Matematika III – 10. přednáška
Náhodné veličiny – základní vlastnosti a typy

Michal Bulant

Masarykova univerzita
Fakulta informatiky

24. 11. 2014

Obsah přednášky

Doporučené zdroje

- Jan Slovák, Martin Panák, Michal Bulant, *Matematika drsně a svižně*, MU Brno, 2013, 774 s. (též jako e-text).
- Karel Zvára, Josef Štěpán, **Pravděpodobnost a matematická statistika**, Matfyzpress, 4. vydání, 2006, 230 stran, ISBN 80-867-3271-1.
- 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, **Statistika**, Masarykova univerzita, 2004, distanční studijní opora ESF, `http://www.math.muni.cz/~budikova/esf/Statistika.zip`.
- Marie Budíková, Tomáš Lerch, Štěpán Mikoláš, **Základní statistické metody**, Masarykova univerzita, 2005, 170 stran, ISBN 80-210-3886-1.

Náhodné veličiny

Vraťme se k jednoduchému a názornému příkladu statistik kolem výsledků studentů v daném předmětu, který 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 (v případě `MB103` celá čísla od 0 do 45), 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 Ω všech studentů definovánu funkci bodového ohodnocení $X : \Omega \to \mathbb{R}$. Je to typický příklad **náhodné veličiny**.

U každé náhodné veličiny potřebujeme umět pracovat s vhodnou množinou jevů. Zpravidla požadujeme, abychom mohli pracovat s pravděpodobnostmi příslušnosti hodnoty X do předem zadaného intervalu.

Přirozenější interpretací výsledku pokusu je totiž často spíše než zjištění, zda konkrétní náhodný jev *nastal či nenastal*, nějaká hodnota:

- součet bodů na dvou kostkách,
- počet bakterií v daném množství roztoku nebo
- počet studentů, kteří uspěli u zkoušky nebo kteří získali alespoň 5 bodů z konkrétního příkladu.

Od pravděpodobnostního prostoru $(\Omega, \mathcal{A}, P)$ tedy potřebujeme přejít k obdobné dvojici $(\mathbb{R}, \mathcal{B})$ tak, abychom podmnožinám $\mathbb{R}$, ležícím v σ-algebře $\mathcal{B}$ byli schopni přiřadit pravděpodobnost odvozenou z $(\Omega, \mathcal{A}, P)$.

Na prostoru $\mathbb{R}^k$ uvažujme nejmenší jevové pole $\mathcal{B}$ obsahující všechny k–rozměrné intervaly. Množinám v $\mathcal{B}$ říkáme **borelovské množiny** (nebo také měřitelné množiny) na $\mathbb{R}^k$.
Speciálně pro $k = 1$ jde o množiny, které obdržíme **z intervalů konečnými průniky a nejvýše spočetnými sjednoceními**.

Definice

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

$$P_X(B) = P(X^{-1}(B)) = P(\{\omega \in \Omega; X(\omega) \in B\})$$

se nazývá **rozdělení pravděpodobnosti** náhodné veličiny X.
Náhodný vektor $(X_1, \dots, X_k)$ na $(\Omega, \mathcal{A}, P)$ je k–tice náhodných veličin.

Příklady k procvičení

Příklad

Hodíme jedenkrát kostkou, množina elementárních jevů je $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. Jevovým polem nechť je $\mathcal{A} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4, \omega_5, \omega_6\}, \Omega\}$.
Zjistěte jestli zobrazení $X : \Omega \to \mathbb{R}$ dané předpisem

a) $X(\omega_i) = i$ pro každé $i \in \{1, 2, 3, 4, 5, 6\}$,

b) $X(\omega_1) = X(\omega_2) = -2, X(\omega_3) = X(\omega_4) = X(\omega_5) = X(\omega_6) = 3$,

je náhodnou veličinou vzhledem k $\mathcal{A}$.

Příklad

Je dáno jevové pole $(\Omega, \mathcal{A})$, kde $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ a

$$\mathcal{A} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4, \omega_5\}, \{\omega_1, \omega_2, \omega_3\},$$
$$\{\omega_1, \omega_2, \omega_4, \omega_5\}, \{\omega_3, \omega_4, \omega_5\}, \Omega\}.$$

Najděte nějaké (co nejobecnější) zobrazení $X : \Omega \to \mathbb{R}$, které bude náhodnou veličinou vzhledem k $\mathcal{A}$.

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

Definice

Distribuční funkcí (*distribution, cumulative density function*) náhodné veličiny X je funkce $F : \mathbb{R} \to \mathbb{R}$ definovaná pro všechny $x \in \mathbb{R}$ vztahem

$$F(x) = P(X \leq x).$$

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

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

Diskrétní náhodné veličiny

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

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

Evidentně $\sum_{i=1}^{n} f(x_i) = 1$.
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 nekonečnými řadami).

Příklady k procvičení

Příklad

Nechť $\Omega = \{\omega_1, \omega_2, \omega_3\}$ a $\mathcal{A} = \{\Omega, \emptyset, \{\omega_3\}, \{\omega_1, \omega_2\}\}$. Určete všechny pravděpodobnostní funkce zobrazující $\mathcal{A}$ do množiny $\{0, 1, \theta, 1-\theta\}$.

Příklad

Třikrát nezávisle na sobě hodíme mincí. Náhodná veličina X udává počet hlav, které padnou při těchto hodech. Určete pravděpodobnostní a distribuční funkci náhodné veličiny X.

Příklad

Pravděpodobnost, že výrobek bude vyhovovat všem technickým požadavkům, je 0,9. Popište rozdělení náhodné veličiny udávající počet nevyhovujících výrobků mezi 3 výrobky.

Spojité náhodné veličiny

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

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

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

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

Definice

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

Příklady k procvičení

Příklad

Rozhodněte, které z následujících funkcí jsou hustotami (mimo vymezený interval je vždy funkce nulová, c je vhodná konstanta – v případě, že jde o hustotu, tuto konstantu určete):

1. c pro $x \in (a, b)$,
2. cx pro $x \in (0, 1)$,
3. cx pro $x \in (-1, 2)$,
4. $cx \sin x$ pro $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$,
5. ce^x pro $x \in (0, \infty)$,
6. ce^{-x} pro $x \in (0, \infty)$,
7. $\frac{c}{1+x^2}$.

Příklad

V lese tvaru trojúhelníka s vrcholy v bodech $(-1, 0), (1, 0)$ a $(0, \sqrt{3})$ se ztratilo dítě. Pravděpodobnost výskytu dítěte v určité části lesa je úměrná velikosti této části, nikoliv umístění této části. Určete

1. rozdělení vzdálenosti dítěte od zvolené strany lesa,
2. rozdělení vzdálenosti dítěte od nejbližší strany lesa.

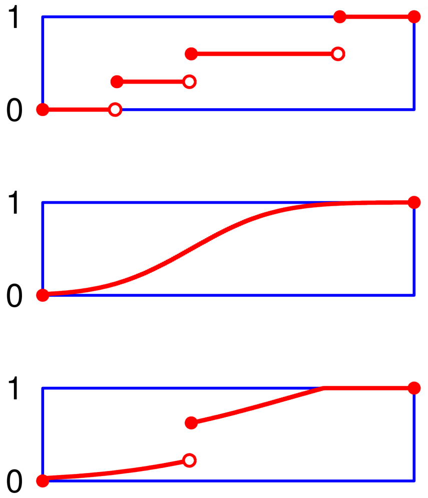
Vlastnosti distribuční funkce

Věta

Nechť X je náhodná veličina, $F(x)$ je její distribuční funkce.

1. *F je neklesající.*
2. *F je zprava spojitá, $\lim_{x\to-\infty} F(x) = 0$ a $\lim_{x\to\infty} F(x) = 1$.*
3. *Je-li X diskrétní s hodnotami $x_1, \dots, x_n$, pak je $F(x)$ po částech konstantní, $F(x) = \sum_{x_i \le x} P(X = x_i)$ a $F(x) = 1$ kdykoliv $x \ge x_n$.*
4. *Je-li X spojitá, pak je $F(x)$ diferencovatelná a její derivace se rovná hustotě X, tj. platí $F'(x) = f(x)$.*

Distribuční funkce - příklady



Obdobně definujeme distribuční funkce a hustotu a pravděpodobnostní funkci pro spojité a diskrétní náhodné **vektory**.
Hovoříme také o **simultánních pravděpodobnostních funkcích a hustotách**.
Pro dvě proměnné (vektor (X, Y) náhodných veličin):

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

u diskrétních a pro všechny $a, b \in \mathbb{R}$ pro spojité:

$$P(-\infty < X \leq a, -\infty < Y \leq b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x, y) dxdy.$$

Marginální rozložení pro jednu z proměnných obdržíme tak, že přes ostatní posčítáme nebo zintegrujeme.
Náhodné veličiny X a Y jsou **stochasticky nezávislé**, jestliže je jejich simultánní distribuční funkce splňují

$$F(x, y) = G(x) \cdot H(y)$$

kde G a H jsou distribuční funkce veličin X a Y.

Příklady k procvičení

Příklad

V zásilce s 10 výrobky je 8 kvalitních (z nich je 5 první jakosti a 3 jsou druhé jakosti) a 2 zmetky. Ze zásilky vybereme bez vracení 2 výrobky. Náhodná veličina X nechť značí počet vybraných kvalitních výrobků a Y počet vybraných výrobků první jakosti. Určete sdruženou i marginální pravděpodobností funkci a rozhodněte, zda jsou náhodné veličiny X a Y stochasticky nezávislé.

Příklad

Spojitý náhodný vektor (X, Y) má hustotu

$$f(x, y) = 24x^2y(1 - x)$$

pro $0 \leq x, y < 1$ a jinde nulovou. Dokažte, že X a Y jsou stochasticky nezávislé.

Rovnoměrné (diskrétní) rozdělení popisuje náhodnou veličinu, která může nabývat konečně mnoha hodnot se stejnou pravděpodobností, značíme $X \sim Rd(n)$ (např. $X \sim Rd(6)$ odpovídá hodu kostkou).

Alternativní rozdělení popisuje pokus se dvěma možnými výsledky, často nazývanýni *zdar*, resp. *nezdar*. Náhodná veličina $X \sim A(p)$ nabývá hodnoty 1 (*zdar)* s pravděpodobností p. Distribuční a pravděpodobnostní funkce jsou tedy tvaru:

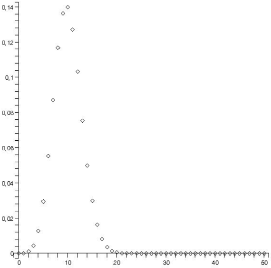
$$F_X(t) = \begin{cases} 0 & t < 0 \\ 1 - p & 0 \leq t < 1 \\ 1 & t > 1 \end{cases} \qquad f_X(t) = \begin{cases} p & t = 1 \\ 1 - p & t = 0 \\ 0 & \text{jinak} \end{cases}.$$

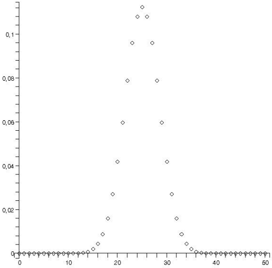
Binomické rozdělení $\mathrm{Bi}(n, p)$ odpovídá n–krát nezávisle opakovanému pokusu popsanému alternativním rozdělením, přičemž naše náhodná veličina měří počet zdarů. Je tedy

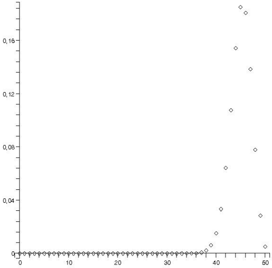
$$f_X(t) = \begin{cases} \binom{n}{t} p^t (1-p)^{n-t} & t \in \{0, 1, \ldots, n\} \\ 0 & \text{jinak} \end{cases}.$$

Binomické rozdělení

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







Binomické rozdělení

S binomickým rozdělením se potkáváme 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 $\mathrm{Bi}(r, 1/n)$.

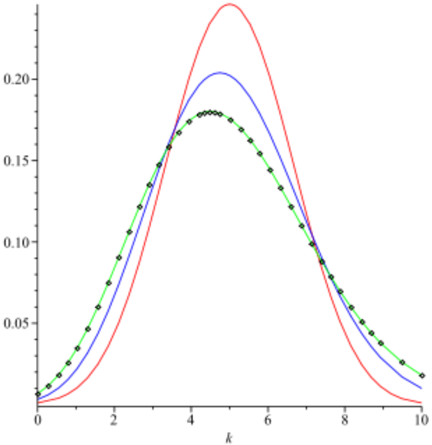
Binomické $\to$ Poissonovo rozdělení

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 λ, můžeme dobře vyjádřit chování našeho rozdělení veličin X_n při limitním přechodu $n \to \infty$. Takovéto chování popisuje např. fyzikální soustavy s velikým počtem molekul plynu. Standardní úpravy vedou při $\lim_{n\to\infty} r_n/n = \lambda$ k výsledku:

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

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

Binomické → Poissonovo rozdělení



Tečky znázorňují Poissonovo rozdělení Po(5), červeně Bi($10, \frac{1}{2}$), modře Bi($20, \frac{1}{4}$) a zeleně Bi($1000, \frac{1}{200}$)

Poissonovo rozdělení Po(λ)

Poissonovo rozdělení popisuje náhodné veličiny s pravděpodobnostní funkcí

$$f_X(k) = \begin{cases} \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda} & k \in \mathbb{N} \\ 0 & \text{jinak.} \end{cases}$$

Jak jsme odvodili výše, toto diskrétní rozdělení (rozložené do nekonečně mnoha bodů) dobře aproximuje binomická rozdělení Bi(n, λ) pro konstantní $\lambda > 0$ a veliká n.
Snadno ověříme

$$\sum_{k=0}^{\infty} f_X(k) = \sum_k \frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda} \sum_k \frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda+\lambda} = 1.$$

Poissonovo rozdělení

Dobře modeluje výskyt jevů:

- s očekávanou konstantní hustotou na jednotku objemu – např. bakterie ve vzorku (popis očekávaného výskytu k bakterií při rozdělení vzorku na n stejných částí)
- rozdělení událostí, které se vyskytují náhodně v čase a bez závislosti na předchozí historii – v praxi jsou takové procesy často spojeny s poruchovostí strojů a zařízení

Příklad

- počet branek ve fotbalovém zápase (za 90 minut)
- počet telefonních hovorů za minutu na call centru
- počet aut přijíždějících na křižovatku
- ...

Geometrické rozdělení

Geometrické rozdělení má náhodná veličina $X \sim \mathrm{Ge}(p)$, která udává celkový počet *nezdarů*, které v posloupnosti opakovaných pokusů předcházejí prvnímu *zdaru*, přičemž pravděpodobnost úspěchu v každém pokusu je rovna p .

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

Hypergeometrické rozdělení. Mějme N předmětů, z nichž právě M má danou vlastnost. Z těchto N předmětů náhodně vybereme n předmětů bez vracení. Náhodná veličina $X \sim \mathrm{Hg}(N, M, n)$ udává počet vybraných prvků s danou vlastností. Zřejmě tato náhodná velišina může nabývat pouze celočíselných hodnot z intervalu $[\max\{0, M-N+n\}, \min\{n, M\}]$. Pro t z tohoto intervalu pak

$$f_X(t) = \frac{\binom{M}{t}\binom{N-M}{n-t}}{\binom{N}{n}}.$$

Typy spojitých náhodných veličin

Rovnoměrné spojité rozdělení $\mathrm{Rs}(a, b)$ je nejjednodušším příkladem spojitého rozdělení. Ilustruje, že při jednoduše formulovaném požadavku na chování rozdělení nám nezbude moc prostoru pro jeho definici. Nyní chceme, aby pravděpodobnost každé hodnoty v předem daném intervalu $(a, b) \subset \mathbb{R}$ byla stejná, tj. hustota f_X našeho rozdělení náhodné veličiny X má být konstantní. Pak ovšem jsou pro libovolná reálná čísla $-\infty < a < b < \infty$ jen jediné možné hodnoty

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

Exponenciální rozdělení $\mathrm{Ex}(\lambda)$ je dalším rozdělením, které je snadno určeno požadovanými vlastnostmi náhodné veličiny. Předpokládejme, že sledujeme náhodný jev, jehož výskyty v nepřekrývajících se časových 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 jistě $\ln P(t+s) = \ln P(t) + \ln P(s)$, takže limitním přechodem

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

Označme si spočtenou derivaci zprava v nule jako $-\lambda \in \mathbb{R}$. Pak tedy pro $P(t)$ platí $\ln P(t) = -\lambda t + C$ a počáteční podmínka dává jediné řešení

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

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

Nyní uvažme náhodnou veličinu X udávající (náhodný) okamžik, kdy náš jev poprvé nastane (je vidět analogie s geometrickým rozdělením?). Zřejmě tedy je distribuční funkce rozdělení pro X dána

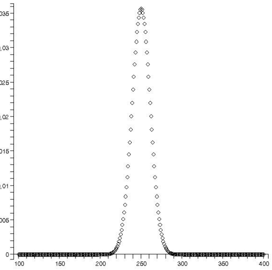
$$F_X(t) = 1 - P(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

Je vidět, že skutečně jde rostoucí funkci s hodnotami mezi nulou a jedničkou a správnými limitami v $\pm\infty$.
Hustotu tohoto rozdělení dostaneme derivováním distribuční funkce, tj.

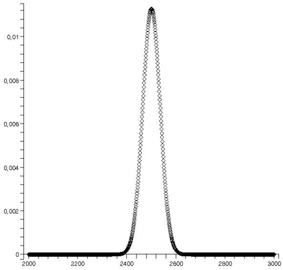
$$f_X = \begin{cases} \lambda e^{-\lambda t} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

Normální rozdělení

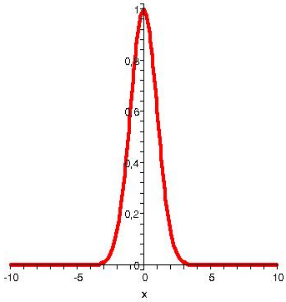
Jde o nejdůležitější rozdělení. Uveďme nejprve motivaci pro jeho zavedení.
Pokud budeme v **binomickém rozdělení** $\mathrm{Bi}(n, p)$ zvyšovat n při zachování úspěšnosti p, bude mít pravděpodobnostní funkce pořád přibližně stejný tvar.



$\mathrm{Bi}(500, 0.5)$



$\mathrm{Bi}(5000, 0.5)$



graf funkce $e^{-x^2/2}$

Normální rozdělení $N(0,1)$

Vzhledem k uvedené motivaci se nabízí hledat vhodné spojité rozdělení, které by mělo hustotu danou nějakou obdobnou funkcí. Protože je $\mathrm{e}^{-x^2/2}$ vždy kladná funkce, potřebovali bychom spočíst $\int_a^b \mathrm{e}^{-x^2/2}\,dx$ což není pomocí elementárních funkcí možné. Je však možné (i když ne úplně snadné) ověřit, že příslušný nevlastní integrál konverguje k hodnotě

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

Odtud vyplývá, že hustota rozdělení náhodné veličiny může být

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

Rozdělení s touto hustotou se nazývá **normální rozdělení** $\mathrm{N}(0,1)$.

Normální rozdělení $N(0,1)$

Příslušnou distribuční funkci

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

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**.

Abychom uměli přesněji zformulovat asymptotickou blízkost normálního a binomického rozdělení pro $n \to \infty$, musíme si vytvořit další nástroje pro práci s náhodnými veličinami. Budeme k tomu používat funkce dvojím různým způsobem.

Příklady k procvičení

Příklad

Nechť má X binomické rozdělení s parametry $n = 4, p = 2/3$. Určete rozdělení transformované náhodné veličiny $Y = (X - 2)^2$ a nakreslete graf její distribuční funkce.

Příklad

Mějme náhodnou veličinu X hustoty $f(x) = 2xe^{-x^2}$ pro $x > 0$ (a jinde nulové). Určete hustotu pravděpodobnosti náhodné veličiny $Y = X^2$.

Transformace náhodných veličin

Příklad

Nechť má náhodná veličina X rovnoměrné rozdělení na intervalu $\langle 0, r \rangle$. Určete distribuční funkci a hustotu pravděpodobnosti rozdělení objemu koule o poloměru X.

Řešení

Určeme nejprve distribuční funkci F (pro $0 < d < \frac{4}{3}\pi r^3$)

$$F(d) = P\left[\frac{4}{3}\pi X^3 \leq d\right] = P\left[X \leq \sqrt[3]{\frac{3d}{4\pi}}\right] = \frac{\sqrt[3]{\frac{3d}{4\pi}}}{r},$$

celkem

$$F(x) = \begin{cases} 0 & \text{pro} \quad x \leq 0 \\ \sqrt[3]{\frac{3}{4\pi r^3}}\, x^{\frac{1}{3}} & \text{pro} \quad 0 < x < \frac{4}{3}\pi r^3 \\ 1 & \text{pro} \quad x \geq \frac{4}{3}\pi r^3 \end{cases}$$

Příklad (rozdělení $\chi^2(1)$)

Nechť Z má normované normální rozdělení. Určete hustotu transformované náhodné veličiny $X = Z^2$.

Řešení

Zřejmě je pro $x \leq 0$ distribuční funkce nulová, pro $x > 0$ dostáváme: $F_X(x) = P[Z^2 < x] = P[-\sqrt{x} < Z < \sqrt{x}] =$

$$= \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z^2}{2}} \, \mathrm{d}z = 2 \int_0^x \frac{1}{2\sqrt{2\pi}} t^{-\frac{1}{2}} \mathrm{e}^{-\frac{t}{2}} \, \mathrm{d}t$$

a derivací podle x dostaneme hustotu $f_X(x) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} \mathrm{e}^{-\frac{x}{2}}$. Rozdělení náhodné veličiny s touto hustotou se nazývá (Pearsonovo) χ^2 rozdělení s jedním stupněm volnosti a značí se $X \sim \chi^2(1)$.