mathcal{A}, \lambda)$ s Lebesgueovou mírou uvažujte posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$,

$$X_1 = \boldsymbol{1}_{[0;\frac{1}{2})}, \quad X_2 = \boldsymbol{1}_{[\frac{1}{2};1]},$$
$$X_3 = \boldsymbol{1}_{[0;\frac{1}{4})}, \quad X_4 = \boldsymbol{1}_{[\frac{1}{4};\frac{1}{2})}, \quad X_5 = \boldsymbol{1}_{[\frac{1}{2};\frac{3}{4})}, \quad X_6 = \boldsymbol{1}_{[\frac{3}{4};1]},$$
$$X_7 = \boldsymbol{1}_{[0;\frac{1}{8})}, \quad \dots, \quad X_{14} = \boldsymbol{1}_{[\frac{7}{8};1]}, \quad \dots$$

Ověřte, že $X_n \xrightarrow{\lambda} 0$, ale $X_n \overset{\text{s.j.}}{\not\longrightarrow} 0$.

Příklad

Uvažujme Lebesgueovu míru $(\Omega = \mathbb{R}, \mathcal{A} = \mathcal{B}, \lambda)$ na $\mathbb{R}$ a posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$, $X_n = \boldsymbol{1}_{(n,\,\infty)}$.

Ověřte, že $X_n \longrightarrow 0$ a $X_n \xrightarrow{\text{s.v.}} 0$, ale $X_n \overset{\lambda}{\not\longrightarrow} 0$.
Lebesgueova míra λ na $\mathbb{R}$ totiž není konečná.

Příklad

Uvažujme prostor $(\Omega = \mathbb{N}, \mathcal{A} = 2^{\Omega}, \mu)$ s čítací mírou μ a posloupnost náhodných veličin $\{X_n\}_{n=1}^{\infty}$, $X_n = \boldsymbol{1}_{\{1,\,\dots,\,n\}}$.

Ověřte, že $X_n \longrightarrow 1$ a $X_n \xrightarrow{\text{s.v.}} 1$, ale $X_n \overset{\mu}{\not\longrightarrow} 1$.
Čítací míra μ na $\mathbb{N}$ totiž není konečná.

Zákony velkých čísel

Tvrzení (Slabý zákon velkých čísel)

Nechť $X_1, X_2, \ldots$ *je posloupnost náhodných veličin se stejnou střední hodnotou* $\mu = \mathrm{E}(X_k)$ *a se stejným rozptylem* $\sigma^2 = \mathrm{Var}(X_k)$*, který je shora omezený. Potom pro každé* $\varepsilon > 0$ *platí*

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \geq \varepsilon\right) = 0,$$

tzn. posloupnost částečných průměrů $\overline{S_n} = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$ *konverguje pro* $n \to \infty$ *v pravděpodobnosti k* μ*,* $\overline{S_n} \xrightarrow{P} \mu$*.*

Tvrzení (Silný zákon velkých čísel)

Nechť $X_1, X_2, \ldots$ *je posloupnost náhodných veličin se stejnou střední hodnotou* $\mu = \mathrm{E}(X_k)$ *a se stejným čtvrtým centrálním momentem* $\mathrm{E}\left[(X_k - \mu)^4\right]$*. Potom platí*

$$P\left(\lim_{n\to\infty} \overline{S_n} = \mu\right) = 1, \qquad \textit{tj.} \qquad \overline{S_n} \xrightarrow{\text{s.j.}} \mu.$$

Definice (i.d. a i.i.d. náhodné veličiny)

Náhodné veličiny v systému $\{X_k\}_{k\in I}$ jsou **i.d.** *(identically distributed)*, pokud pro libovolnou borelovsky měřitelnou funkci $h : \mathbb{R} \to \mathbb{R}$ střední hodnota $\mathrm{E}\left[h(X_k)\right]$ nezávisí na volbě $k \in I$, tj. pokud pro všechna $k \in I$ stejná.

Náhodné veličiny v systému $\{X_k\}_{k\in I}$ jsou **i.i.d.** *(independent identically distributed)*, pokud jsou **i.d.** a nezávislé.

Tvrzení (Slabý a silný zákon velkých čísel pro i.i.d.)

*Nechť $X_1, X_2, \ldots$ je posloupnost **i.i.d.** náhodných veličin s konečnou (stejnou) střední hodnotou. Potom platí:*

- $\overline{S_n} \xrightarrow{P} \mu$,
- $\overline{S_n} \xrightarrow{\text{s.j.}} \mu$.

Rozdělení pravděpodobnosti

Definice (Rozdělení pravděpodobnosti)

Nechť X je náhodná veličina na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$. **Rozdělení pravděpodobnosti** *(probability distribution, law)* náhodné veličiny X je zobrazení $P_X : \mathcal{B} \to [0;1]$, dané předpisem

$$P_X(B) = P(X \in B) = P\left(X^{-1}(B)\right) = P(\{\omega \in \Omega : X(\omega) \in B\}) \qquad \text{pro } B \in \mathcal{B}.$$

Trojice $(\mathbb{R}, \mathcal{B}, P_X)$ je pravděpodobnostní prostor na $\mathbb{R}$.

Definice (Distribuční funkce)

Distribuční funkce *(cumulative distribution function)* náhodné veličiny X je funkce $F_X : \mathbb{R} \to [0;1]$, dané předpisem

$$F(x) = F_X(x) = P_X((-\infty, x]) = P(X \leq x) \qquad \text{pro } x \in \mathbb{R}.$$

Tvrzení (Vlastnosti distribuční funkce)

Pro distribuční funkci F_X náhodné veličiny X platí:

- F_X *je neklesající,*
- F_X *je zprava spojitá,*
- $\lim_{x \to -\infty} F_X(x) = 0, \qquad \lim_{x \to \infty} F_X(x) = 1.$

Tvrzení (Change of variable)

Nechť X je náhodná veličina na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ a $h : \mathbb{R} \to \mathbb{R}$ je borelovsky měřitelná funkce. Potom platí

$$\mathrm{E}_P\big(h(X)\big) = \int_\Omega h(X)\,\mathrm{d}P = \int_\mathbb{R} h(x)\,\mathrm{d}P_X = \int_\mathbb{R} h(x)\,\mathrm{d}F_X = \mathrm{E}_{P_X}(h),$$

pokud alespoň jedna střední hodnota existuje.

Věta říká, že střední hodnotu transformované náhodné veličiny $h(X)$ vzhledem k pravděpodobnosti P lze počítat jako střední hodnotu funkce h vzhledem k rozdělení pravděpodobnosti P_X. Borelovsky měřitelná funkce h je totiž vlastně náhodnou veličinou na měřitelném prostoru $(\mathbb{R}, \mathcal{B})$.

Tvrzení (Důsledek)

Nechť X, Y *jsou náhodné veličiny na obecně různých pravděpodobnostních prostorech. Potom* $P_X = P_Y$*, právě když* $\mathrm{E}\big(h(X)\big) = \mathrm{E}\big(h(Y)\big)$ *pro všechny borelovsky měřitelné funkce* $h : \mathbb{R} \to \mathbb{R}$*, pro které alespoň jedna ze středních hodnot existuje.*

Tvrzení (Důsledek)

Nechť X, Y *jsou náhodné veličiny na stejném pravděpodobnostním prostoru a nechť* $P(X = Y) = 1$*.*
Potom platí $\mathrm{E}\big(h(X)\big) = \mathrm{E}\big(h(Y)\big)$ *pro všechny borelovsky měřitelné funkce* $h : \mathbb{R} \to \mathbb{R}$*, pro které alespoň jedna ze středních hodnot existuje.*

Tvrzení

Nechť $P_1, P_2, \dots$ *je posloupnost rozdělení pravděpodobnosti na* $(\mathbb{R}, \mathcal{B})$*.*
Potom existuje pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$ *a na něm náhodné veličiny* $X_1, X_2, \dots$ *tak, že* $P_{X_n} = P_n$*,* $n = 1, 2, \dots$*.*

Uvažujme náhodnou veličinu X, $P(X = c) = 1$ pro nějaké pevně zvolené $x \in \mathbb{R}$. Rozdělení pravděpodobnosti P_X nazýváme **degenerované (koncentrované)** v bodě c, a značíme $P_X = \delta_c$. Platí

$$P_X(B) = \delta_c(B) = \mathbf{1}_B(c) = \begin{cases} 1, & c \in B \\ 0, & c \notin B \end{cases}, \qquad \text{pro } B \in \mathbb{B}.$$

Pro distribuční funkci takové náhodné veličiny s δ_c rozdělením pravděpodobnosti platí

$$F_X(x) = \begin{cases} 0, & x < c \\ 1, & x \geq c \end{cases}, \qquad \text{pro } x \in \mathbb{R}.$$

Spočítáme střední hodnotu borelovské transformace h:

$$\mathrm{E}_P\big(h(X)\big) = \mathrm{E}_{P_X}(h) = \int_{\mathbb{R}} h(x)\,\mathrm{d}\delta_c = h(c),$$

tzn. střední hodnota je vlastně vyhodnocení funkce h v bodě c.

Nyní již lehce dokážeme známé vlastnosti:

$$\mathrm{E}(X) = c, \qquad \mathrm{E}(X^2) = c^2, \qquad \mathrm{Var}(X) = 0.$$

Tvrzení

Nechť $P_{X_1}, P_{X_2}, \ldots$ *je nejvýše spočetná posloupnost rozdělení pravděpodobnosti a* $a_1, a_2, \dots \geq 0$ *nezáporné konstanty. Potom pro* $P_X = \sum_{n=1}^{\infty} a_n P_{X_n}$ *a libovolnou borelovsky měřitelnou funkci* $h : \mathbb{R} \to \mathbb{R}$ *platí*

$$\int_{\mathbb{R}} h \, \mathrm{d}P_X = \sum_{n=1}^{\infty} a_n \int_{\mathbb{R}} h \, \mathrm{d}P_{X_n}.$$

Pokud navíc $\sum_{n=1}^{\infty} a_n = 1$*, je* P_X *rozdělení pravděpodobnosti a platí*

$$\mathrm{E}_{P_X}(h) = \sum_{n=1}^{\infty} a_n \, \mathrm{E}_{P_{X_n}}(h).$$

Věta říká, že pokud je P_X konvexní kombinací nejvýše spočetně mnoha rozdělení pravděpodobnosti, lze střední hodnotu vzhledem k P_X počítat jako stejnou konvexní kombinaci středních hodnot vzhledem k jednotlivým rozdělením pravděpodobnosti.

Uvažujme náhodnou veličinu $X \sim \mathrm{Po}(\lambda)$. Víme, že

$$P(X = x) = \mathrm{e}^{-\lambda} \frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ \ldots$$

Pro každou $B \in \mathcal{B}$ nyní lze psát

$$P_X(B) = \sum_{x \in B} \mathrm{e}^{-\lambda} \frac{\lambda^x}{x!} = \sum_{x=0}^{\infty} \underbrace{\mathrm{e}^{-\lambda} \frac{\lambda^x}{x!}}_{a_x} \delta_x(B).$$

Máme konvexní kombinaci $P_X = \sum_{x=1}^{\infty} a_x \, \delta_x$ a podle předchozích vět tedy dostáváme známý vztah

$$\mathrm{E}_{P_X}(h) = \sum_{x=1}^{\infty} a_x \, \mathrm{E}_{\delta_x}(h) = \sum_{x=1}^{\infty} \mathrm{e}^{-\lambda} \frac{\lambda^x}{x!} \, h(x).$$

Uvědomte si, že uvedený postup lze použít pro každou náhodnou veličinu diskrétního typu.

Pro libovolnou borelovsky měřitelnou funkci $f : \mathbb{R} \to \mathbb{R}$ s vlastnostmi $f(x) \geq 0$, $\int_{\mathbb{R}} f(x)\,\mathrm{d}\lambda(x) = 1$, kde λ je lebesgueova míra na $\mathbb{R}$, definujeme rozdělení pravděpodobnosti

$$P_Y(B) = \int_{\mathbb{R}} f(x)\,\mathbf{1}_B(x)\,\mathrm{d}\lambda(x), \qquad B \in \mathcal{B}.$$

Taková funkce f se nazývá **hustota (rozdělení) pravděpodobnosti (vzhledem k Lebesgueově míře)** *(pdf, probability density function)*.

Často se zkráceně píše

$$P_Y(B) = \int_B f = \int_B f\,\mathrm{d}\lambda,$$

dokonce lze symbolicky psát $\mathrm{d}P_Y = f\,\mathrm{d}\lambda$, resp. $f = \dfrac{\mathrm{d}P_Y}{\mathrm{d}\lambda}$.

Později zjistíme, že hustota pravděpodobnosti f je vlastně Radonovou-Nikodymovou derivací rozdělení pravděpodobnosti P_Y vzhledem k Lebesgueově míře λ a odtud plyne i pojmenování, že Y je náhodnou veličinou absolutně spojitého typu.

Tvrzení

Nechť rozdělení pravděpodobnosti P_Y má hustotu f vzhledem k Lebesgueově míře λ. Pro libovolnou borelovsky měřitelnou funkci $h : \mathbb{R} \to \mathbb{R}$ potom platí

$$\mathrm{E}_{P_X}\big(h\big) = \int_{\mathbb{R}} h \,\mathrm{d}P_X = \int_{\mathbb{R}} f(x)\, h(x)\, \mathrm{d}\lambda(x),$$

pokud alespoň jedna strana rovnosti existuje.

Integrál vpravo v tvrzení věty je přitom Lebesgueův. Z teorie míry je známo, že existuje-li Riemannův integrál, oba integrály mají stejnou hodnotu,

$$\int_{B} f(x)\, h(x)\, \mathrm{d}\lambda(x) = \int_{B} f(x)\, h(x)\, \mathrm{d}x.$$

Riemannův integrál existuje např. pokud je množina B konečným sjednocením intervalů, přesněji pokud hranice množiny B má nulovou míru, $\lambda(\partial B) = 0$.

Příklad

Uvažujte pravděpodobnostní prostor $(\Omega = [0;1], \mathcal{A}, \lambda)$ s Lebesgueovou mírou na intervalu $[0;1]$ a náhodnou veličinu X,

$$X(\omega) = \begin{cases} 1, & 0 \le \omega < \frac{1}{4} \\ 2\omega^2, & \frac{1}{4} \le \omega < \frac{3}{4} \\ \omega^2, & \frac{3}{4} \le \omega \le 1 \end{cases}.$$

Spočítejte pravděpodobnosti $\lambda(X \in [0;1])$ a $\lambda(X \in [\frac{1}{2};1])$.

$$\left[\frac{1+2\sqrt{2}}{4} \doteq 0{,}957; \frac{\sqrt{2}}{2} \doteq 0{,}707\right]$$

Příklad

Nechť $P_X = \frac{1}{4}\delta_1 + \frac{1}{4}\delta_2 + \frac{1}{2}P_Y$, kde $Y \sim \mathrm{N}(0;1)$.
Ověřte výpočtem, že $\mathrm{E}(X) = \frac{3}{4}$, $\mathrm{E}(X^2) = \frac{7}{4}$, $\mathrm{Var}(X) = \frac{19}{16}$.

Příklad

Nechť $Y \sim \mathrm{N}(0;1)$ s hustotou pravděpodobnosti $\varphi(y) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}y^2\right]$.

Pomocí předchozí věty odvoďte známé vzorce $\mathrm{E}(Y) = \int\limits_{\infty}^{\infty} y\,\varphi(y)\,\mathrm{d}y = 0$,

$\mathrm{E}(Y^2) = \int\limits_{\infty}^{\infty} y^2 \varphi(y)\,\mathrm{d}y = 1$,

Příklad

Uvažujte dvě nezávislé náhodné veličiny Y, Z, $Y \sim \mathrm{N}(0;1)$,
$P(Z = 0) = P(Z = 1) = \frac{1}{2}$. Nechť $X = Y Z$.
Zapište rozdělení pravděpodobnosti P_X. $\left[P_X = \frac{1}{2}\delta_0 + \frac{1}{2}P_Y\right]$

Příklad

Pro $X \sim \mathrm{Po}(\lambda = 5)$ odvoďte a číselně spočítejte $\mathrm{E}(X)$, $\mathrm{E}(X^2)$, $\mathrm{Var}(X)$ a $\mathrm{E}(3^X)$.
$\left[5 = \lambda; 30; 5 = \lambda; \mathrm{e}^{10} \doteq 22026{,}47\right]$

Příklad

Spočítejte $\mathrm{E}(X)$, $\mathrm{E}(X^2)$ a $\mathrm{Var}(X)$ pro $P_X = \frac{1}{2}\delta_1 + \frac{1}{2}\lambda$, kde λ je Lebesgueova míra na intervalu $[0;1]$.

$$\left[\tfrac{3}{4}; \tfrac{2}{3}; \tfrac{5}{48}\right]$$

Příklad

Spočítejte $\mathrm{E}(X)$, $\mathrm{E}(X^2)$ a $\mathrm{Var}(X)$ pro $P_X = \frac{1}{3}\delta_2 + \frac{2}{3}P_Y$, kde $Y \sim \mathrm{N}(0;1)$.

$$\left[\tfrac{2}{3}; 2; \tfrac{14}{9}\right]$$

Příklad

Uvažujte dvě nezávislé náhodné veličiny Y, Z, $Y \sim \mathrm{N}(0;1)$, $P(Z=-1) = P(Z=1) = \frac{1}{2}$. Nechť $X = YZ$.

1. Zapište rozdělení pravděpodobnosti P_X. $[P_X = P_Y,$ tj. $X \sim \mathrm{N}(0;1)]$
2. Ověřte, že $P(|X| = |Y|) = 1$.
3. Ověřte, že X, Y nejsou nezávislé.
4. Ověřte, že $\mathrm{Cov}(X, Y) = 0$.
5. Zamyslete se nad výsledky předchozích dvou podúkolů.

Lebesgueova-Stieljesova míra, Lebesgueův-Stieljesův integrál

Tvrzení

Nechť $F : \mathbb{R} \to \mathbb{R}$ je neklesající a zprava spojitá funkce a $\mathcal{J}$ je systém intervalů $\mathcal{J} = \{(a, b];\ a, b, \in \mathbb{R}, a \leq b\}$. Potom množinová funkce $\mu_F : \mathcal{J} \to \mathbb{R}$ definovaná předpisem $\mu_F\big((a, b]\big) = F(b) - F(a)$ je σ-aditivní.

Tvrzení

Nechť $F : \mathbb{R} \to \mathbb{R}$ je neklesající a zprava spojitá funkce. Potom existuje právě jedna míra μ_F definovaná na $\mathcal{B}$, taková, že $\mu_F\big((a, b]\big) = F(b) - F(a)$ pro $a, b \in \mathbb{R}$, $a \leq b$.

Definice

Míra $\mu_F : \mathcal{B} \to \mathbb{R}$ z předchozí věty se nazývá **Lebesgueova-Stieltjesova míra indukovaná funkcí F**.

Lebesgueova míra λ je speciálním případem Lebesgueovy-Stieltjesovy míry μ_F pro identickou funkci F, $F(x) = x$. V tom případě totiž máme

$$\mu_F\big((a, b]\big) = F(b) - F(a) = b - a = \lambda\big((a, b]\big).$$

(D)

Uvažujme posloupnost nezáporných čísel $a_1, a_2, \dots \geq 0$, jejíž řada $\sum_{n=1}^{\infty} a_n$ konverguje. Nechť $x_1, x_2, \dots$ je lib. posloupnost reálných čísel. Položme $F(x) = \sum\limits_{x_n \leq x} a_n$. Potom $F(x)$ je neklesající a zprava spojitá.

(S)

Nechť $f : \mathbb{R} \to \mathbb{R}$ je (lebesgueovsky) integrovatelná nezáporná funkce. Potom funkce $F(x) = \int\limits_{-\infty}^{x} f(u)\,\mathrm{d}u$ je neklesající a zprava spojitá.

Podmínku neklesající a zprava spojité funkce splňuje mj. každá distribuční funkce F. Náhodným veličinám odpovídají určité distribuční funkce a tedy i určité Lebesgueovy-Stieltjesovy míry μ_F.

Tvrzení

Nechť $F : \mathbb{R} \to \mathbb{R}$ je neklesající a zprava spojitá funkce a nechť $\lim\limits_{x\to-\infty} F(x) = 0$, $\lim\limits_{x\to\infty} F(x) = 1$. Potom existuje náhodná veličina X taková, že F je její distribuční funkcí, tzn. $F_X = F$.

Definice

Nechť $F : \mathbb{R} \to \mathbb{R}$ je neklesající a zprava spojitá funkce, μ_F je touto funkcí indukovaná Lebesgueova-Stieltjesova míra a $h : \mathbb{R} \to \mathbb{R}$ je borelovsky měřitelná funkce. Integrál $\int h \, \mathrm{d}\mu_F$ se nazývá **Lebesgueův-Stieltjesův integrál** funkce h vzhledem k funkci F a označuje se také $\int h \, \mathrm{d}\mu_F = \int h \, \mathrm{d}F$. Pro $B \in \mathcal{B}$ definujeme

$$\int_B h(x) \, \mathrm{d}F(x) = \int_B h \, \mathrm{d}\mu_F = \int_{\mathbb{R}} h(x) \, \mathbf{1}_B \, \mathrm{d}\mu_F = \int_{-\infty}^{\infty} h(x) \, \mathbf{1}_B(x) \, \mathrm{d}\mu_F(x).$$

Lebesgueův-Stieltjesův integrál je speciálním případem integrálu podle míry, $\int h \, \mathrm{d}\mu$, pro $\mu = \mu_F$.

Lebesgueův integrál je speciálním případem Lebesgueova-Stieltjesova integrálu pro $\mu_F = \lambda$, $F(x) = x$.

Za podmínek **(D)** je borelovsky měřitelná funkce h integrovatelná vzhledem k funkci F, právě když řada $\sum_{n=1}^{\infty} h(x_n)\, p_n$ absolutně konverguje, tj. když $\sum_{n=1}^{\infty} |h(x_n)| p_n < \infty$. V tom případě máme

$$\int_{-\infty}^{\infty} h(x)\, \mathrm{d}F(x) = \sum_{n=1}^{\infty} h(x_n)\, p_n = \sum_{n=1}^{\infty} h(x_n)\, \Delta F(x_n).$$

Za podmínek **(S)** je borelovsky měřitelná funkce h integrovatelná vzhledem k funkci F, právě když existuje integrál $\int_{\infty}^{\infty} h(x) f(x)\, \mathrm{d}x$. V tom případě máme

$$\int_{-\infty}^{\infty} h(x)\, \mathrm{d}F(x) = \int_{\infty}^{\infty} h(x) f(x)\, \mathrm{d}x.$$

a pro $B \in \mathcal{B}$ dále $$\int_{B} h(x)\, \mathrm{d}F(x) = \int_{\infty}^{\infty} h(x) f(x)\, \mathbf{1}_B(x)\, \mathrm{d}x.$$

Je-li F diferencovatelná, máme $\mathrm{d}F(x) = F'(x)\, \mathrm{d}x = f(x)\, \mathrm{d}x$.

Záměna integrálu a derivace, momentová vytvořující funkce

Uvažujme posloupnost náhodných veličin $X_1, X_2, \ldots$ konvergujících skoro jistě, $X_n \xrightarrow{\text{s.j.}} X$. Znamená to také, že $\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X)$? Víme, že obecně to neplatí.

Věta o monotonní konvergenci říká, že za podmínek $X_n \nearrow X$ a $\mathrm{E}(X_1) > -\infty$ opravdu platí $\lim_{n\to\infty} \mathrm{E}(X_n) = \mathrm{E}(X)$.

Tvrzení (O dominované konvergenci)

Nechť $X, X_1, X_2, \ldots$ jsou náhodné veličiny na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$, nechť $X_n \xrightarrow{\text{s.j.}} X$ a existuje náhodná veličina Y, $\mathrm{E}(Y) < \infty$, $|X_n| \leq Y$ pro $n \in \mathbb{N}$. Potom platí $\lim_{n\to\infty} \mathrm{E}(X_n) = X$.

Tvrzení platí speciálně pro konstantní náhodnou veličinu $Y = c$, tzn. pro omezené náhodné veličiny X_n.

Tvrzení (záměna pořadí derivace a integrálu)

Nechť $\left\{X_t(\omega);\ t \in (a, b)\right\}$ *je množina náhodných veličin na pravděpodobnostním prostoru* $(\Omega, \mathcal{A}, P)$ *závislých na parametru* t*, s konečnými středními hodnotami. Nechť pro všechna* $\omega \in \Omega$ *a* $t \in (a, b)$ *existuje derivace* $\dfrac{\partial}{\partial t} X_t(\omega) = X_t'(\omega)$*. Potom* X_t' *jsou opět náhodné veličiny.*
Dále nechť existuje náhodná veličina Y*,* $\mathrm{E}(Y) < \infty$*,* $|X_t'(\omega)| \le Y(\omega)$ *pro všechna* $\omega \in \Omega$ *a* $t \in (a, b)$*. Potom funkce* $\phi(t) = \mathrm{E}(X_t)$ *je diferencovatelná, má konečné derivace* $\phi'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\phi(t)$ *a platí*

$$\phi'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{E}(X_t) = \mathrm{E}(X_t') \qquad \text{pro } t \in (a, b).$$

Definice (Momentová vytvořující funkce)

Momentová vytvořující funkce *(mgf, moment generating function)* náhodné veličiny X na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je funkce

$$M_X(s) = \mathrm{E}_P\left(\mathrm{e}^{sX}\right) = \mathrm{E}_{P_X}\left(\mathrm{e}^{sx}\right), \qquad s \in \mathbb{R}.$$

Vlastnosti momentové vytvořující funkce:

- $M_X(0) = 1$,
- pro $s \neq 0$ může být $M_X(s) = \infty$,
- pro nezávislé náhodné veličiny X, Y je $M_{X+Y}(s) = M_X(s)\, M_Y(s)$,
- pro $Y \sim N(0;1)$ je $M_Y(s) = \exp\left[-\frac{1}{2}s^2\right]$.

Tvrzení (výpočet momentů)

Nechť X *je náhodná veličina s momentovou vytvořující funkcí* $M_X(s)$*,* $M_X(s) < \infty$ *pro* $|s| < \delta$ *pro nějaké* $\delta > 0$*. Potom* $\mathrm{E}(|X^n|) < \infty$ *pro všechna* $n \in \mathbb{N}_0$ *a platí*

$$M_X(s) = \sum_{n=0}^{\infty} \mathrm{E}(X^n) \frac{s^n}{n!},$$

$$M_X^{(k)}(0) = \frac{\partial^k}{\partial s^k} M_X(s) \bigg|_{s=0} = \mathrm{E}(X^k), \qquad k \in \mathbb{N}.$$

Tvrzení (Fubiniova věta)

Nechť μ je σ-konečná míra na $(\Omega_1, \mathcal{A}_1)$, ν je σ-konečná míra na $(\Omega_2, \mathcal{A}_2)$ a nechť $\mu \times \nu$ je součinová míra na $\mathcal{A}_1 \times \mathcal{A}_1$. Pokud $h(x, y) : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}$ je funkce měřitelná vzhledem k $\mu \times \nu$, potom platí

$$\int_{\Omega_1 \times \Omega_2} h \, \mathrm{d}(\mu \times \nu) = \int_{\Omega_1} \left(\int_{\Omega_2} h(x, y) \, \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) = \int_{\Omega_2} \left(\int_{\Omega_1} h(x, y) \, \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y),$$

pokud $\int_{\Omega_1 \times \Omega_2} h^+ \, \mathrm{d}(\mu \times \nu) < \infty$ nebo $\int_{\Omega_1 \times \Omega_2} h^- \, \mathrm{d}(\mu \times \nu) < \infty$.
Vnitřní integrály přitom mohou být nekonečné nebo nedefinované na množině míry nula.

***σ*-linearita střední hodnoty**

Nechť $\mu = P$ je pravděpodobnost, ν je čítací míra na $\mathbb{N}$ a $X_1, X_2, \ldots$ jsou náhodné veličiny s $\sum_{n=1}^{\infty} \mathrm{E}\left(|X_n|\right) < \infty$. Potom platí $\mathrm{E}\left(\sum_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} \mathrm{E}\left(X_n\right)$.

Konvoluce

Nechť X, Y jsou nezávislé náhodné veličiny s rozděleními pravděpodobnosti $\mu = P_X, \nu = P_Y$.

Potom $X + Y$ má rozdělení pravděpodobnosti $P_X * P_Y$, kde

$$(P_X * P_Y)(A) = \int_{\mathbb{R}} P_X(A - y)\,\mathrm{d}P_Y(y) = \int_{\mathbb{R}} P_Y(A - x)\,\mathrm{d}P_X(x), \qquad A \subseteq \mathbb{R},$$

kde $A - y = \{a - y;\ a \in A\}$.

Navíc, pokud P_X má hustotu f a P_Y má hustotu g vzhledem k Lebesgueově míře λ, potom $P_X * P_Y$ má vzhledem k λ hustotu $f * g$,

$$(f * g)(z) = \int_{\mathbb{R}} f(z - y)\,g(y)\,\mathrm{d}\lambda(y) = \int_{\mathbb{R}} f(x)\,g(z - x)\,\mathrm{d}\lambda(x), \qquad x \in \mathbb{R}.$$

Slabá konvergence, konvergence v distribuci

Definice

Posloupnost rozdělení pravděpodobnosti $P_1, P_2, \ldots$ náhodných veličin $X_1, X_2, \ldots$ **konverguje slabě** *(weak convergence)* k rozdělení pravděpodobnosti P_X náhodné veličiny X, pokud pro všechny omezené spojité funkce $h : \mathbb{R} \to \mathbb{R}$ platí

$$\lim_{n\to\infty} \int_{\mathbb{R}} h \, \mathrm{d}P_n = \int_{\mathbb{R}} h \, \mathrm{d}P_X.$$

Slabá konvergence rozdělení pravděpodobnosti je ekvivalentní konvergenci náhodných veličin **v distribuci** *(in distribution)*, $X_n \xrightarrow{\mathrm{d}} X$, definované jako bodová konvergence distribučních funkcí $\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P(X_n \leq x) = F_X(x)$ pro všechna $x \in \mathbb{R}$, kde je $F(x)$ spojitá.

Tvrzení

Následující vyjádření jsou ekvivalentní:

1. *posloupnost rozdělení* $P_1, P_2, \ldots$ *slabě konverguje k* P_X,
2. $\lim\limits_{n\to\infty} F_n(x) = F_X(x)$ *pro všechna* $x \in \mathbb{R}$, $P_X(\{x\}) = 0$,
3. $\lim\limits_{n\to\infty} P_N(B) = P_X(B)$ *pro všechny* $B \in \mathcal{B}$, *s nulovou mírou hranice,* $P_X(\partial B) = 0$.
4. $\lim\limits_{n\to\infty} \int\limits_{\mathbb{R}} h \, \mathrm{d}P_n = \int\limits_{\mathbb{R}} h \, \mathrm{d}P_X$ *pro všechny omezené borelovské funkce* h *s nulovou mírou bodů nespojitosti,*
5. ***Skorochodova věta (Skorohod's theorem):***
   na stejném pravděpodobnostním prostoru existují náhodné veličiny $Y, Y_1, Y_2, \ldots$, *takové, že* $P_Y = P_X$, $P_{Y_n} = P_n$, $n = 1, 2, \ldots$,
   a $Y_n \xrightarrow{\text{s.j.}} Y$.

Tvrzení (postačující podmínka konvergence v distribuci)

Pokud $X_n \xrightarrow{P} X$, *potom* $X_n \xrightarrow{\mathrm{d}} X$.

Shrňme si poznatky o konvergencích:

$$X_n \longrightarrow X \;\Rightarrow\; X_n \xrightarrow{\text{s.j.}} X \;\Rightarrow\; X_n \xrightarrow{P} X \;\Rightarrow\; X_n \xrightarrow{\text{d}} X.$$

Opačná poslední implikace neplatí obecně, pouze ve tvaru Skorochodovy věty: $X_n \xrightarrow{\text{d}} X \;\Rightarrow\; Y_n \xrightarrow{\text{s.j.}} Y$. Konvergence v distribuci totiž neříká nic o vztahu náhodných veličin $X, X_1, X_2, \ldots$, hovoří jen o jejich rozdělení pravděpodobnosti.

Speciálně v případě konstantní náhodné veličiny $X = c$ však ekvivalence platí v celém řetězci konvergencí.

Příklad

Uvažujme i.i.d. náhodné veličiny $X, X_1, X_2, \ldots$, $P(X = \pm 1) = \frac{1}{2}$, $P(X_n = \pm 1) = \frac{1}{2}$, $n = 1, 2, \ldots$

Zřejmě $X_n \xrightarrow{\text{d}} X$. Ale $P(|X_n - X| \geq 2) = \frac{1}{2}$, $\lim\limits_{n\to\infty} P(|X_n - X| \geq 2) = \frac{1}{2} \neq 0$, takže $X_n \overset{P}{\not\to} X$ a tedy $X_n \overset{\text{s.j.}}{\not\to} X$.

Následující tvrzení je velmi užitečným nástrojem ve statistice.

Tvrzení (Slutského věta, Slutsky theorem)

Nechť $X, X_1, X_2, \ldots, Y_1, Y_2, \ldots$ *jsou náhodné veličiny, takové, že* $X_n \xrightarrow{\mathrm{d}} X$, $Y_n \xrightarrow{P} c$, $c \in \mathbb{R}$.
Potom platí:

- $X_n + Y_n \xrightarrow{\mathrm{d}} X + c, \quad X_n - Y_n \xrightarrow{\mathrm{d}} X - c,$
- $X_n Y_n \xrightarrow{\mathrm{d}} c X,$
- $\dfrac{X_n}{Y_n} \xrightarrow{\mathrm{d}} \dfrac{X}{c}$ *pro* $c \neq 0$.

Charakteristická funkce, centrální limitní věta

Definice (Charakteristická funkce)

Charakteristická funkce *(characteristic function)* náhodné veličiny X na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je funkce $\psi_X : \mathbb{R} \to \mathbb{C}$,

$$\psi_X(t) = \mathrm{E}_P\left(\mathrm{e}^{i\,t\,X}\right) = \mathrm{E}_{P_X}\left(\mathrm{e}^{i\,t\,x}\right), \qquad t \in \mathbb{R}.$$

Vlastnosti charakteristické funkce:

- $\psi_X(0) = 1$,
- pro nezávislé náhodné veličiny X, Y je $\psi_{X+Y}(t) = \psi_X(t)\,\psi_Y(t)$,
- protože $|\mathrm{e}^{itx}| = 1$, vždy platí $|\psi_X(t)| \leq 1$,
- ψ_X je vždy stejnoměrně spojitá funkce,
- pro $Y \sim N(0;1)$ je $\psi_Y(t) = \exp\left[-\frac{1}{2}t^2\right]$.

Tvrzení (výpočet momentů)

Nechť X je náhodná veličina s konečným n-tým momentem, tj. $\mathrm{E}(|X^n|) < \infty$, a s charakteristickou funkcí $\psi_X(t)$. Potom pro $k = 0, 1, \ldots, n$ platí

$$\psi_X^{(k)}(t) = \mathrm{E}\left[(iX)^k \, \mathrm{e}^{itX}\right], \qquad \psi_X^{(k)}(0) = \frac{\partial^k}{\partial t^k}\psi_X(t)\bigg|_{t=0} = i^k \, \mathrm{E}(X^k).$$

Tvrzení (Fourierova věta o jednoznačnosti)

Nechť X, Y jsou náhodné veličiny. Potom $P_X = P_Y$, právě když $\psi_X = \psi_Y$.

Tvrzení (Fourierova věta o inverzi)

Nechť ψ_X je charakteristická funkce náhodné veličiny X s rozdělením pravděpodobnosti P_X. Pro $a, b, \in \mathbb{R}$, $a < b$, $P_X(\{a\}) = P_X(\{b\}) = 0$, potom platí

$$P_X([a, b]) = \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-i t a} - e^{-i t b}}{i t} \psi_X(t)\,\mathrm{d}t.$$

Tvrzení (O spojitosti)

Nechť $P, P_1, P_2, \ldots$ jsou rozdělení náhodných veličin $X, X_1, X_2, \ldots$ s odpovídajícími charakteristickými funkcemi $\psi, \psi_1, \psi_2, \ldots$. Potom $\{P_n\}_{n=1}^{n}$ konverguje slabě k P, tj. $X_n \xrightarrow{d} X$, právě když $\lim_{n\to\infty} \psi_n(t) = \psi(t)$ pro všechna $t \in \mathbb{R}$, tj. právě když posloupnost charakteristických funkcí konverguje bodově.

Tvrzení (Centrální limitní věta, Central Limit Theorem)

Nechť $X_1, X_2, \ldots$ je posloupnost náhodných veličin s konečnou střední hodnotou m a konečným rozptylem s^2. Označme částečné součty $S_n = \sum_{k=1}^{n} X_k$. Potom pro $n \to \infty$ rozdělení pravděpodobnosti náhodných veličin $\frac{S_n - m\,n}{s\sqrt{n}}$ slabě konverguje k $\mathrm{N}(0;1)$ rozdělení pravděpodobnosti,

$$\frac{S_n - m\,n}{s\sqrt{n}} \xrightarrow{\mathrm{d}} Y, \qquad Y \sim \mathrm{N}(0;1).$$

- Pro každé $x \in \mathbb{R}$: $\displaystyle \lim_{n\to\infty} P\left(\frac{S_n - m\,n}{s\sqrt{n}} \le x\right) = \Phi(x).$
- $\displaystyle \frac{S_n - m\,n}{\sqrt{n}} \xrightarrow{\mathrm{d}} Y, \qquad Y \sim \mathrm{N}(0;s^2).$

Slabou konvergenci rozdělení pravděpodobnosti, resp. konvergenci náhodných veličin v distribuci, lze dokázat i bez explicitního využití charakteristické funkce nebo Věty o spojitosti, pomocí tzv. metody momentů.

Definice

Označme **momenty rozdělení pravděpodobnosti** P_X:

$$\alpha_k = \int_{\mathbb{R}} x^k \,\mathrm{d}P_X(x), \qquad k = 1, 2, \ldots$$

Předpokládejme, že všechny tyto momenty α_k existují a jsou konečné. Pokud P_X je **jediné** rozdělení pravděpodobnosti, které má tyto momenty, říkáme, že P_X je **určené svými momenty**.

Která rozdělení pravděpodobnosti jsou určena svými momenty?

Tvrzení

Nechť X je náhodná veličina s momentovou vytvořující funkcí $M_X(s)$, která je omeze na na nějakém okolí počátku, tj. $M_X(s) < \infty$ pro $|s| < \delta$ pro nějaké $\delta > 0$. Potom rozdělení pravděpodobnosti P_X je určené svými momenty, tedy i určené svojí momentovou vytvořující funkcí.

Tvrzení

Nechť rozdělení pravděpodobnosti P_X je určené svými momenty. Dále nechť $P_1, P_2, \ldots$ jsou rozdělení pravděpodobnosti posloupnosti náhodných veličin $X_1, X_2, \ldots$, taková, že $\int_{\mathbb{R}} x^k \,\mathrm{d}P_n(x) < \infty$ pro každé $k, n \in \mathbb{N}$ a $\lim_{n\to\infty} \int_{\mathbb{R}} x^k \,\mathrm{d}P_n(x) = \int_{\mathbb{R}} x^k \,\mathrm{d}P_X(x)$ pro každé $k \in \mathbb{N}$. Potom $X_n \xrightarrow{\mathrm{d}} X$, tzn. posloupnost rozdělení pravděpodobnosti slabě konverguje k P_X.

Tvrzení vlastně říká, že z konvergence momentů plyne slabá konvergence, resp. konvergence v distribuci.

Radonova-Nikodymova derivace, dekompozice rozdělení pravděpodobnosti

Definice (Absolutní spojitost měr)

Nechť μ, ν jsou σ-konečné míry na měřitelném prostoru $(\Omega, \mathcal{A})$. Míra μ je **absolutně spojitá** *(absolutely continuous)* vzhledem k míře ν, $\mu \ll \nu$, pokud pro všechny $A \in \mathcal{A}$ platí: $\nu(A) = 0 \Rightarrow \mu(A) = 0$.

Tvrzení (Radonova-Nikodymova věta)

Nechť μ, ν jsou σ-konečné míry na měřitelném prostoru $(\Omega, \mathcal{A})$. Potom $\mu \ll \nu$, právě když existuje nezáporná měřitelná funkce $f : \Omega \to \mathbb{R}$, taková, že

$$\mu(A) = \int_A f(\omega)\,\mathrm{d}\nu(\omega) = \int_\Omega f(\omega)\,\mathbf{1}_A(\omega)\,\mathrm{d}\nu(\omega) \quad \text{pro } A \in \mathcal{A}.$$

Tato funkce f je přitom určena jednoznačně až na množinu míry nula.

Definice (Radonova-Nikodymova derivace)

Funkce f z Radonovy-Nikodymovy věty se nazývá **Radonova-Nikodymova derivace** a často se značí $f(\omega) = \frac{\mathrm{d}\mu}{\mathrm{d}\nu}(\omega)$.

Radonova-Nikodymova derivace je vlastně náhodnou veličinou.

Pro každou náhodnou veličinu (měřitelnou funkci) $X : \Omega \to \mathbb{R}$ a každou $A \in \mathcal{A}$ platí

$$\int_A X(\omega)\,\mathrm{d}\mu(\omega) = \int_A X(\omega)\,\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(\omega)\,\mathrm{d}\nu(\omega).$$

Pokud obě uvažované míry jsou pravděpodobnostní, $P = \mu$, $Q = \nu$, $P \ll Q$, dostáváme

$$\int_A X(\omega)\,\mathrm{d}P(\omega) = \int_A X(\omega)\,\frac{\mathrm{d}P}{\mathrm{d}Q}(\omega)\,\mathrm{d}Q(\omega).$$

Speciálně pro $A = \Omega$ pak obdržíme vzorec pro výpočet střední hodnoty náhodné veličiny vzhledem ke dvěma pravděpodobnostním mírám $P \ll Q$,

$$\mathrm{E}_P\,(X) = \int_\Omega X(\omega)\,\mathrm{d}P(\omega) = \int_\Omega X(\omega)\,\frac{\mathrm{d}P}{\mathrm{d}Q}(\omega)\,\mathrm{d}Q(\omega) = \mathrm{E}_Q\left(X\,\frac{\mathrm{d}P}{\mathrm{d}Q}\right).$$

Definice

Rozdělení pravděpodobnosti P_X je:

- **diskrétní**,
  pokud $P_X(\mathbb{R}) = \sum_{x \in M} P_X(\{x\})$ pro nejvýše spočetnou množinu $M \subset \mathbb{R}$.
- **absolutně spojité (vzhledem k Lebesgueově míře λ)**,
  pokud existuje nezáporná borelovsky měřitelná funkce $f : \mathbb{R} \to \mathbb{R}$, taková, že pro každou $B \in \mathcal{B}$ platí
  $$P_X(B) = \int_B f(x)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}} f(x)\,\mathbf{1}_B(x)\,\mathrm{d}\lambda(x).$$
- **singulární spojité (vzhledem k Lebesgueově míře λ)**,
  pokud $P_X(\{x\}) = 0$ pro všechna $x \in \mathbb{R}$, ale přitom existuje množina $M \subseteq \mathbb{R}$: $\lambda(M) = 0$, $P_X(\overline{M}) = 0$.

Tvrzení (Lebesgueova dekompozice rozdělení pravděpodobnosti)

Každé rozdělení pravděpodobnosti P_X lze jednoznačně rozložit do tvaru

$$P_X = P_D + P_{AS} + P_S, \qquad \text{kde}$$

- P_D *je diskrétní rozdělení pravděpodobnosti,*
- P_{AS} *je rozdělení pravděpodobnosti absolutně spojité vzhledem k* λ*,*
- P_S *je rozdělení pravděpodobnosti singulární spojité vzhledem k* λ*.*

Tvrzení (Lebesgueova dekompozice měr)

Pokud μ, ν *jsou dvě* σ*-konečné míry na měřitelném prostoru* $(\Omega, \mathcal{A})$*, potom* μ *lze jednoznačně rozložit do tvaru*

$$\mu = \mu_{AS} + \mu_S, \qquad \text{kde}$$

- μ_{AS} *je míra absolutně spojitá vzhledem k* ν*,* $\mu_{AS} \ll \nu$*,*
- μ_S *je singulární míra vzhledem k* ν*,*
  tj. existuje množina $M \subseteq \mathbb{R}$*:* $\nu(M) = 0$*,* $\mu_S(\overline{M}) = 0$*.*

Podmíněná pravděpodobnost, podmíněná střední hodnota

Víme, že pro podmíněnou pravděpodobnost platí

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)}, \qquad \text{pokud } P(C) > 0.$$

Obecněji, pokud Y je náhodná veličina a $P(C) > 0$, lze definovat $Y|C$, rozdělení pravděpodobnosti náhodné veličiny Y podmíněné jevem C, předpisem

$$P_{Y \mid C}(B) = P(Y \in B \mid C) = \frac{P\left((Y \in B) \cap C\right)}{P(C)}, \qquad B \in \mathcal{B}.$$

Podobně lze definovat i podmíněnou střední hodnotu,

$$\mathrm{E}(Y \mid C) = \int_{\Omega} y \, \mathrm{d}P_{Y|C} = \frac{\mathrm{E}(Y \boldsymbol{1}_C)}{P(C)}.$$

Uvedený postup však zcela selhává v případě $P(C) = 0$. Podmíněné pravděpodobnosti a podmíněné střední hodnoty i za podmínky s nulovou pravděpodobností však v praxi běžně potřebujeme, např. při výpočtech marginálních charakteristik ve vícerozměrných rozděleních. Pro korektní definici tak musíme zvolit zcela jiný, neintuitivní, přístup.

Definice

Nechť $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor. Množina $\mathcal{H}$ je **sub-σ-algebra**, pokud $\mathcal{H} \subseteq \mathcal{A}$ a je σ-algebrou na Ω.
Náhodná veličina X na $(\Omega, \mathcal{A}, P)$ je **$\mathcal{H}$-měřitelná**, pokud

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{H} \qquad \text{pro všechna } x \in \mathbb{R}.$$

Dále definujeme

$$\sigma(X) = \big\{\{X \in B\} : B \in \mathcal{B}\big\} = \big\{\{\omega \in \Omega : X(\omega) \in B\} : B \in \mathcal{B}\big\}.$$

Uvědomte si, že $\sigma(X)$ je vždy sub-σ-algebrou.

Definice (podmiňování náhodnou veličinou)

Nechť X, Y jsou náhodné veličiny na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$, $\mathrm{E}(|Y|) < \infty$, a nechť $A \in \mathcal{A}$ je jev.

- **Podmíněná střední hodnota** $\mathrm{E}(Y \mid X)$ náhodné veličiny Y při dané X je náhodná veličina $\mathrm{E}(Y \mid X)(\omega)$, pokud je $\sigma(X)$-měřitelná a pro všechny $B \in \mathcal{B}$ splňuje
$$\mathrm{E}\big[\mathrm{E}(Y \mid X)\, \mathbf{1}_{\{X \in B\}}\big] = \mathrm{E}\big[Y\, \mathbf{1}_{\{X \in B\}}\big].$$
- **Podmíněná pravděpodobnost** $P(A \mid X)$ jevu A při dané X je náhodná veličina $P(A \mid X)(\omega) = \mathrm{E}(\mathbf{1}_A \mid X)(\omega)$, pokud je $\sigma(X)$-měřitelná a pro všechny $B \in \mathcal{B}$ splňuje
$$\mathrm{E}\big[P(A \mid X)\, \mathbf{1}_{\{X \in B\}}\big] = P\big[A \cap \{X \in B\}\big].$$

Podmíněnou pravděpodobnost i podmíněnou střední hodnotu zavádíme jako náhodné veličiny na $\sigma(X)$, což je zcela neintuitivní. Jsme totiž zvyklí s pravděpodobností i se střední hodnotou pracovat jako s čísly. Tento nový přístup však umožňuje překonat problémy s podmiňováním jevy s nulovou pravděpodobností.

Při volbě $B = \mathbb{R}$ dostáváme

$$\mathrm{E}\big[P(A\,|\,X)\,\mathbf{1}_{\{X\in\mathbb{R}\}}\big] = P\big[A \cap \{X \in \mathbb{R}\}\big] = P(A \cap \Omega) = P(A),$$
$$\mathrm{E}\big[\mathrm{E}(Y\,|\,X)\,\mathbf{1}_{\{X\in\mathbb{R}\}}\big] = \mathrm{E}\big[Y\,\mathbf{1}_{\{X\in\mathbb{R}\}}\big] = \mathrm{E}(Y\cdot 1) = \mathrm{E}(Y).$$

To znamená, že v průměru přes všechny možné hodnoty náhodné veličiny X se $P(A\,|\,X)$ chová jako $P(A)$ a $\mathrm{E}(Y\,|\,X)$ jako $\mathrm{E}(Y)$.

Již víme, že střední hodnoty nejsou ovlivněny změnami na množině míry nula. Podmíněné pravděpodobnosti a podmíněné střední hodnoty jsou proto určeny jednoznačně až na množinu míry nula.

Tvrzení

Nechť X, Y jsou náhodné veličiny na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ a $A \in \mathcal{A}$ je jev. Potom podmíněná pravděpodobnost $P(A\,|\,X)$ a podmíněná střední hodnota $\mathrm{E}(Y\,|\,X)$ vždy existují a jsou určeny jednoznačně až na množinu míry nula.

Konstrukce

Pro $H \in \sigma(X)$ zavedeme následující míry:

- P_0 jako restrikci P na $\sigma(X)$, tj. $P_0(H) = P(H)$,
- ν předpisem $\nu(H) = P(A \cap H)$,
- μ^+, μ^- předpisy $\mu^+(H) = \mathrm{E}(Y^+ \boldsymbol{1}_E)$ a $\mu^-(H) = \mathrm{E}(Y^- \boldsymbol{1}_E)$.

Všechny zavedené míry jsou přitom absolutně spojité vzhledem k P_0, $\nu \ll P_0$, $\mu^+ \ll P_0$ a $\mu^- \ll P_0$. Podle Radonovy-Nikodymovy tedy věty existují Radonovy-Nikodymovy derivace $\frac{\mathrm{d}\nu}{\mathrm{d}P_0}(\omega)$, $\frac{\mathrm{d}\mu^+}{\mathrm{d}P_0}(\omega)$ a $\frac{\mathrm{d}\mu^-}{\mathrm{d}P_0}(\omega)$ a jsou určeny jednoznačně až na množinu míry (P_0) nula.

Nyní definujeme

- $\mathrm{E}(Y \mid X)(\omega) = \frac{\mathrm{d}\mu^+}{\mathrm{d}P_0}(\omega) - \frac{\mathrm{d}\mu^-}{\mathrm{d}P_0}(\omega)$,
- $P(A \mid X)(\omega) = \frac{\mathrm{d}\nu}{\mathrm{d}P_0}(\omega)$.

Definice (podmiňování sub-σ-algebrou)

Obecně definujeme podmíněnou střední hodnotu $\mathrm{E}(Y \,|\, \mathcal{H})(\omega)$ a podmíněnou pravděpodobnost $P(A \,|\, \mathcal{H})(\omega)$ za podmínky dané sub-σ-algebrou $\mathcal{H}$ jako $\mathcal{H}$-měřitelné náhodné veličiny splňující pro všechny $H \in \mathcal{H}$ podmínky

$$\mathrm{E}\big[\mathrm{E}(Y \,|\, \mathcal{H})\,\boldsymbol{1}_H\big] = \mathrm{E}\big[Y\,\boldsymbol{1}_H\big],$$

$$\mathrm{E}\big[P(A \,|\, \mathcal{H})\,\boldsymbol{1}_H\big] = \mathrm{E}\big[\mathrm{E}(\boldsymbol{1}_A \,|\, \mathcal{H})\,\boldsymbol{1}_H\big] = P\big[A \cap H\big].$$

Konstrukce

Pro $H \in \mathcal{H}$ zavedeme následující míry:

- P_0 jako restrikci P na $\mathcal{H}$, tj. $P_0(H) = P(H)$,
- ν předpisem $\nu(H) = P(A \cap H)$,
- μ^+, μ^- předpisy $\mu^+(H) = \mathrm{E}(Y^+\,\boldsymbol{1}_H)$ a $\mu^-(H) = \mathrm{E}(Y^-\,\boldsymbol{1}_H)$.

Nyní definujeme

- $\mathrm{E}(Y \,|\, \mathcal{H})(\omega) = \dfrac{\mathrm{d}\mu^+}{\mathrm{d}P_0}(\omega) - \dfrac{\mathrm{d}\mu^-}{\mathrm{d}P_0}(\omega)$,
- $P(A \,|\, \mathcal{H})(\omega) = \dfrac{\mathrm{d}\nu}{\mathrm{d}P_0}(\omega)$.

Při volbě $\mathcal{H} = \sigma(X)$ zjistíme, že vlastně platí

$$P(A \mid \sigma(X)) = P(A \mid X), \qquad \mathrm{E}(A \mid \sigma(X)) = \mathrm{E}(A \mid X).$$

Při volbě $\mathcal{H} = \{\varnothing, \Omega\}$ obdržíme konstantní náhodné veličiny

$$P(A \mid \{\varnothing, \Omega\})(\omega) = P(A), \qquad \mathrm{E}(A \mid \{\varnothing, \Omega\})(\omega) = \mathrm{E}(Y).$$

Při volbě $\mathcal{H} = \mathcal{A}$ obdržíme

$$P(A \mid \mathcal{A})(\omega) = \boldsymbol{1}_A(\omega), \qquad \mathrm{E}(A \mid \mathcal{A})(\omega) = Y(\omega).$$

Definice

Podmíněný rozptyl náhodné veličiny Y při dané náhodné veličině X, resp. při dané sub-σ-algebře $\mathcal{H}$, je

$$\mathrm{Var}(Y \mid X) = \mathrm{E}\Big[\big(Y - \mathrm{E}(Y \mid X)\big)^2 \,\Big|\, X\Big],$$
$$\mathrm{Var}(Y \mid \mathcal{H}) = \mathrm{E}\Big[\big(Y - \mathrm{E}(Y \mid \mathcal{H})\big)^2 \,\Big|\, \mathcal{H}\Big].$$

Podmíněný rozptyl kvantifikuje rozptyl Y při znalosti X.

Tvrzení

Pokud $\mathrm{Var}(Y) < \infty$, *potom platí*

$$\mathrm{Var}(Y) = \mathrm{E}\big[\mathrm{Var}(Y \mid X)\big] + \mathrm{Var}\big[\mathrm{E}(Y \mid X)\big],$$
$$\mathrm{Var}(Y) = \mathrm{E}\big[\mathrm{Var}(Y \mid \mathcal{H})\big] + \mathrm{Var}\big[\mathrm{E}(Y \mid \mathcal{H})\big].$$

Nepodmíněný rozptyl lze rozložit na střední hodnotu podmíněného rozptyl a rozptyl podmíněné střední hodnoty.

Tvrzení

Nechť X, Y *jsou náhodné veličiny na pravděpodobnostním prostoru* $(\Omega, \mathcal{A}, P)$*,* $\mathrm{E}(Y) < \infty$*,* $\mathrm{E}(X\,Y) < \infty$*,* $\mathcal{H}$ *je sub-*σ*-algebra a nechť* X *je* $\mathcal{H}$*-měřitelná.*
Potom s pravděpodobností 1 platí

$$\mathrm{E}(X\,Y \,|\, \mathcal{H}) = X\,\mathrm{E}(Y \,|\, \mathcal{H}).$$

Tvrzení

Nechť X *je náhodná veličina na pravděpodobnostním prostoru* $(\Omega, \mathcal{A}, P)$*,* $\mathrm{E}(X) < \infty$ *a nechť* $\mathcal{H}_1 \subseteq \mathcal{H}_2$ *jsou dvě sub-*σ*-algebry.*
Potom s pravděpodobností 1 platí

$$\mathrm{E}\Big[\mathrm{E}(X \,|\, \mathcal{H}_2) \,\Big|\, \mathcal{H}_1\Big] = \mathrm{E}(X \,|\, \mathcal{H}_1).$$

Martingal

Definice (Diskrétní martingal)

Náhodný proces s diskrétním časem $X_0(\omega), X_1(\omega), X_2(\omega), \ldots$ je **martingal** *(martingale)*, pokud pro všechna $n \in \mathbb{N}_0$ platí

- $\mathrm{E}\big(|X_n|\big) < \infty$,
- $\mathrm{E}\big(X_{n+1} \,|\, X_0, X_1, \ldots, X_n\big) = \mathrm{E}\big(X_{n+1} \,|\, \sigma(X_0, X_1, \ldots, X_n)\big) = X_n$.

Literatura

Použité zdroje a další studijní literatura:

1. Rosenthal J. S. (2006) *A First Look at Rigorous Probability Theory*
2. Roussas G. G. (2014) *An Introduction to Measure-Theoretic Probability*
3. Riečan B. (2015) *Miniteória pravdepodobnosti*
4. Riečan B. (1972) *O pravdepodobnosti a miere*