

# Matematika III – 8. týden

## Pravděpodobnost a náhodné veličiny

Jan Slovák

Masarykova univerzita  
Fakulta informatiky

9. 11. – 13. 11. 2014

# Obsah přednášky

- 1 Literatura
- 2 Pravděpodobnost
- 3 Náhodné veličiny
- 4 Pravděpodobnostní funkce a hustoty

# Plán přednášky

- 1 Literatura
- 2 Pravděpodobnost
- 3 Náhodné veličiny
- 4 Pravděpodobnostní funkce a hustoty

## Kde je dobré číst?

- Karel Zvára, Josef Štěpán, Pravděpodobnost a matematická pravděpodobnost statistika, Matfyzpress, 2006, 230pp.
- J. Slovák, M. Panák, M. Bulant, Matematika drsně a svižně, Muni Press, Brno 2013, v+773 s., elektronická edice [www.math.muni.cz/Matematika\\_drsne\\_svizne](http://www.math.muni.cz/Matematika_drsne_svizne)
- Marie Budíková, Štěpán Mikoláš, Pavel Osecký, Teorie pravděpodobnosti a matematická statistika (sbírka příkladů), Masarykova univerzita, 3. vydání, 2004, 117 stran, ISBN 80-210-3313-4.
- Marie Budíková, Tomáš Lerch, Štěpán Mikoláš, Základní statistické metody, Masarykova univerzita, 2005, 170 stran, ISBN 80-210-3886-1.
- Riley, K.F., Hobson, M.P., Bence, S.J. Mathematical Methods for Physics and Engineering, second edition, Cambridge University Press, Cambridge 2004, ISBN 0 521 89067 5, xxiii + 1232 pp.

# Plán přednášky

- 1 Literatura
- 2 **Pravděpodobnost**
- 3 Náhodné veličiny
- 4 Pravděpodobnostní funkce a hustoty

Připomeneme (a trochu zobecníme) pojmy a výsledky z druhé přednášky prvního semestru.

### Definition (Náhodné jevy)

Budeme pracovat s neprázdnou pevně zvolenou množinou  $\Omega$  všech možných výsledků, kterou nazýváme **základní prostor**.

Připomeneme (a trochu zobecníme) pojmy a výsledky z druhé přednášky prvního semestru.

### Definition (Náhodné jevy)

Budeme pracovat s neprázdnou pevně zvolenou množinou  $\Omega$  všech možných výsledků, kterou nazýváme **základní prostor**. Prvky  $\omega \in \Omega$  představují jednotlivé **možné výsledky**.

Připomeneme (a trochu zobecníme) pojmy a výsledky z druhé přednášky prvního semestru.

### Definition (Náhodné jevy)

Budeme pracovat s neprázdnou pevně zvolenou množinou  $\Omega$  všech možných výsledků, kterou nazýváme **základní prostor**.

Prvky  $\omega \in \Omega$  představují jednotlivé **možné výsledky**.

Systém podmnožin  $\mathcal{A}$  základního prostoru se nazývá **jevové pole** a jeho prvky se nazývají **jevy**, jestliže

- $\Omega \in \mathcal{A}$ , tj. základní prostor, je jevem,
- je-li  $A, B \in \mathcal{A}$ , pak  $A \setminus B \in \mathcal{A}$ , tj. pro každé dva jevy je jevem i jejich množinový rozdíl,
- je-li  $A_i \in \mathcal{A}$ ,  $i \in I$  nejvýše spočetný systém jevů, pak také jejich sjednocení je jevem, tj.  $\cup_{i \in I} A_i \in \mathcal{A}$ .



- Komplement  $A^c = \Omega \setminus A$  jevu  $A$  je jevem, který nazýváme *opačný jev* k jevu  $A$ .

- Komplement  $A^c = \Omega \setminus A$  jevu  $A$  je jevem, který nazýváme *opačný jev* k jevu  $A$ .
- Průnik dvou jevů opět jevem, protože pro každé dvě podmnožiny  $A, B \subset \Omega$  platí

$$A \setminus (\Omega \setminus B) = A \cap B.$$

- Komplement  $A^c = \Omega \setminus A$  jevu  $A$  je jevem, který nazýváme *opačný jev* k jevu  $A$ .
- Průnik dvou jevů opět jevem, protože pro každé dvě podmnožiny  $A, B \subset \Omega$  platí

$$A \setminus (\Omega \setminus B) = A \cap B.$$

Jevové pole je tedy systém podmnožin základního prostoru uzavřený na konečné průniky, spočetná sjednocení a množinové rozdíly. Jednotlivé množiny  $A \in \mathcal{A}$  nazýváme **náhodné jevy** (vzhledem k  $\mathcal{A}$ ).

Terminologie připomíná souvislosti s popisem skutečných jevů a jejich statistickým popisem:

- celý základní prostor  $\Omega$  se nazývá **jistý jev**, prázdná podmnožina  $\emptyset \in \mathcal{A}$  se nazývá **nemožný jev**,

Terminologie připomíná souvislosti s popisem skutečných jevů a jejich statistickým popisem:

- celý základní prostor  $\Omega$  se nazývá **jistý jev**, prázdná podmnožina  $\emptyset \in \mathcal{A}$  se nazývá **nemožný jev**,
- jednoprvkové podmnožiny  $\{\omega\} \in \Omega$  se nazývají **elementární jevy**,

Terminologie připomíná souvislosti s popisem skutečných jevů a jejich statistickým popisem:

- celý základní prostor  $\Omega$  se nazývá **jistý jev**, prázdná podmnožina  $\emptyset \in \mathcal{A}$  se nazývá **nemožný jev**,
- jednoprvkové podmnožiny  $\{\omega\} \in \Omega$  se nazývají **elementární jevy**,
- **společné nastoupení jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcap_{i \in I} A_i$ ,  
**nastoupení alespoň jednoho z jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcup_{i \in I} A_i$ ,

Terminologie připomíná souvislosti s popisem skutečných jevů a jejich statistickým popisem:

- celý základní prostor  $\Omega$  se nazývá **jistý jev**, prázdná podmnožina  $\emptyset \in \mathcal{A}$  se nazývá **nemožný jev**,
- jednoprvkové podmnožiny  $\{\omega\} \in \Omega$  se nazývají **elementární jevy**,
- **společné nastoupení jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcap_{i \in I} A_i$ ,  
**nastoupení alespoň jednoho z jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcup_{i \in I} A_i$ ,
- $A, B \in \mathcal{A}$  jsou **neslučitelné jevy**, je-li  $A \cap B = \emptyset$ ,

Terminologie připomíná souvislosti s popisem skutečných jevů a jejich statistickým popisem:

- celý základní prostor  $\Omega$  se nazývá **jistý jev**, prázdná podmnožina  $\emptyset \in \mathcal{A}$  se nazývá **nemožný jev**,
- jednoprvkové podmnožiny  $\{\omega\} \in \Omega$  se nazývají **elementární jevy**,
- **společné nastoupení jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcap_{i \in I} A_i$ ,  
**nastoupení alespoň jednoho z jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcup_{i \in I} A_i$ ,
- $A, B \in \mathcal{A}$  jsou **neslučitelné jevy**, je-li  $A \cap B = \emptyset$ ,
- jev  $A$  má za **důsledek** jev  $B$ , když  $A \subset B$ ,



Terminologie připomíná souvislosti s popisem skutečných jevů a jejich statistickým popisem:

- celý základní prostor  $\Omega$  se nazývá **jistý jev**, prázdná podmnožina  $\emptyset \in \mathcal{A}$  se nazývá **nemožný jev**,
- jednoprvkové podmnožiny  $\{\omega\} \in \Omega$  se nazývají **elementární jevy**,
- **společné nastoupení jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcap_{i \in I} A_i$ , **nastoupení alespoň jednoho z jevů**  $A_i, i \in I$ , odpovídá jevu  $\bigcup_{i \in I} A_i$ ,
- $A, B \in \mathcal{A}$  jsou **neslučitelné jevy**, je-li  $A \cap B = \emptyset$ ,
- jev  $A$  má za **důsledek** jev  $B$ , když  $A \subset B$ ,
- je-li  $A \in \mathcal{A}$ , pak se jev  $B = \Omega \setminus A$  nazývá **opačný jev k jevu**  $A$ , píšeme  $B = A^c$ .

## Definition (Pravděpodobnost)

**Pravděpodobnostní prostor** je jevové pole  $\mathcal{A}$  podmnožin (konečného) základního prostoru  $\Omega$ , na kterém je definována skalární funkce  $P : \mathcal{A} \rightarrow \mathbb{R}$  s následujícími vlastnosti:

- je nezáporná, tj.  $P(A) \geq 0$  pro všechny jevy  $A$ ,
- je aditivní, tj.  $P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$ , pro každý nejvýše spočetný systém po dvou disjunktních jevů,
- pravděpodobnost jistého jevu je 1.

Funkci  $P$  nazýváme **pravděpodobností** na jevovém poli  $(\Omega, \mathcal{A})$ .

## Definition (Pravděpodobnost)

**Pravděpodobnostní prostor** je jevové pole  $\mathcal{A}$  podmnožin (konečného) základního prostoru  $\Omega$ , na kterém je definována skalární funkce  $P : \mathcal{A} \rightarrow \mathbb{R}$  s následujícími vlastnosti:

- je nezáporná, tj.  $P(A) \geq 0$  pro všechny jevy  $A$ ,
- je aditivní, tj.  $P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$ , pro každý nejvýše spočetný systém po dvou disjunktních jevů,
- pravděpodobnost jistého jevu je 1.

Funkci  $P$  nazýváme **pravděpodobností** na jevovém poli  $(\Omega, \mathcal{A})$ .

## Důsledky

Pro všechny jevy platí  $P(A^c) = 1 - P(A)$ .

## Definition (Pravděpodobnost)

**Pravděpodobnostní prostor** je jevové pole  $\mathcal{A}$  podmnožin (konečného) základního prostoru  $\Omega$ , na kterém je definována skalární funkce  $P : \mathcal{A} \rightarrow \mathbb{R}$  s následujícími vlastnosti:

- je nezáporná, tj.  $P(A) \geq 0$  pro všechny jevy  $A$ ,
- je aditivní, tj.  $P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$ , pro každý nejvýše spočetný systém po dvou disjunktních jevů,
- pravděpodobnost jistého jevu je 1.

Funkci  $P$  nazýváme **pravděpodobností** na jevovém poli  $(\Omega, \mathcal{A})$ .

## Důsledky

Pro všechny jevy platí  $P(A^c) = 1 - P(A)$ .

Additivnost platí pro jakýkoliv spočetný počet neslučitelných jevů  $A_i \subset \Omega$ ,  $i \in I$ , tj.

$$P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i), \text{ kdykoliv je } A_i \cap A_j = \emptyset, i \neq j, i, j \in I.$$

Připomeňme si klasickou konečnou pravděpodobnost.

Připomeňme si klasickou konečnou pravděpodobnost.

### Definition

Nechť  $\Omega$  je konečný základní prostor a necht' jevové pole  $\mathcal{A}$  je právě systém všech podmnožin v  $\Omega$ . **Klasická pravděpodobnost** je pravděpodobnostní prostor  $(\Omega, \mathcal{A}, P)$  s pravděpodobnostní funkcí  $P : \mathcal{A} \rightarrow \mathbb{R}$ ,

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

Zjevně takto zadaná funkce skutečně definuje pravděpodobnost.

## Peterburgský paradox (Bernoulli, 1738)

Typický příklad klasické pravděpodobnosti jsou jevy související s házením mincí. Představme si následující pravidla kasina:

## Peterburgský paradox (Bernoulli, 1738)

Typický příklad klasické pravděpodobnosti jsou jevy související s házením mincí. Představme si následující pravidla kasina: Návštěvník zaplatí vklad  $C$  a poté hází mincí. Je-li  $T$  počet hodů potřebných k první hlavě, pak obdrží výhru  $2^T$ . Jaká je „fér hodnota“ pro vklad  $C$ ?



## Peterburgský paradox (Bernoulli, 1738)

Typický příklad klasické pravděpodobnosti jsou jevy související s házením mincí. Představme si následující pravidla kasina: Návštěvník zaplatí vklad  $C$  a poté hází mincí. Je-li  $T$  počet hodů potřebných k první hlavě, pak obdrží výhru  $2^T$ . Jaká je „férová hodnota“ pro vklad  $C$ ?

Pravděpodobnost, že padne hlava je u férové mince  $1/2$ , je proto  $P(T = k) = 2^{-k}$ . Pravděpodobnost, že po nějakém konečném počtu hodů hra skončí je dána součtem  $\sum_{k=1}^{\infty} 2^{-k} = 1$ . Proto je úpravděpodobnost jevu, že stále padá orel nulová.

Sečteme-li všechny pravděpodobnosti výsledků vynásobených výhrami  $2^k$ , dostaneme  $\sum_1^{\infty} 1 = \infty$ . Zdá se proto, že se vyplatí vložit i velký vklad...

# Peterburgský paradox (Bernoulli, 1738)

Typický příklad klasické pravděpodobnosti jsou jevy související s házením mincí. Představme si následující pravidla kasina: Návštěvník zaplatí vklad  $C$  a poté hází mincí. Je-li  $T$  počet hodů potřebných k první hlavě, pak obdrží výhru  $2^T$ . Jaká je „férová hodnota“ pro vklad  $C$ ?

Pravděpodobnost, že padne hlava je u férové mince  $1/2$ , je proto  $P(T = k) = 2^{-k}$ . Pravděpodobnost, že po nějakém konečném počtu hodů hra skončí je dána součtem  $\sum_{k=1}^{\infty} 2^{-k} = 1$ . Proto je úpravděpodobnost jevu, že stále padá orel nulová.

Sečteme-li všechny pravděpodobnosti výsledků vynásobených výhrami  $2^k$ , dostaneme  $\sum_1^{\infty} 1 = \infty$ . Zdá se proto, že se vyplatí vložit i velký vklad...

Ve skutečnosti simulací hry zjistíme, že nezávisle na počtu pokusů se prakticky všechny výhry budou pohybovat v rozmezí  $T$  do 6. Důvodem je, že vysoké výhry jsou velice nepravděpodobné a proto je při reálných úvahách nelze brát vážně.

# Podmíněná pravděpodobnost

Obvyklé je také klást dotazy s dodatečnou podmínkou. Např. „jaká je pravděpodobnost, že při hodu dvěma kostkami padly dvě pětky, je-li součet hodnot deset?“. Připomeneme, že formalizovat takové úvahy umíme následovně.

## Definition

Nechť  $H$  je jev s nenulovou pravděpodobností v jevovém poli  $\mathcal{A}$  v pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$ . **Podmíněná pravděpodobnost**  $P(A|H)$  jevu  $A \in \mathcal{A}$  vzhledem k hypotéze  $H$  je definována vztahem

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

Definice odpovídá požadavku, že jevy  $A$  a  $H$  nastanou zároveň, za předpokladu, že  $A$  nastal s pravděpodobností  $P(A \cap H)/P(A)$ .

# Podmíněná pravděpodobnost

Obvyklé je také klást dotazy s dodatečnou podmínkou. Např. „jaká je pravděpodobnost, že při hodu dvěma kostkami padly dvě pětky, je-li součet hodnot deset?“. Připomeneme, že formalizovat takové úvahy umíme následovně.

## Definition

Nechť  $H$  je jev s nenulovou pravděpodobností v jevovém poli  $\mathcal{A}$  v pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$ . **Podmíněná pravděpodobnost**  $P(A|H)$  jevu  $A \in \mathcal{A}$  vzhledem k hypotéze  $H$  je definována vztahem

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

Definice odpovídá požadavku, že jevy  $A$  a  $H$  nastanou zároveň, za předpokladu, že  $A$  nastal s pravděpodobností  $P(A \cap H)/P(A)$ . Je také vidět přímo z definice, hypotéza  $H$  a jev  $A$  jsou nezávislé tehdy a jen tehdy, je-li  $P(A) = P(A|H)$ .

# Bayesovy věty

Přepsáním formule pro podmíněnou pravděpodobnost dostáváme

$$P(A \cap B) = P(B \cap A) = P(A)P(B|A) = P(B)P(A|B).$$

## Theorem (Bayesovy věty)

*Pro pravděpodobnost jevů  $A$  a  $B$  platí*

- 1  $P(A|B) = \frac{P(A)P(B|A)}{P(B)}.$
- 2  $P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A')P(B|A')}.$

# Bayesovy věty

Přepsáním formule pro podmíněnou pravděpodobnost dostáváme

$$P(A \cap B) = P(B \cap A) = P(A)P(B|A) = P(B)P(A|B).$$

## Theorem (Bayesovy věty)

*Pro pravděpodobnost jevů  $A$  a  $B$  platí*

- 1  $P(A|B) = \frac{P(A)P(B|A)}{P(B)}.$
- 2  $P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A')P(B|A')}.$

## Důkaz.

První tvrzení je přepsáním předchozí formule, druhé z prvního plyne dosazením  $P(B) = P(A)P(B|A) + P(A')P(B|A')$ . □

## Příklad – testování

Předpokládejme, že předpokladem přijetí studentů na univerzitu jsou testy způsobilosti ke studiu. Inteligentní osoba v něm má 99% úspěšnost. Zároveň předpokládejme, že úspěšnost neinteligentních osob je 0.5%.

## Příklad – testování

Předpokládejme, že předpokladem přijetí studentů na univerzitu jsou testy způsobilosti ke studiu. Inteligentní osoba v něm má 99% úspěšnost. Zároveň předpokládejme, že úspěšnost neinteligentních osob je 0.5%.

S jakou pravděpodobností je náhodně vybraný student/ka univerzity inteligentní, jestliže je v populaci je  $p$  promile inteligentních osob (tj.  $p$  osob z tisíce považujeme za inteligentní).



## Příklad – testování

Předpokládejme, že předpokladem přijetí studentů na univerzitu jsou testy způsobilosti ke studiu. Inteligentní osoba v něm má 99% úspěšnost. Zároveň předpokládejme, že úspěšnost neinteligentních osob je 0.5%.

S jakou pravděpodobností je náhodně vybraný student/ka univerzity inteligentní, jestliže je v populaci je  $p$  promile inteligentních osob (tj.  $p$  osob z tisíce považujeme za inteligentní).

Označme  $A$  jev, že je daná osoba je inteligentní, a  $B$  jev, že prošla testem. Dle Bayesovy věty je hledaná pravděpodobnost

$$P(A|B) = \frac{p/1000 \cdot 99/100}{p/1000 \cdot 99/100 + (1000 - p)/1000 \cdot 5/1000}$$

Jestliže zvolíme za  $p$  nějaké konkrétní četnosti, dostaneme příslušné očekávatelné spolehlivosti testu. V následující tabulce je spočten výsledek pro několik  $p$ :

$p$	500	100	10	1	0.1
$P(A B)$	0.99	0,96	0.67	0.17	0.02

Pokud stejné číselné zadání použijeme pro screening některé nemoci, řekněme HIV pozitivitu, dostáváme hrozné výsledky!

Výsledek asi neodpovídá naší intuici a může se zdát šokující ve vztahu k použití takovýchto testů.

Výsledek asi neodpovídá naší intuici a může se zdát šokující ve vztahu k použití takovýchto testů.

Evidentně prostý výběr náhodné osoby a použití jediného testu, byť velmi citlivého, specifického a účinného, nejsou vhodné ani na otestování skutečného stavu populace, ani na preventivní vyšetření jednotlivců, pokud nemáme další podpůrné informace a lepší nástroje.

Výsledek asi neodpovídá naší intuici a může se zdát šokující ve vztahu k použití takovýchto testů.

Evidentně prostý výběr náhodné osoby a použití jediného testu, byť velmi citlivého, specifického a účinného, nejsou vhodné ani na otestování skutečného stavu populace, ani na preventivní vyšetření jednotlivců, pokud nemáme další podpůrné informace a lepší nástroje.

Právě matematická statistika dává nástroje na kvalifikovanější postupy v medicínské i průmyslové diagnostice, ekonomických modelech, vyhodnocování experimentálních dat atd.

# Plán přednášky

- 1 Literatura
- 2 Pravděpodobnost
- 3 Náhodné veličiny**
- 4 Pravděpodobnostní funkce a hustoty

Vraťme se k jednoduchému a názornému příkladu statistik kolem výsledků studentů<sup>1</sup> v daném předmětu. Je a není podobný klasické pravděpodobnosti a s ní související statistice při házení kostkou.

---

<sup>1</sup>Myslíme samozřejmě na „studenty a studentky“, pro zestručnění textu ale používám podobně jako v legislativních textech bezpohlavní označení „student“

Vraťme se k jednoduchému a názornému příkladu statistik kolem výsledků studentů<sup>1</sup> v daném předmětu. Je a není podobný klasické pravděpodobnosti a s ní související statistice při házení kostkou. Na jedné straně jsme připustili pouze konečný počet možných bodových hodnocení (celá čísla od 0 do 20), zároveň ale není patrně vhodné představovat si výsledky jednotlivých studentů jako analogii nezávislého házení kostkou (to by byla skutečně divně vedená přednáška).

---

<sup>1</sup>Myslíme samozřejmě na „studenty a studentky“, pro zestručnění textu ale používám podobně jako v legislativních textech bezpohlavní označení „student“



Vraťme se k jednoduchému a názornému příkladu statistik kolem výsledků studentů<sup>1</sup> v daném předmětu. Je a není podobný klasické pravděpodobnosti a s ní související statistice při házení kostkou. Na jedné straně jsme připustili pouze konečný počet možných bodových hodnocení (celá čísla od 0 do 20), zároveň ale není patrně vhodné představovat si výsledky jednotlivých studentů jako analogii nezávislého házení kostkou (to by byla skutečně divně vedená přednáška).

Místo toho máme na základním prostoru  $\Omega$  všech studentů definovanou funkci bodového ohodnocení  $X : \Omega \rightarrow \mathbb{R}$ . Je to typický příklad **náhodné veličiny**.

S každou náhodnou veličinou potřebujeme umět pracovat s vhodnou množinou jevů. Zpravidla požadujeme, abychom mohli pracovat s pravděpodobnostmi příslušnosti hodnoty  $X$  do předem zadaného intervalu.

---

<sup>1</sup>Myslíme samozřejmě na „studenty a studentky“, pro zestručnění textu ale používám podobně jako v legislativních textech bezpohlavní označení „student“

Na prostoru  $\mathbb{R}^k$  uvažujme nejmenší jevové pole  $\mathcal{B}$  obsahující všechny  $k$ -rozměrné intervaly. Množinám v  $\mathcal{B}$  říkáme **Borelovské množiny** na  $\mathbb{R}^k$ .

Na prostoru  $\mathbb{R}^k$  uvažujme nejmenší jevové pole  $\mathcal{B}$  obsahující všechny  $k$ -rozměrné intervaly. Množinám v  $\mathcal{B}$  říkáme **Borelovské množiny** na  $\mathbb{R}^k$ .

### Definition (Náhodné veličiny a distribuční funkce)

**Náhodná veličina**  $X$  na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  je taková funkce  $X : \Omega \rightarrow \mathbb{R}$ , že vzor  $X^{-1}(B)$  patří do  $\mathcal{A}$  pro každou Borelovskou množinu  $B \in \mathcal{B}$  na  $\mathbb{R}$ .

Na prostoru  $\mathbb{R}^k$  uvažujme nejmenší jevové pole  $\mathcal{B}$  obsahující všechny  $k$ -rozměrné intervaly. Množinám v  $\mathcal{B}$  říkáme **Borelovské množiny** na  $\mathbb{R}^k$ .

### Definition (Náhodné veličiny a distribuční funkce)

**Náhodná veličina**  $X$  na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  je taková funkce  $X : \Omega \rightarrow \mathbb{R}$ , že vzor  $X^{-1}(B)$  patří do  $\mathcal{A}$  pro každou Borelovskou množinu  $B \in \mathcal{B}$  na  $\mathbb{R}$ .

**Náhodný vektor**  $(X_1, \dots, X_k)$  na  $(\Omega, \mathcal{A}, P)$  je  $k$ -tice náhodných veličin.

Definice náhodné veličiny zajišťuje, že pro všechny  $-\infty \leq a \leq b \leq \infty$  existuje pravděpodobnost  $P(a \leq X < b)$ , kde používáme stručné značení pro jev  $A = (\omega \in \Omega; a \leq X(\omega) < b)$ .

### Definition

**Distribuční funkcí** náhodné veličiny  $X$  je funkce  $F : \mathbb{R} \rightarrow \mathbb{R}$  definovaná pro všechny  $x \in \mathbb{R}$  vztahem

$$F(x) = P(X < x).$$

Definice náhodné veličiny zajišťuje, že pro všechny  $-\infty \leq a \leq b \leq \infty$  existuje pravděpodobnost  $P(a \leq X < b)$ , kde používáme stručné značení pro jev  $A = (\omega \in \Omega; a \leq X(\omega) < b)$ .

### Definition

**Distribuční funkcí** náhodné veličiny  $X$  je funkce  $F : \mathbb{R} \rightarrow \mathbb{R}$  definovaná pro všechny  $x \in \mathbb{R}$  vztahem

$$F(x) = P(X < x).$$

**Distribuční funkcí** náhodného vektoru  $(X_1, \dots, X_k)$  je funkce  $F : \mathbb{R}^k \rightarrow \mathbb{R}$  definovaná pro všechny  $(x_1, \dots, x_k) \in \mathbb{R}^k$  vztahem

$$F(x) = P(X_1 < x_1 \wedge \dots \wedge X_k < x_k).$$

# Diskrétní náhodné veličiny

Předpokládejme, že pro náhodná veličina  $X$  na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  nabývá jen konečně mnoha hodnot  $x_1, x_2, \dots, x_n \in \mathbb{R}$ . Pak existuje tzv. **pravděpodobnostní funkce**  $f(x)$  taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

Evidentně  $\sum_1^n f(x_i) = 1$ .

# Diskrétní náhodné veličiny

Předpokládejme, že pro náhodná veličina  $X$  na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  nabývá jen konečně mnoha hodnot  $x_1, x_2, \dots, x_n \in \mathbb{R}$ . Pak existuje tzv. **pravděpodobnostní funkce**  $f(x)$  taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

Evidentně  $\sum_1^n f(x_i) = 1$ .

Takové náhodné veličině se říká **diskrétní**.



# Diskrétní náhodné veličiny

Předpokládejme, že pro náhodná veličina  $X$  na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  nabývá jen konečně mnoha hodnot  $x_1, x_2, \dots, x_n \in \mathbb{R}$ . Pak existuje tzv. **pravděpodobnostní funkce**  $f(x)$  taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

Evidentně  $\sum_1^n f(x_i) = 1$ .

Takové náhodné veličině se říká **diskrétní**.

Každá náhodná veličina definovaná pro klasickou pravděpodobnost je diskrétní.

# Diskrétní náhodné veličiny

Předpokládejme, že pro náhodná veličina  $X$  na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  nabývá jen konečně mnoha hodnot  $x_1, x_2, \dots, x_n \in \mathbb{R}$ . Pak existuje tzv. **pravděpodobnostní funkce**  $f(x)$  taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

Evidentně  $\sum_1^n f(x_i) = 1$ .

Takové náhodné veličině se říká **diskrétní**.

Každá náhodná veličina definovaná pro klasickou pravděpodobnost je diskrétní. Obdobně lze definici pravděpodobnostní funkce rozšířit na veličiny se spočetně mnoha hodnotami (pracujeme pak s absolutně konvergentními nekonečnými řadami :-)

# Spojité náhodné veličiny

I když hodnoty náhodné veličiny  $X$  nejsou diskrétní, můžeme postupovat podobně s užitím nástrojů diferenciálního a integrálního počtu. Intuitivně lze uvažovat takto: **hustotu**  $f(x)$  **pravděpodobnosti** pro  $X$  si představíme jako

$$P(x \leq X < x + dx) = f(x)dx.$$

# Spojité náhodné veličiny

I když hodnoty náhodné veličiny  $X$  nejsou diskrétní, můžeme postupovat podobně s užitím nástrojů diferenciálního a integrálního počtu. Intuitivně lze uvažovat takto: **hustotu**  $f(x)$  **pravděpodobnosti** pro  $X$  si představíme jako

$$P(x \leq X < x + dx) = f(x)dx.$$

To znamená, že chceme pro  $-\infty \leq a \leq b \leq \infty$

$$P(a \leq X < b) = \int_a^b f(x)dx. \quad (*)$$

# Spojité náhodné veličiny

I když hodnoty náhodné veličiny  $X$  nejsou diskrétní, můžeme postupovat podobně s užitím nástrojů diferenciálního a integrálního počtu. Intuitivně lze uvažovat takto: **hustotu**  $f(x)$  **pravděpodobnosti** pro  $X$  si představíme jako

$$P(x \leq X < x + dx) = f(x)dx.$$

To znamená, že chceme pro  $-\infty \leq a \leq b \leq \infty$

$$P(a \leq X < b) = \int_a^b f(x)dx. \quad (*)$$

## Definition

Náhodná veličina  $X$ , pro kterou existuje její **hustota pravděpodobnosti** splňující (\*), se nazývá **spojitá**.

## Theorem

Pro každou náhodnou veličinu  $X$  má její distribuční funkce  $F : \mathbb{R} \rightarrow [0, 1]$  následující vlastnosti

- 1  $F$  je neklesající funkce;
- 2  $F$  má v každém bodě  $x \in \mathbb{R}$  limitu zleva i limitu zprava;
- 3  $F$  je zleva spojitá;
- 4 v nevlastních bodech má  $F$  limity

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

- 5 pravděpodobnost, že  $X$  nabývá právě hodnotu  $x$  je dána

$$P(X = x) = \lim_{y \rightarrow x+} F(y) - F(x). \quad (2)$$

- 6 Distribuční funkce náhodné veličiny má vždy nejvýše spočetně mnoho bodů nespojitosti.

Důkaz věty je založený na pozorování vyplývajícím vcelku jednoduše z axiomů pravděpodobnosti:

### Theorem

*Uvažme pravděpodobnostní prostor  $(\Omega, \mathcal{A}, P)$  a neklesající řetězec jevů  $A_1 \subset A_2 \subset \dots$ . Pak platí*

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

*Pokud je naopak  $A_1 \supset A_2 \supset A_3 \supset \dots$ , potom platí*

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} P(A_i).$$

# Plán přednášky

- 1 Literatura
- 2 Pravděpodobnost
- 3 Náhodné veličiny
- 4 Pravděpodobnostní funkce a hustoty**



## Diskrétní náhodné veličiny

Jestliže náhodná veličina  $X$  na pravděpodobnostním prostoru  $(\Omega, \mathcal{A}, P)$  nabývá jen konečně nebo spočetně mnoha hodnot  $x_1, x_2, \dots \in \mathbb{R}$ , pak existuje **pravděpodobnostní funkce**  $f(x)$  taková, že

$$f(x) = \begin{cases} P(X = x_i) & x = x_i \\ 0 & \text{jinak.} \end{cases}$$

## Spojité náhodné veličiny

**Hustota**  $f(x)$  **pravděpodobnosti** pro náhodnou veličinu  $X$  je funkce splňující pro  $-\infty \leq a \leq b \leq \infty$

$$P(a < X < b) = \int_a^b f(x) dx. \quad (*)$$

Náhodná veličina  $X$ , pro kterou existuje její **hustota pravděpodobnosti** splňující (\*), se nazývá **spojitá**.

# Degenerované a alternativní rozdělení.

**Degenerované rozdělení**  $D(\mu)$  odpovídá konstantní hodnotě  $X = \mu$ . Distribuční funkce  $F_X$  a pravděpodobnostní funkce  $f_X$  jsou tedy rovny

$$F_X(t) = \begin{cases} 0 & t \leq \mu \\ 1 & t > \mu \end{cases} \quad f_X(t) = \begin{cases} 1 & t = \mu \\ 0 & \text{jinak} \end{cases} .$$

## Degenerované a alternativní rozdělení.

**Degenerované rozdělení**  $D(\mu)$  odpovídá konstantní hodnotě  $X = \mu$ . Distribuční funkce  $F_X$  a pravděpodobnostní funkce  $f_X$  jsou tedy rovny

$$F_X(t) = \begin{cases} 0 & t \leq \mu \\ 1 & t > \mu \end{cases} \quad f_X(t) = \begin{cases} 1 & t = \mu \\ 0 & \text{jinak} \end{cases}.$$

**Alternativní rozdělení**  $A(p)$  popisuje pokus s pouze dvěma možnými výsledky, kterým budeme říkat zdar a nezdar. Náhodné veličině  $X$  pro určitost přiřadíme hodnotu 0 pro nezdar a 1 pro zdar. Pokud má zdar pravděpodobnost  $p$ , pak nezdar musí mít pravděpodobnost  $1 - p$ . Jsou tedy distribuční a pravděpodobnostní funkce tvaru:

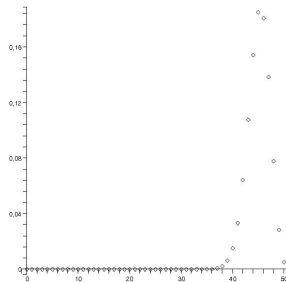
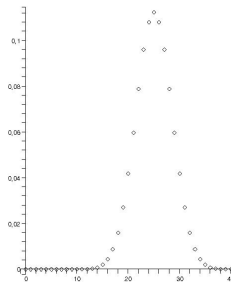
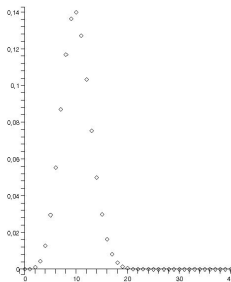
$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ 1 - p & 0 < t \leq 1 \\ 1 & t > 1 \end{cases} \quad f_X(t) = \begin{cases} p & t = 1 \\ 1 - p & t = 0 \\ 0 & \text{jinak} \end{cases}.$$

# Binomické rozdělení $Bi(n, p)$

odpovídá  $n$ -krát nezávisle opakovanému pokusu popsanému alternativním rozdělením, přičemž naše náhodná veličina měří počet zdarů. Je tedy zjevné, že pravděpodobnostní funkce bude mít nenulové hodnoty právě v celých číslech  $0, \dots, n$  odpovídajícím celkovému počtu úspěchů v pokusech (a nezáleží nám na pořadí). Je tedy

$$f_X(t) = \begin{cases} \binom{n}{t} p^t (1-p)^{1-t} & t \in \{0, 1, \dots, n\} \\ 0 & \text{jinak} \end{cases}.$$

Na obrázku jsou pravděpodobnostní funkce pro  $Bi(50, 0.2)$ ,  $Bi(50, 0.5)$  a  $Bi(50, 0.9)$ . Rozdělení pravděpodobnosti dobře odpovídá intuici, že nejvíce výsledků bude blízko u hodnoty  $np$ :



S binomickým rozdělením se setkáváme velice často v praktických úlohách. Jednou z nich je popis náhodné veličiny, která popisuje počet  $X$  předmětů v jedné zvolené přihrádce z  $n$  možných, do nichž jsme náhodně rozdělili  $r$  předmětů.

Umístění kteréhokoliv předmětu do pevně zvolené přihrádky má pravděpodobnost  $1/n$  (každá z nich je stejně pravděpodobná).

Zjevně tedy bude pro jakýkoliv počet  $k = 0, \dots, r$

$$P(X = k) = \binom{r}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{r-k} = \binom{r}{k} \frac{(n-1)^{r-k}}{n^r},$$

jde proto o rozložení  $X$  typu  $\text{Bi}(r, 1/n)$ .

Jestliže nám bude vzrůstat počet přihrádek  $n$  společně s počtem předmětů  $r_n$  tak, že v průměru nám na každou přihrádku bude připadat (přibližně) stejný počet prvků  $\lambda$ , můžeme dobře vyjádřit chování našeho rozdělení veličin  $X_n$  při limitním přechodu  $n \rightarrow \infty$ :

Jestliže nám bude vzrůstat počet přihrádek  $n$  společně s počtem předmětů  $r_n$  tak, že v průměru nám na každou přihrádku bude připadat (přibližně) stejný počet prvků  $\lambda$ , můžeme dobře vyjádřit chování našeho rozdělení veličin  $X_n$  při limitním přechodu  $n \rightarrow \infty$ :

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_n = k) &= \lim_{n \rightarrow \infty} \binom{r_n}{k} \frac{(n-1)^{r_n-k}}{n^{r_n}} \\ &= \lim_{n \rightarrow \infty} \frac{r_n(r_n-1)\dots(r_n-k+1)}{(n-1)^k} \frac{1}{k!} \left(1 - \frac{1}{n}\right)^{r_n} \\ &= \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 + \frac{-\frac{r_n}{n}}{r_n}\right)^{r_n} \\ &= \frac{\lambda^k}{k!} e^{-\lambda} \end{aligned}$$

protože obecně funkce  $(1 + x/n)^n$  konvergují stejnoměrně k funkci  $e^x$  na každém omezeném intervalu v  $\mathbb{R}$ .

To dává **Poissonovo rozdělení**  $Po(\lambda)$ .



# Poissonovo rozdělení

Poissonovo rozdělení  $Po(\lambda)$  popisuje např. události, které se vyskytují náhodně v čase a přitom pravděpodobnost výskytu v následujícím časovém intervalu o jednotkové délce nezávisí na předchozí historii a je rovna stále stejné hodnotě  $\lambda$ .

V praxi jsou takové procesy spojeny např. s poruchovostí strojů a zařízení.

# Poissonovo rozdělení

Poissonovo rozdělení  $Po(\lambda)$  popisuje např. události, které se vyskytují náhodně v čase a přitom pravděpodobnost výskytu v následujícím časovém intervalu o jednotkové délce nezávisí na předchozí historii a je rovna stále stejné hodnotě  $\lambda$ .

V praxi jsou takové procesy spojeny např. s poruchovostí strojů a zařízení.

## Theorem (Poissonova věta)

*Jsou-li  $X_n \sim \text{Bi}(n, p_n)$  a  $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda$  je konečná, pak*

$$\lim_{n \rightarrow \infty} P(X_n = k) = P(X = k),$$

*kde  $X \sim \text{Po}(\lambda)$ .*

# Příklady spojitých rozdělání

Nejjednodušší je tzv. **rovnoměrné rozdělání**. Jestliže chceme, aby pravděpodobnost každé hodnoty v předem daném intervalu  $(a, b) \subset \mathbb{R}$  byla stejná, pak hustota  $f_X$  našeho rozdělání náhodné veličiny  $X$  má být konstantní. Pak ovšem jsou pro libovolná reálná čísla  $-\infty < a < b < \infty$  jen jediné možné hodnoty

$$f_X(t) = \begin{cases} 0 & t \leq a \\ \frac{1}{b-a} & t \in (a, b) \\ 0 & t \geq b, \end{cases} \quad F_X(t) = \begin{cases} 0 & t \leq a \\ \frac{t-a}{b-a} & t \in (a, b) \\ 1 & t \geq b. \end{cases}$$

# Exponenciální rozdělení $\text{ex}(\lambda)$

je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme výskyt náhodného jevu tak, že výskyty v nepřekrývajících se intervalech jsou nezávislé. Je-li tedy  $P(t)$  pravděpodobnost, že jev **nenastane** během intervalu délky  $t$ , pak nutně  $P(t + s) = P(t)P(s)$  pro všechna  $t, s > 0$ . Předpokládejme navíc diferencovatelnost funkce  $P$  a  $P(0) = 1$ . Pak  $\ln P(t + s) = \ln P(t) + \ln P(s)$ .

# Exponenciální rozdělení $\text{ex}(\lambda)$

je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme výskyt náhodného jevu tak, že výskyty v nepřekrývajících se intervalech jsou nezávislé. Je-li tedy  $P(t)$  pravděpodobnost, že jev **nenastane** během intervalu délky  $t$ , pak nutně  $P(t+s) = P(t)P(s)$  pro všechna  $t, s > 0$ . Předpokládejme navíc diferencovatelnost funkce  $P$  a  $P(0) = 1$ . Pak  $\ln P(t+s) = \ln P(t) + \ln P(s)$ . Limitním přechodem:

$$\lim_{s \rightarrow 0_+} \frac{\ln P(t+s) - \ln P(t)}{s} = (\ln P)'(0) = -\lambda.$$

Odtud vyplývá diferenciální rovnice

$$(\ln P(t))' = -\lambda.$$

Odtud dostáváme  $\ln P(t) = -\lambda t + C$  a počáteční podmínka dává jediné řešení

$$P(t) = e^{-\lambda t}.$$

Všimněme si, že z definice našich objektů vyplývá, že  $\lambda > 0$ .

Odtud dostáváme  $\ln P(t) = -\lambda t + C$  a počáteční podmínka dává jediné řešení

$$P(t) = e^{-\lambda t}.$$

Všimněme si, že z definice našich objektů vyplývá, že  $\lambda > 0$ . Uvažme náhodnou veličinu  $X$  udávající okamžik, kdy náš jev poprvé **nastane**. Zřejmě tedy je distribuční funkce rozdělení pro  $X$  dána

$$F_X(t) = 1 - P(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

Je vidět, že je to rostoucí funkce s hodnotami mezi nulou a jedničkou a správnými limitami v  $\pm\infty$ .

Hustotu tohoto rozdělení dostaneme derivováním distribuční funkce, tj.

$$f_X = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

# Normální rozdělení

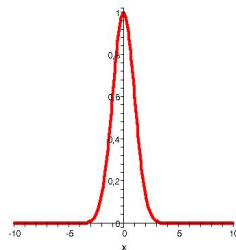
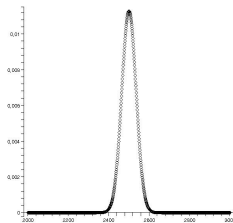
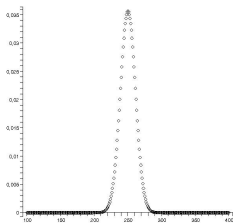
je ze všech nejdůležitější.



# Normální rozdělení

je ze všech nejdůležitější.

Jestliže v binomiálním rozdělení zachováme konstantní úspěšnost  $p$ , ale budeme přidávat počet pokusů  $n$ , bude pravděpodobnostní funkce kupodivu pořád mít podobný tvar (i když jiné rozměry). Na obrázku při rostoucím  $n$  se budou vynesené bodové hodnoty slévat do křivky, pro hodnoty  $Bi(500, 0.5)$  a  $Bi(5000, 0.5)$  je výsledek vidět na obrázku níže. Třetí křivka na obrázku je grafem funkce  $f(x) = e^{-x^2/2}$ .



Hledáme-li podobné spojitě rozdělení, potřebovali bychom spočítat  $\int_a^b e^{-x^2/2} dx$  což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Hledáme-li podobné spojité rozdělení, potřebovali bychom spočítat  $\int_a^b e^{-x^2/2} dx$  což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Rozdělení s touto hustotou se nazývá **normální rozdělení**  $N(0, 1)$ . Příslušnou distribuční funkci

$$F_X(x) = \int_{-\infty}^x e^{-x^2/2} dx$$

nelze vyjádřit pomocí elementárních funkcí, přesto se s ní numericky běžně počítá (pomocí tabulek nebo softwarových aplikací).

Hledáme-li podobné spojitě rozdělení, potřebovali bychom spočítat  $\int_a^b e^{-x^2/2} dx$  což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}.$$

Odtud vyplývá, že možná hustota rozdělení náhodného rozdělení může být

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Rozdělení s touto hustotou se nazývá **normální rozdělení**  $N(0, 1)$ . Příslušnou distribuční funkci

$$F_X(x) = \int_{-\infty}^x e^{-x^2/2} dx$$

nelze vyjádřit pomocí elementárních funkcí, přesto se s ní numericky běžně počítá (pomocí tabulek nebo softwarových aplikací). Hustotě  $f_X$  se také často říká **Gaussova křivka**.