

Náhodné veličiny

Zavedení náhodné veličiny a transformované náhodné veličiny

Výsledky náhodného pokusu lze popsat reálnými čísly (resp. reálnými vektory) pomocí nějakého zobrazení $X: \Omega \rightarrow \mathbb{R}^n$, tj. zobrazení, které možnému výsledku ω přiřadí číslo $X(\omega)$ nebo několik čísel $X_1(\omega), \dots, X_n(\omega)$. Např. při hodu kostkou lze poloze kostky číslem i nahoru (tj. možnému výsledku ω_i) přiřadit číslo i , $i = 1, 2, \dots, 6$.

Pokud bude toto zobrazení splňovat podmínku tzv. měřitelnosti, tj. pro všechna reálná čísla x je množina

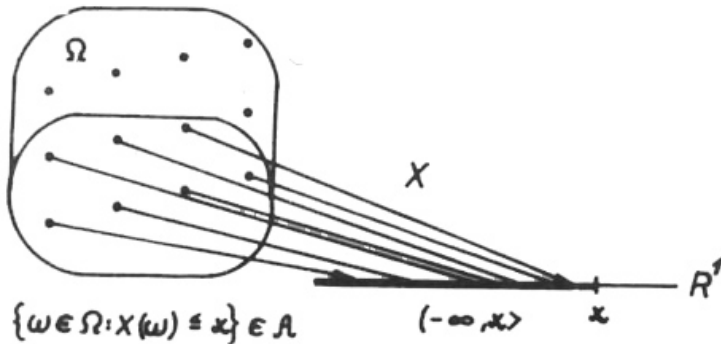
$\{\omega \in \Omega : X(\omega) \leq x\}$ jev vzhledem k jevovému poli \mathcal{A} , pak se X nazývá **náhodná veličina** a $\mathbf{X} = (X_1, \dots, X_n)$ se nazývá **náhodný vektor** (se složkami X_1, \dots, X_n). Číslo $X(\omega)$ se nazývá **číselná realizace** náhodné veličiny X příslušná možnému výsledku ω .

(Nehrozí-li nebezpečí nedorozumění, náhodnou veličinu i její číselnou realizaci značíme týmž symbolem X .)

V některých situacích potřebujeme náhodnou veličinu X transformovat pomocí funkce g na **transformovanou náhodnou**

veličinu $Y = g(X)$. Např. X – průměr náhodně vybrané kuličky do kuličkového ložiska, $Y = \frac{4}{3}\pi \left(\frac{X}{2}\right)^3$ - objem kuličky.

Ilustrace náhodné veličiny



Vztah mezi znakem a náhodnou veličinou

Pojem „znak“, který jsme zavedli v popisné statistice, je sice blízký pojmu „náhodná veličina“, ale není s ním totožný. Znak může být považován za náhodnou veličinu, jestliže jeho hodnoty zjišťujeme na objektech, které byly vybrány ze základního souboru náhodně.

Simultánní a marginální náhodný vektor

Jestliže z náhodného vektoru (X_1, \dots, X_n) vybereme některé složky, např. X_k, \dots, X_l , dostaneme **marginální náhodný vektor** (X_k, \dots, X_l) . Původní náhodný vektor se v této souvislosti nazývá **simultánní náhodný vektor**.

Např. výsledky pěti zkoušek na konci semestru považujeme za náhodné veličiny X_1, \dots, X_5 , přičemž X_3, X_4 jsou známky ze dvou nejdůležitějších předmětů. Náhodný vektor (X_1, \dots, X_5) je simultánní, náhodný vektor (X_3, X_4) je marginální.

Zapisování jevů pomocí náhodných veličin

Zápis $\omega \in \Omega \Rightarrow X(\omega) \in B$ znamená jev, že náhodná veličina X se realizovala v číselné množině B . Zkráceně píšeme $X \in B$.

Je-li $B =]-\infty, x]$ nebo $B = \{x\}$, jev $X \in B$ znamená, že náhodná veličina X se realizovala hodnotou nejvýše x a jev $X = x$ znamená, že náhodná veličina X se realizovala hodnotou x .

Popis pravděpodobnostního chování náhodné veličiny

Při pozorování realizací náhodné veličiny si povšimneme, že některé její hodnoty se vyskytují s větší pravděpodobností, jiné s menší. Pravděpodobnostní chování náhodné veličiny X budeme popisovat pomocí **distribuční funkce**, která udává pravděpodobnost jevu, že náhodná veličina X se realizuje hodnotou nejvýše x :

$$\forall x \in \mathbb{R} : \Phi(x) = P(X \leq x)$$

Je to zidealizovaný protějšek empirické distribuční funkce zavedené v popisné statistice:

$$\forall x \in \mathbb{R} : F_n(x) = \frac{N(X \leq x)}{n}$$

Lze očekávat, že s rostoucím rozsahem výběrového souboru se budou hodnoty empirické distribuční funkce $F(x)$ ustalovat kolem hodnot distribuční funkce $\Phi(x)$. Vlastnosti empirické distribuční funkce se přenášejí i na distribuční funkci – je to funkce neklesající, zprava spojitá a normovaná v tom smyslu, že

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

Podobně se definuje i **simultánní distribuční funkce náhodného vektoru** (X_1, \dots, X_n) :

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^n : \Phi_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = P(X_1 \leq x_1 \wedge \dots \wedge X_n \leq x_n)$$

Distribuční funkce libovolného marginálního náhodného vektoru se nazývá **marginální distribuční funkce**.

Největší význam mají jednorozměrné marginální distribuční funkce $\Phi_1(x_1), \dots, \Phi_n(x_n)$ jednotlivých složek náhodného vektoru.

Diskrétní náhodné veličiny

V praxi mají značný význam náhodné veličiny, které nabývají pouze konečně nebo spočetně mnoha hodnot – jsou to **diskrétní náhodné veličiny**.

Příklady diskrétních náhodných veličin: počet chyb, jichž se dopustí nějaké zařízení za určitou dobu, počet zákazníků ve frontě, počet zmetků v denní produkci apod.

Pravděpodobnostní chování diskrétní náhodné veličiny popisujeme **pravděpodobnostní funkcí**:

$$\forall x \in \mathbb{R} : \pi(x) = P\{K = x\}$$

Je to zidealizovaný protějšek četnostní funkce zavedené v popisné statistice v souvislosti s bodovým rozložením četností:

$$\forall x \in \mathbb{R} : p(x) = \frac{N(x)}{n}$$

S rostoucím rozsahem výběrového souboru se budou hodnoty četnostní funkce ustalovat kolem hodnot pravděpodobnostní funkce.

Vlastnosti četnostní funkce se přenášejí i na pravděpodobnostní funkci, tedy pravděpodobnostní funkce

je nezáporná $\forall x \in \mathbb{R} : \pi(x) \geq 0$,

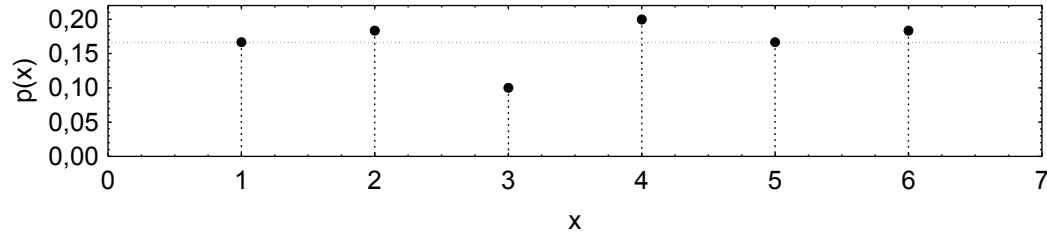
je normovaná $\sum_{x=-\infty}^{\infty} \pi(x) = 1$,

s distribuční funkcí je spjata součtovým vztahem $\forall x \in \mathbb{R} : \Phi(x) = \sum_{t \leq x} \pi(t)$

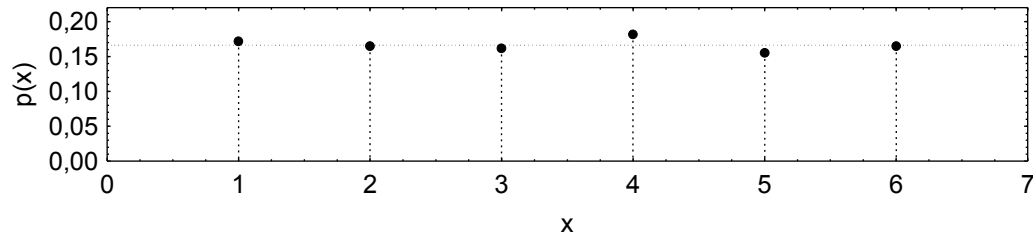
Ilustrace vztahu mezi četnostní funkcí a pravděpodobnostní funkcí

Provedeme n hodů kostkou. Zajímáme se o četnostní funkci počtu ok.

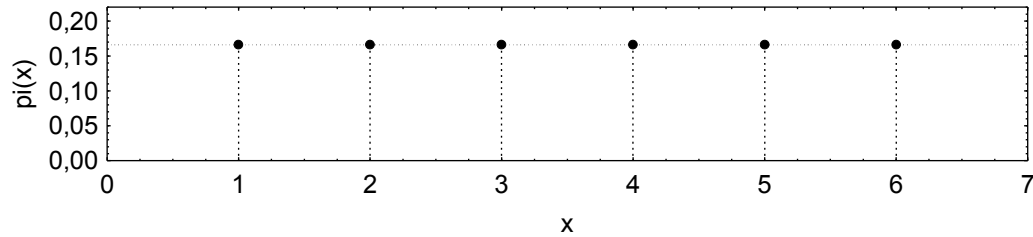
$n = 60$



$n = 600$



$n \rightarrow \infty$



Diskrétní náhodný vektor

Jestliže všechny složky náhodného vektoru (X_1, \dots, X_n) jsou diskrétní náhodné veličiny, hovoříme o **diskrétním náhodném vektoru**. Jeho pravděpodobnostní chování je popsáno **simultánní pravděpodobnostní funkcí**:

$$\forall \xi_1, \dots, \xi_n \in \mathcal{X}^n : \pi(\xi_1, \dots, \xi_n) = \mathbb{P}(X_1 = \xi_1 \wedge \dots \wedge X_n = \xi_n)$$

Pravděpodobnostní funkce libovolného diskrétního marginálního náhodného vektoru se nazývá marginální pravděpodobnostní funkce. Největší význam mají jednorozměrné marginální pravděpodobnostní funkce $\pi_1(x_1), \dots, \pi_n(x_n)$ jednotlivých složek náhodného vektoru. Získáme je tak, že hodnoty simultánní pravděpodobnostní funkce sečteme přes všech $n-1$ přebývajících proměnných.

Spojité náhodné veličiny

Dalším velmi důležitým typem veličin jsou **spojité náhodné veličiny** – ty nabývají všech hodnot z nějakého intervalu.

Příklady spojitých náhodných veličin: rozměry sériově vyráběných výrobků, hektarový výnos nějaké zemědělské plodiny, výsledky různých fyzikálních a chemických měření apod.

Pravděpodobnostní chování spojitě náhodné veličiny popisujeme **hustotou pravděpodobnosti** $\varphi(x)$, což je zidealizovaný protějšek hustoty četnosti $f(x)$ zavedené v popisné statistice v souvislosti s intervalovým rozložením četností. S rostoucím rozsahem výběrového souboru a klesajícími šířkami třídících intervalů se budou hodnoty hustoty četnosti ustalovat kolem hodnot hustoty pravděpodobnosti.

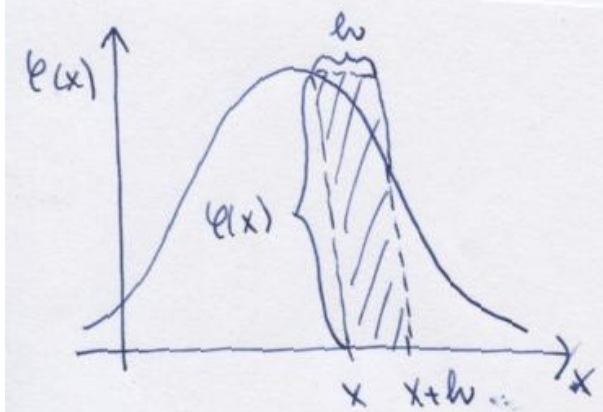
Vlastnosti hustoty četnosti se přenášejí i na hustotu pravděpodobnosti, tedy hustota pravděpodobnosti

je nezáporná $\forall x \in \mathbb{R} : \varphi(x) \geq 0$,

je normovaná $\int_{-\infty}^{\infty} \varphi(x) dx = 1$,

s distribuční funkcí je spjata integrálním vztahem $\forall x \in \mathbb{R} : \Phi(x) = \int_{-\infty}^x \varphi(t) dt$.

Pozor – na rozdíl od pravděpodobnostní funkce diskrétní náhodné veličiny nemá hustota pravděpodobnosti spojité náhodné veličiny význam pravděpodobnosti! Její význam lze odvodit z integrálního vztahu mezi distribuční funkcí a hustotou pravděpodobnosti.



Pravděpodobnost, že náhodná veličina se bude realizovat v intervalu $(x, x+h]$, je:

$$P\{x < X \leq x+h\} = P\{x+h \leq X < x\} = \int_{-\infty}^{x+h} \varphi(t) dt - \int_{-\infty}^x \varphi(t) dt = \int_x^{x+h} \varphi(t) dt$$

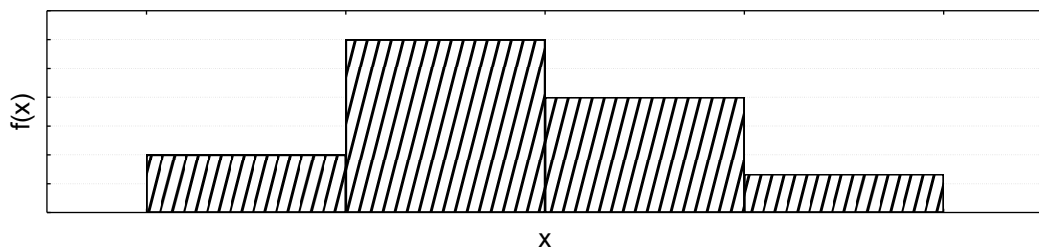
Bude-li h dostatečně malé číslo, lze plochu pod grafem hustoty nahradit

obsahem obdélníka o stranách $\varphi(x)$ a h , tj. $P\{x < X \leq x+h\} \approx \varphi(x) \cdot h$

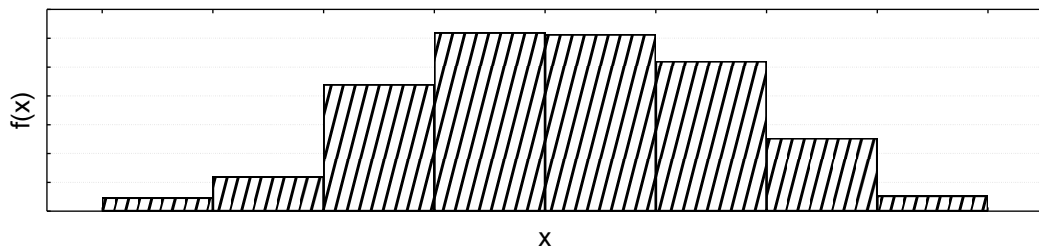
Ilustrace vztahu mezi hustotou četnosti a hustotou pravděpodobnosti

Náhodně vybereme n sériově vyráběných součástek, změříme jejich délku a budeme se zajímat hustotu četnosti odchylek těchto měření od deklarované délky součástky.

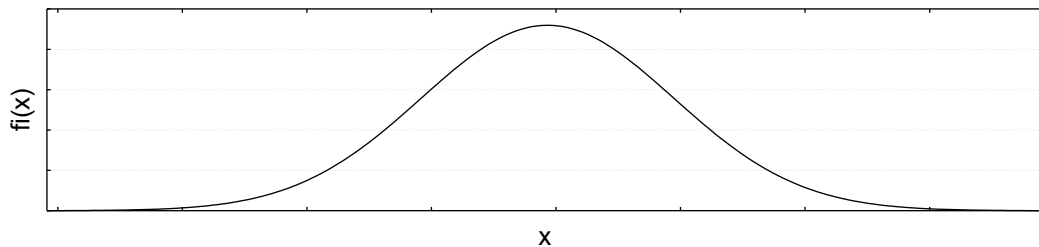
$n = 40, r = 4$



$n = 400, r = 8$



$n \rightarrow \infty, r \rightarrow \infty$



Spojité náhodný vektor

Jestliže všechny složky náhodného vektoru (X_1, \dots, X_n) jsou spojité náhodné veličiny, hovoříme o **spojitém náhodném vektoru**. Jeho pravděpodobnostní chování je popsáno **simultánní hustotou pravděpodobnosti** $\varphi(x_1, \dots, x_n)$.

Hustota pravděpodobnosti libovolného spojitého marginálního náhodného vektoru se nazývá marginální hustota pravděpodobnosti. Největší význam mají jednorozměrné marginální hustoty pravděpodobnosti $\varphi_1(x_1), \dots, \varphi_n(x_n)$ jednotlivých složek náhodného vektoru. Získáme je tak, že simultánní hustotu pravděpodobnosti integrujeme přes všech $n-1$ přebývajících proměnných.

Stochasticky nezávislé náhodné veličiny

Při provedení pokusu se může stát, že se realizace jedné náhodné veličiny Y dají jednoznačně určit ze známé realizace druhé náhodné veličiny X , tedy je mezi nimi funkční vztah $Y = g(X)$. Takové náhodné veličiny se nazývají deterministicky závislé.

Jejich protipólem jsou náhodné veličiny stochasticky nezávislé: informace o realizaci jedné z nich nijak nemění šance, s nimiž při témž pokusu očekáváme realizaci druhé.

Např. náhodný pokus spočívá v hodů dvěma kostkami. Náhodná veličina X udává počet ok, která padla na 1. kostce a náhodná veličina Y udává počet ok, která padla na druhé kostce. Náhodné veličiny X, Y jsou stochasticky nezávislé.

Stochastickou nezávislost náhodných veličin zavádíme na základě analogie s četnostní nezávislostí znaků v daném výběrovém souboru, která se používá v popisné statistice. Musí platit multiplikativní vztah:

$$\forall \{x, y\} \in \Omega^2 : p_{\{x, y\}} = p_1(x) \cdot p_2(y) \text{ pro bodové rozložení četností,}$$

$$\forall \{x, y\} \in \Omega^2 : f_{\{x, y\}} = f_1(x) \cdot f_2(y) \text{ pro intervalové rozložení četností.}$$

V počtu pravděpodobnosti nahradíme četnostní funkci pravděpodobnostní funkcí resp. hustotu četnosti nahradíme hustotou pravděpodobnosti. Místo dvou náhodných veličin X, Y můžeme uvažovat n náhodných veličin:

Náhodné veličiny X_1, \dots, X_n jsou stochasticky nezávislé, když platí:

$$\forall \{x_1, \dots, x_n\} \in \Omega^n : \pi_{\{x_1, \dots, x_n\}} = \pi_1(x_1) \cdot \dots \cdot \pi_n(x_n) \text{ v diskrétním případě,}$$

$$\forall \{x_1, \dots, x_n\} \in \Omega^n : \varphi_{\{x_1, \dots, x_n\}} = \varphi_1(x_1) \cdot \dots \cdot \varphi_n(x_n) \text{ ve spojitém případě,}$$

$$\forall \{x_1, \dots, x_n\} \in \Omega^n : \Phi_{\{x_1, \dots, x_n\}} = \Phi_1(x_1) \cdot \dots \cdot \Phi_n(x_n) \text{ v obecném případě.}$$

Vybraná rozložení diskrétních a spojitých náhodných veličin

Motivace

Nyní se seznámíme s přehledem důležitých pravděpodobnostních funkcí a hustot pravděpodobnosti. Uvedeme nejenom analytické vyjádření těchto funkcí, ale též jejich grafy. Vysvětlíme rovněž, v jakých situacích se lze s uvedenými rozloženími pravděpodobností setkat. Zvláštní pozornost budeme věnovat normálnímu rozložení, které hraje velkou roli v celé řadě praktických aplikací počtu pravděpodobnosti a jak uvidíme později, i v matematické statistice.

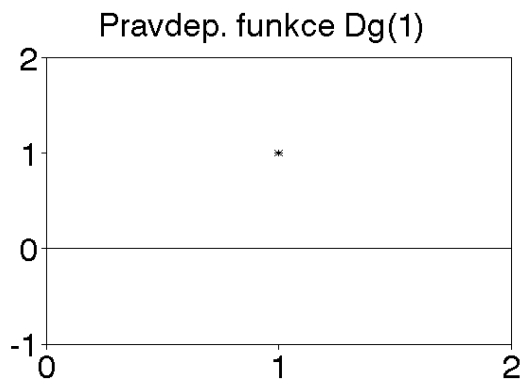
Označení

Známe-li distribuční funkci $\Phi(x)$ náhodné veličiny X (resp. pravděpodobnostní funkci $\pi(x)$ v diskrétním případě resp. hustotu pravděpodobnosti $\varphi(x)$ ve spojitém případě), pak řekneme, že známe rozložení pravděpodobností (zkráceně rozložení) náhodné veličiny X . Toto rozložení závisí na nějakém parametru ϑ , což nejčastěji je reálné číslo nebo reálný vektor.

Zápis $X \sim L(\vartheta)$ čteme: náhodná veličina X má rozložení L s parametrem ϑ .

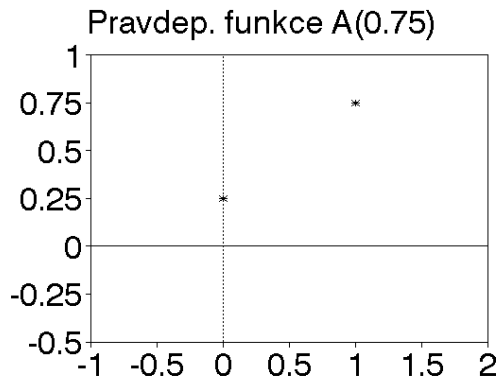
Degenerované rozložení: Náhodná veličina X nabývá pouze konstantní hodnoty μ , píšeme $X \sim \text{Dg}(\mu)$.

$$\pi(x) = \begin{cases} 1 & \text{pro } x = \mu \\ 0 & \text{jinak} \end{cases}$$



Alternativní rozložení: Náhodná veličina X udává počet úspěchů v jednom pokusu, přičemž pravděpodobnost úspěchu je ϑ . Píšeme $X \sim A(\vartheta)$.

$$\pi(x) = \begin{cases} 1 - \vartheta & \text{pro } x=0 \\ \vartheta & \text{pro } x=1 \\ 0 & \text{jinak} \end{cases} \quad \text{neboli } \pi(x) = \begin{cases} \vartheta^x (1 - \vartheta)^{1-x} & \text{pro } x=0,1 \\ 0 & \text{jinak} \end{cases}$$

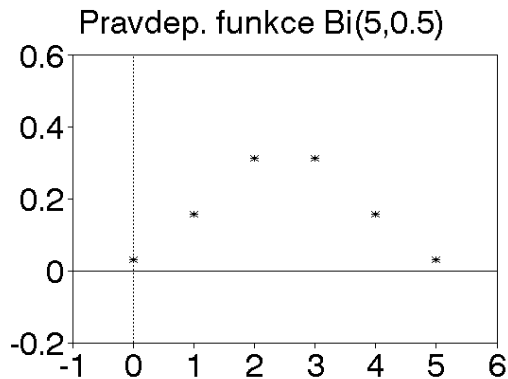


Binomické rozložení: Náhodná veličina X udává počet úspěchů v posloupnosti n nezávislých opakovaných pokusů, přičemž pravděpodobnost úspěchu je v každém pokusu ϑ . Píšeme $X \sim \text{Bi}(n, \vartheta)$.

$$\pi(x) = \begin{cases} \binom{n}{x} \vartheta^x (1-\vartheta)^{n-x} & \text{pro } x = 0, \dots, n \\ 0 & \text{jinak} \end{cases}$$

(Alternativní rozložení je speciálním případem binomického rozložení pro $n = 1$.)

Jsou-li X_1, \dots, X_n stochasticky nezávislé náhodné veličiny, $X_i \sim A(\vartheta)$, $i = 1, \dots, n$, pak $X = \sum_{i=1}^n X_i \sim \text{Bi}(n, \vartheta)$.



Příklad na binomické rozložení pravděpodobností: Firma se účastní čtyř nezávislých výběrových řízení. Pravděpodobnost, že uspěje v kterémkoliv z nich, je pro všechny konkurzy stejná a je rovna 0,7. Jaká je pravděpodobnost, že firma uspěje

- a) právě 2x
- b) aspoň 2x
- c) nejvýše 2x?

Řešení: X ... počet úspěšných konkurzů, $X \sim \text{Bi}(4; 0,7)$

ad a) $P(X = 2) = \binom{4}{2} \cdot 0,7^2 \cdot 0,3^2 = 0,2646$

V systému STATISTICA: =Binom(2;0,7;4)

ad b) $P(X \geq 2) = P(X=2) + P(X=3) + P(X=4) = \binom{4}{2} \cdot 0,7^2 \cdot 0,3^2 + \binom{4}{3} \cdot 0,7^3 \cdot 0,3 + \binom{4}{4} \cdot 0,7^4 = 0,9163$

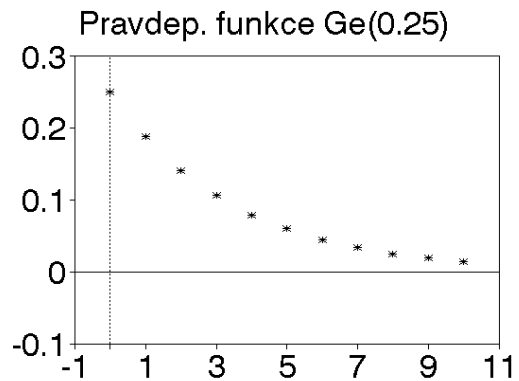
V systému STATISTICA: =1-IBinom(1;0,7;4)

ad c) $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = \binom{4}{0} \cdot 0,3^4 + \binom{4}{1} \cdot 0,7 \cdot 0,3^3 + \binom{4}{2} \cdot 0,7^2 \cdot 0,3^2 = 0,3483$

V systému STATISTICA: =IBinom(2;0,7;4)

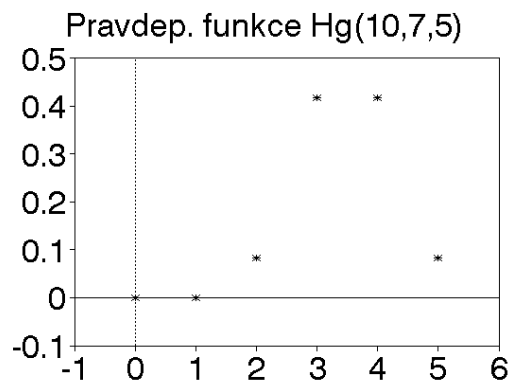
Geometrické rozložení: Náhodná veličina X udává počet neúspěchů v posloupnosti opakovaných nezávislých pokusů předcházejících prvnímu úspěchu, přičemž pravděpodobnost úspěchu je v každém pokusu rovna ϑ . Píšeme $X \sim \text{Ge}(\vartheta)$

$$\pi(x) = \begin{cases} (1 - \vartheta)^x \vartheta & \text{pro } x = 0, 1, \dots \\ 0 & \text{jinak} \end{cases}$$



Hypergeometrické rozložení: V souboru N prvků je M prvků označeno. Náhodně vybereme n prvků bez vracení. Náhodná veličina X udává počet vybraných označených prvků. Píšeme $X \sim \text{Hg}(N, M, n)$

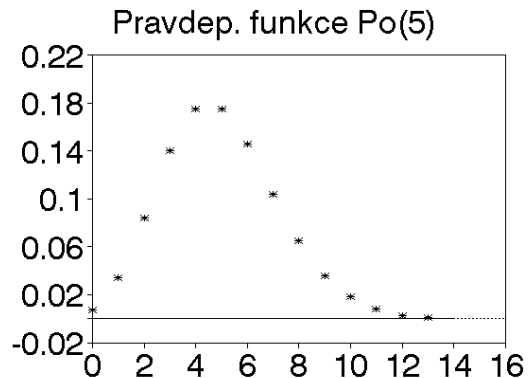
$$\pi(x) = \begin{cases} \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} & \text{pro } x = \max\{0, M - N + n\}, \dots, \min\{M, n\} \\ 0 & \text{jinak} \end{cases}$$



Poissonovo rozložení: Náhodná veličina X udává počet událostí, které nastanou v jednotkovém časovém intervalu (resp. jednotkové oblasti), přičemž události nastávají náhodně, jednotlivě a vzájemně nezávisle. Parametr $\lambda > 0$ je střední počet těchto událostí. Píšeme $X \sim \text{Po}(\lambda)$.

(Poissonovým rozložením se řídí např. počet výzev, které dojdou na telefonní ústřednu během určitého časového intervalu nebo počet mikroorganismů v zorném poli mikroskopu. Jde o tzv. řídce se vyskytující jevy.)

$$\pi(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{pro } x = 0, 1, \dots \\ 0 & \text{jinak} \end{cases}$$



Vztah mezi pravděpodobnostní funkcí binomického a Poissonova rozložení:

Nechť náhodná veličina $X \sim \text{Po}(\lambda)$ a náhodná veličina $Y \sim \text{Bi}(n, \vartheta)$. Nechť $\vartheta \rightarrow 0$ pro $n \rightarrow \infty$ a přitom $n\vartheta \rightarrow \lambda$. Pak pravděpodobnostní funkce náhodné veličiny Y konverguje k pravděpodobnostní funkci náhodné veličiny X , tj.

$$\lim_{n \rightarrow \infty} \binom{n}{y} \vartheta^y (1-\vartheta)^{n-y} = \frac{\lambda^y}{y!} e^{-\lambda}$$

(Aproximace binomického rozložení pomocí Poissonova rozložení je vyhovující, když $n > 30$ a $\vartheta < 0,1$.)

Příklad na Poissonovo rozložení: Dělnice v prádelně obsluhuje 800 vřeten. Pravděpodobnost toho, že se příze přetrhne během časového intervalu délky t , je pro všechna vřetena stejná a je rovna 0,005. Určete pravděpodobnost, že během intervalu délky t dojde k nejvýše 10 přetržením.

Řešení: Y – počet přetržení v časovém intervalu délky t , $Y \sim \text{Bi}(800; 0,005)$.

Přesný výpočet: $P(Y \leq 10) = \sum_{y=0}^{10} \binom{800}{y} 0,005^y (1-0,005)^{800-y} = 0,997239$

Ve STATISTICE: výpočet pomocí funkce $\text{IBinom}(10; 0,005; 800)$.

Aproximativní výpočet: podmínky dobré aproximace jsou splněny, parametr

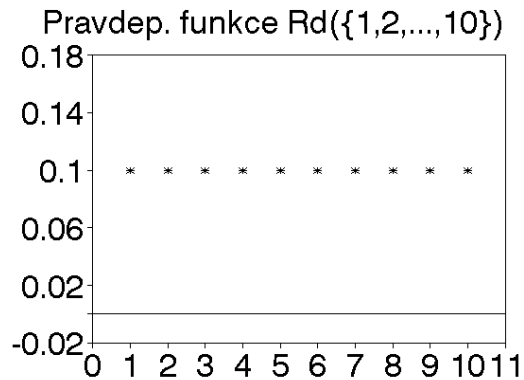
$$\lambda = n\vartheta = 800 \cdot 0,005 = 4, \quad P(Y \leq 10) = \sum_{y=0}^{10} \frac{4^y}{y!} \cdot e^{-4} = 0,9971602$$

Ve STATISTICE: výpočet pomocí funkce $\text{IPoisson}(10; 4)$.

Rovnoměrné diskrétní rozložení: Necht' G je konečná množina o n prvcích. Náhodná veličina X nabývá se stejnou pravděpodobností každé hodnoty z množiny G . Píšeme $X \sim \text{Rd}(G)$

$$\pi(x) = \begin{cases} \frac{1}{n} & \text{pro } x \in G \\ 0 & \text{jinak} \end{cases}$$

(Typickým příkladem je náhodná veličina udávající počet ok při hodu kostkou.)



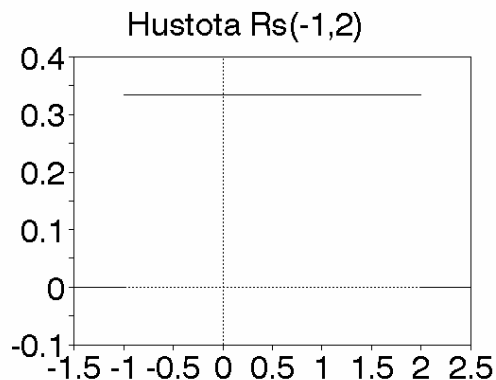
Rovnoměrné spojité rozložení: Předpokládejme, že veličina X

- může nabýt jakékoliv hodnoty mezi čísly a , b
- pravděpodobnost, že nabude hodnoty z jakéhokoliv intervalu v tomto rozmezí je stejná jako pravděpodobnost, že nabude hodnoty z jakéhokoliv jiného intervalu stejné délky.

Jsou-li tyto podmínky splněny, pak X má rovnoměrné spojité rozložení na intervalu (a, b) . Hustota pravděpodobnosti náhodné veličiny X je konstantní na intervalu (a, b) a plocha pod křivkou hustoty tvoří obdélník.

Píšeme $X \sim R_s(a, b)$.

$$\varphi(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in (a, b) \\ 0 & \text{jinak} \end{cases}$$

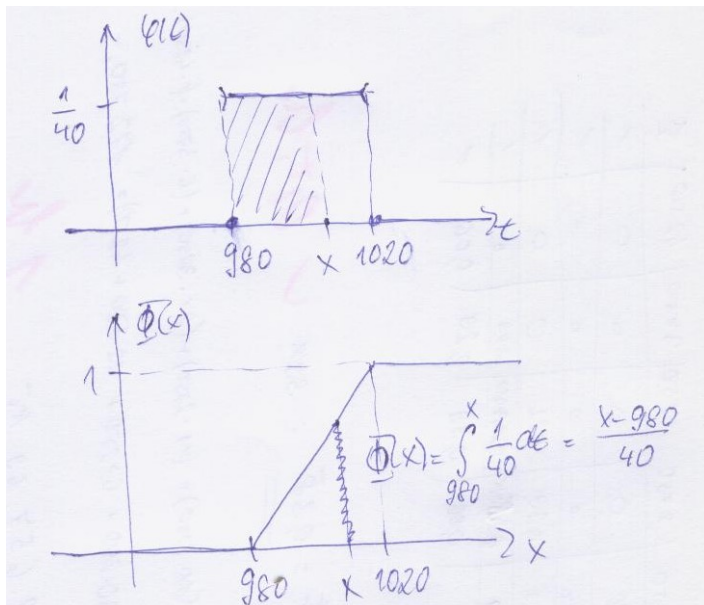


Příklad na rovnoněrné spojité rozložení: Na automatické lince se plní láhve mlékem. Působením náhodných vlivů množství mléka kolísá v intervalu (980 ml, 1020 ml). Každé množství mléka v tomto intervalu považujeme za stejně možné. Jaká je pravděpodobnost, že v náhodně vybrané láhvi bude aspoň 1000 ml mléka?

Řešení:

X – množství mléka v náhodně vybrané láhvi, $X \sim R_s(980, 1020)$,

$$f(x) = \begin{cases} \frac{1}{40} & \text{pro } x \in (980, 1020) \\ 0 & \text{jinak} \end{cases}, \quad P(X \geq 1000) = \int_{1000}^{1020} \frac{1}{40} dx = \frac{1}{40} \left[x \right]_{1000}^{1020} = \frac{20}{40} = 0,5$$

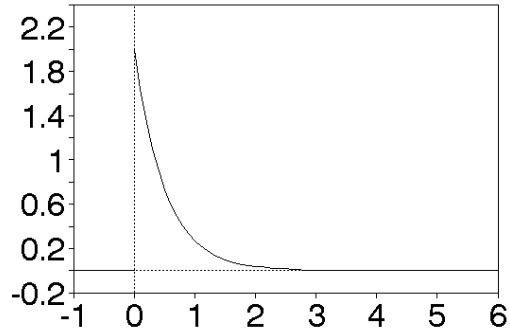


Exponenciální rozložení: Náhodná veličina X udává dobu čekání na příchod nějaké události, která se může dostavit každým okamžikem se stejnou šancí bez ohledu na dosud pročekanou dobu. (Jde o tzv. čekání bez paměti.) Přitom $\frac{1}{\lambda}$

vyjadřuje střední dobu čekání. Píšeme $X \sim \text{Ex}(\lambda)$

$$\varphi(x) = \begin{cases} \lambda e^{-\lambda x} & \text{pro } x > 0 \\ 0 & \text{jinak} \end{cases}, \quad \Phi(x) = \begin{cases} 1 - e^{-\lambda x} & \text{pro } x > 0 \\ 0 & \text{jinak} \end{cases}$$

Hustota $\text{Ex}(2)$



Příklad na exponenciální rozložení: Doba (v minutách) potřebná k obslužení zákazníka v prodejně potravin je náhodná veličina, která se řídí rozložením $\text{Ex}\left(\frac{1}{3}\right)$. Jaká je pravděpodobnost, že doba potřebná k obslužení náhodně vybraného zákazníka v této prodejně bude v rozmezí od 3 do 6 minut?

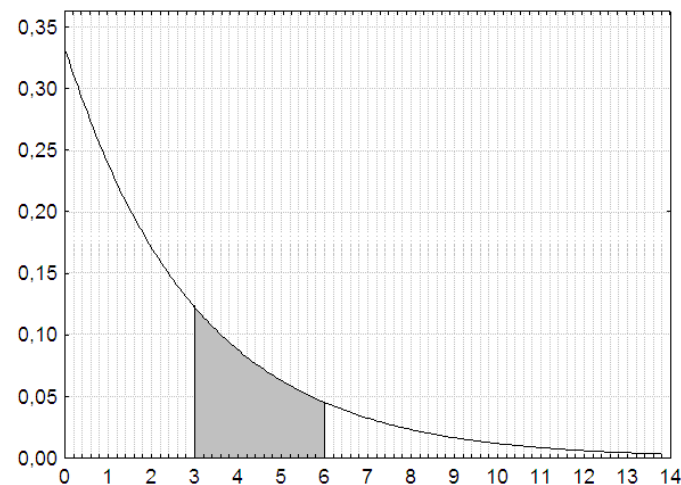
Řešení:

X – doba potřebná k obslužení náhodně vybraného zákazníka, $X \sim \text{Ex}\left(\frac{1}{3}\right)$,

$$P(3 \leq X \leq 6) = \int_3^6 \frac{1}{3} e^{-\frac{x}{3}} dx = \frac{1}{3} \left[-e^{-\frac{x}{3}} \right]_3^6 = -e^{-2} + e^{-1} = 0,233.$$

S pravděpodobností 0,233 bude zákazník obslužen v době od 3 do 6 minut.

(V systému STATISTICA slouží k výpočtu hodnoty distribuční funkce $F(x)$ rozložení $\text{Ex}(\lambda)$ funkce $\text{IExpon}(x;\lambda)$. V našem příkladě tedy do Dlouhého jména proměnné napíšeme $=\text{IExpon}(6;1/3)-\text{IExpon}(3;1/3)$. Dostaneme výsledek 0,2325.)

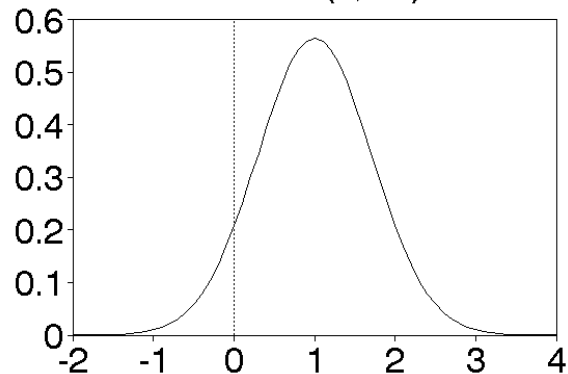


Normální rozložení: Tato náhodná veličina vzniká např. tak, že ke konstantě μ se přičítá velké množství nezávislých náhodných vlivů mírně kolísajících kolem nuly. Proměnlivost těchto vlivů je vyjádřena konstantou $\sigma > 0$.

Píšeme $X \sim N(\mu, \sigma^2)$, hustota $\varphi(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x-\mu}{2\sigma^2}}$. Grafem této hustoty je tzv.

Gaussova křivka.

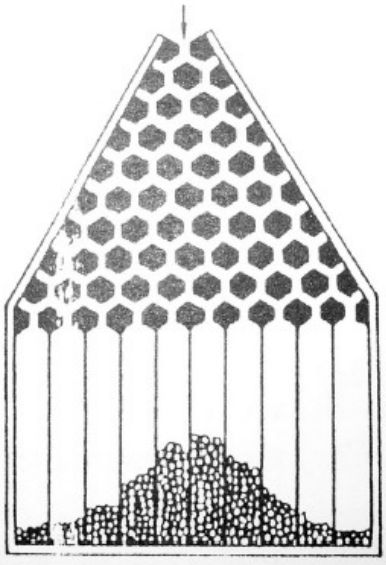
Hustota $N(1,0.5)$



Ilustrace vzniku normálního rozložení pomocí Galtonovy desky:

Deska obsahuje n řad pravidelně uspořádaných klínů, a to tak, že v k -té řadě je právě k klínů. Do otvoru nahoře padají kuličky, které jsou v každé řadě se stejnou pravděpodobností $1/2$ vychylovány vlevo nebo vpravo. Pod poslední radou je $n - 1$ přihrádek, ve kterých se kuličky shromažďují. Nasypeme-li do tohoto systému velké množství kuliček, vytvoří v přihrádkách jakýsi "kopec", jehož tvar je velmi podobný tvaru grafu hustoty náhodné veličiny s normálním rozložením. Náhodné vychylování kuliček jednotlivými řadami překážek je možno chápat jako speciální případ velkého množství chybových faktorů, náhodně působících na nějaký proces, jako působení mnoha blíže nespecifikovatelných vlivů, které ovlivňují zcela náhodně rozložení jeho výsledku.

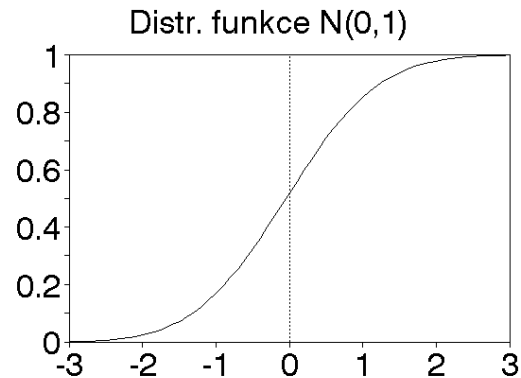
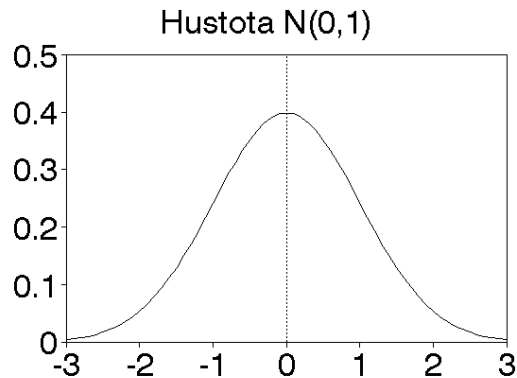
Obrázek



Standardizované normální rozložení:

Pro $\mu = 0$, $\sigma^2 = 1$ se jedná o standardizované normální rozložení, píšeme

$U \sim N(0, 1)$. Hustota pravděpodobnosti má v tomto případě tvar $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$.



$\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$ je tabelována pro $u \geq 0$, pro $u < 0$ se používá přepočtový vzorec $\Phi(-u) = 1 - \Phi(u)$.

Příklad na normální rozložení: Výsledky u přijímacích zkoušek na jistou VŠ jsou normálně rozloženy s parametry $\mu = 550$ bodů, $\sigma = 100$ bodů. S jakou pravděpodobností bude mít náhodně vybraný uchazeč aspoň 600 bodů?

Řešení:

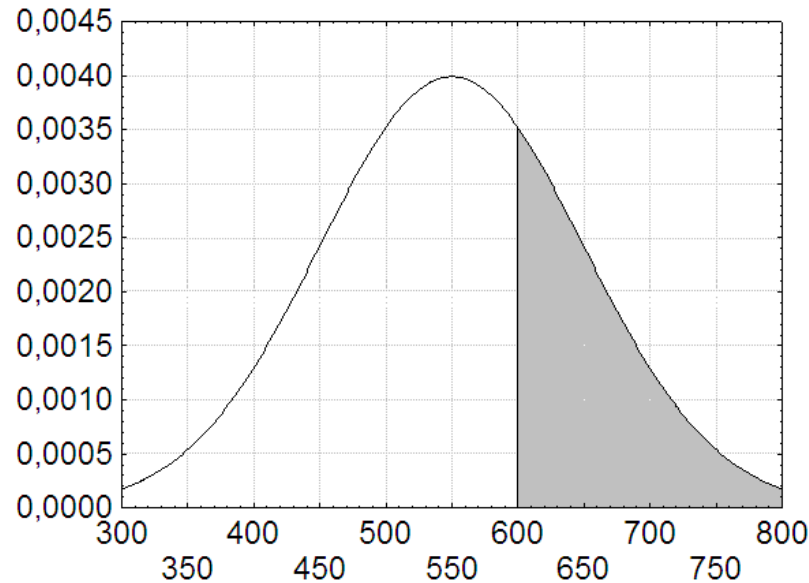
X – výsledek náhodně vybraného uchazeče, $X \sim N(550, 100^2)$,

$$P(X \geq 600) = 1 - P(X \leq 600) + P(X = 600) = 1 - P(X \leq 600) =$$

$$= 1 - P\left\{\frac{X - \mu}{\sigma} \leq \frac{600 - \mu}{\sigma}\right\} = 1 - P\left\{U \leq \frac{600 - 550}{100}\right\} = 1 - \Phi(0,5) = 1 - 0,69146 = 0,30854.$$

Ve STATISTICE: výpočet pomocí funkce $1 - \text{INormal}(600;550;100)$

Náhodně vybraný uchazeč bude mít u zkoušek aspoň 600 bodů s pravděpodobností 0,31.



Některé vlastnosti normálního rozložení:

Jestliže $X \sim N(\mu, \sigma^2)$, pak $U = \frac{X - \mu}{\sigma} \sim N(0, 1)$.

Jestliže $X \sim N(\mu, \sigma^2)$, a $Y = a + bX$, pak $Y \sim N(a + b\mu, b^2\sigma^2)$.

Jestliže X_1, \dots, X_n jsou stochasticky nezávislé náhodné veličiny, $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \dots, n$, $Y = \sum_{i=1}^n X_i$, pak $Y \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$.

Význam normálního rozložení:

Normální rozložení hraje ústřední roli v počtu pravděpodobnosti i matematické statistice. Jeho význam spočívá jednak v tom, že normálním rozložením se řídí pravděpodobnostní chování mnoha náhodných veličin a jednak v tom, že za určitých podmínek konverguje k normálnímu rozložení součet nezávislých náhodných veličin s tímž rozložením (viz centrální limitní věta).

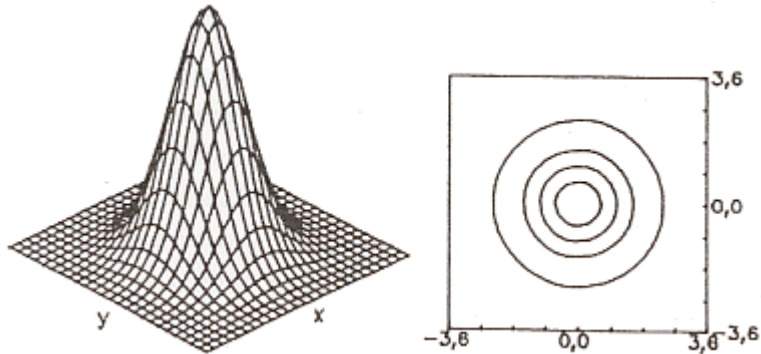
Dvourozměrné normální rozložení: $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right)$

Náhodný vektor $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ vzniká ve dvourozměrných situacích podobně jako skalární náhodná veličina v bodě (e).

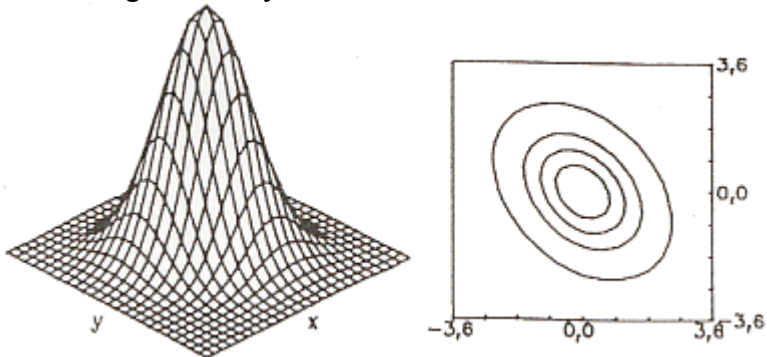
$$\varphi(x_1, x_2) = \frac{1}{\sigma_1 \sigma_2 \sqrt{1 - \rho^2}} e^{-\frac{q(x_1, x_2)}{2}}, \text{ kde } q(x_1, x_2) = \frac{1}{1 - \rho^2} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \frac{x_1 - \mu_1}{\sigma_1} \cdot \frac{x_2 - \mu_2}{\sigma_2} + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right].$$

Pro $\mu_1 = 0, \mu_2 = 0, \sigma_1^2 = 1, \sigma_2^2 = 1, \rho = 0$ se jedná o **standardizované dvourozměrné normální rozložení**.

Vrstevnice a graf hustoty standardizovaného dvourozměrného normálního rozložení

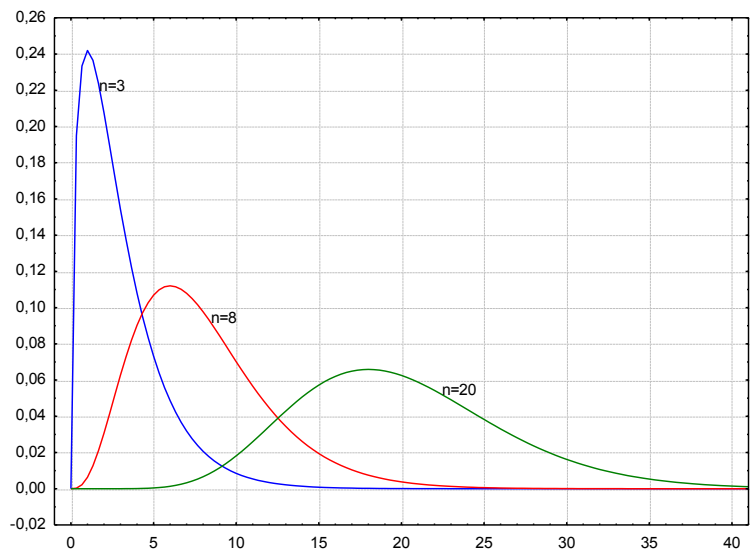


Vrstevnice a graf hustoty dvourozměrného normálního rozložení s parametry $\mu_1 = 0, \mu_2 = 0, \sigma_1^2 = 1, \sigma_2^2 = 1, \rho = -0,75$



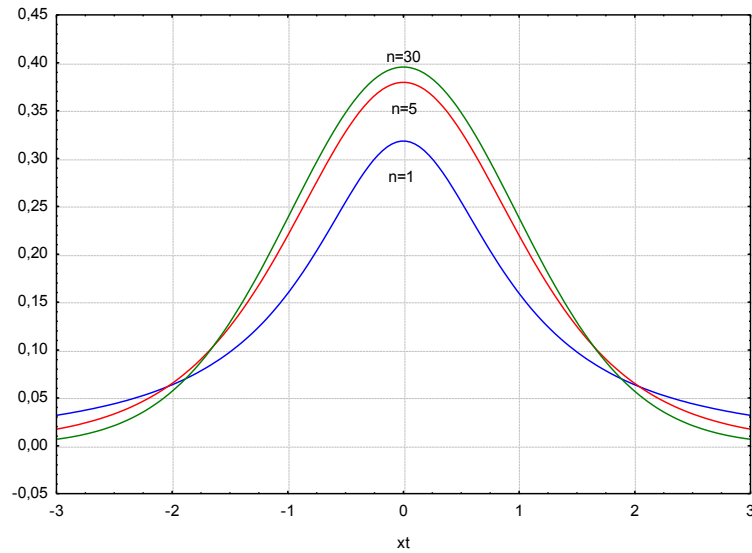
Upozornění: Následující tři rozložení – Pearsonovo, Studentovo a Fisherovo-Snedecorovo – jsou odvozena ze standardizovaného normálního rozložení. Mají velký význam především v matematické statistice při konstrukci intervalů spolehlivosti a testování hypotéz. Vyjádření hustot těchto rozložení neuvádíme, je příliš složité

Pearsonovo rozložení chí-kvadrát s n stupni volnosti: Necht' X_1, \dots, X_n jsou stochasticky nezávislé náhodné veličiny, $X_i \sim N(0, 1)$, $i = 1, \dots, n$. Pak náhodná veličina $X = X_1^2 + \dots + X_n^2 \sim \chi^2(n)$.



Studentovo rozložení s n stupni volnosti: Necht' X_1, X_2 jsou stochasticky nezávislé náhodné veličiny, $X_1 \sim N(0, 1)$, $X_2 \sim \chi^2(n)$.

Pak náhodná veličina $X = \frac{X_1}{\sqrt{\frac{X_2}{n}}} \sim t(n)$.



Fisherovo-Snedecorovo rozložení s n_1 a n_2 stupni volnosti: Necht' X_1, \dots, X_n jsou stochasticky nezávislé náhodné veličiny, $X_i \sim \chi^2(n_i)$, $i = 1, 2$.

Pak náhodná veličina $X = \frac{X_1/n_1}{X_2/n_2} \sim F(n_1, n_2)$.

