Stochasticky nezávislé náhodné veličiny - motivace

10.1. Motivace: 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 hodu 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:

$\forall(x,y)\in R^2 : p(x,y) = p_1(x)p_2(y)$ pro bodové rozložení četností,

$\forall(x,y)\in R^2 : f(x,y) = f_1(x)f_2(y)$ 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 $X_1, \ldots, X_n$ jsou stochasticky nezávislé, když platí:

$\forall(x_1,\ldots,x_n)\in R^n : \pi(x_1,\ldots,x_n) = \pi_1(x_1)\cdot\ldots\cdot\pi_n(x_n)$ v diskrétním případě,

$\forall(x_1,\ldots,x_n)\in R^n : \varphi(x_1,\ldots,x_n) = \varphi_1(x_1)\cdot\ldots\cdot\varphi_n(x_n)$ ve spojitém případě,

$\forall(x_1,\ldots,x_n)\in R^n : \Phi(x_1,\ldots,x_n) = \Phi_1(x_1)\cdot\ldots\cdot\Phi_n(x_n)$ v obecném případě.

Stochasticky nezávislé náhodné veličiny

10.2. Definice:

a) Obecný případ: Řekneme, že náhodné veličiny $X_1, \ldots, X_n$ s marginálními distribučními funkcemi $\Phi_1(x_1), \ldots, \Phi_n(x_n)$ a simultánní distribuční funkcí $\Phi(x_1, \ldots, x_n)$ jsou **stochasticky nezávislé**, právě když

$\forall (x_1, \ldots, x_n) \in R^n : \Phi(x_1, \ldots, x_n) = \Phi_1(x_1) \ldots \Phi_n(x_n).$

b) Diskrétní případ: Řekneme, že diskrétní náhodné veličiny $X_1, \ldots, X_n$ s marginálními pravděpodobnostními funkcemi $\pi_1(x_1), \ldots, \pi_n(x_n)$ a simultánní pravděpodobnostní funkcí $\pi(x_1, \ldots, x_n)$ jsou **stochasticky nezávislé**, právě když

$\forall (x_1, \ldots, x_n) \in R^n : \pi(x_1, \ldots, x_n) = \pi_1(x_1) \ldots \pi_n(x_n).$

c) Spojitý případ: Řekneme, že spojité náhodné veličiny $X_1, \ldots, X_n$ s marginálními hustotami $\varphi_1(x_1), \ldots, \varphi_n(x_n)$ a simultánní hustotou $\varphi(x_1, \ldots, x_n)$ jsou **stochasticky nezávislé**, právě když

$\forall (x_1, \ldots, x_n) \in R^n : \varphi(x_1, \ldots, x_n) = \varphi_1(x_1) \ldots \varphi_n(x_n)$ s případnou výjimkou na množině bodů neovlivňujících integraci.

10.3. Definice:

Řekneme, že posloupnost $\{X_n\}_{n=1}^{\infty}$ je **posloupnost stochasticky nezávislých náhodných veličin**, právě když pro všechna přirozená n jsou stochasticky nezávislé náhodné veličiny $X_1, \ldots, X_n$.

Příklad

10.4. Příklad: Diskrétní náhodný vektor (X_1, X_2) má simultánní pravděpodobnostní funkci $\pi(x_1, x_2)$ danou hodnotami: $\pi(0,0) = \pi(0,2) = \pi(1,1) = \pi(2,0) = \pi(2,2) = 0$, $\pi(0,1) = \pi(1,0) = \pi(1,2) = \pi(2,1) = 0,25$. Jsou náhodné veličiny X_1, X_2 stochasticky nezávislé?

Řešení:

Sestavíme kontingenční tabulku, v níž budou hodnoty simultánní pravděpodobnostní funkce a obou marginálních pravděpodobnostních funkcí.

x_1	x_2			$\pi_1(x_1)$
	0	1	2	
0	0	0,25	0	0,25
1	0,25	0	0,25	0,5
2	0	0,25	0	0,25
$\pi_2(x_2)$	0,25	0,5	0,25	1

Ověříme splnění multiplikativního vztahu $\forall (x_1, x_2) \in R^2$: $\pi(x_1, x_2) = \pi_1(x_1)\, \pi_2(x_2)$. Již pro $x_1 = 0$, $x_2 = 0$ vztah splněn není, protože $\pi(0,0) = 0$, avšak $\pi_1(0) = 0,25$ a $\pi_2(0) = 0,25$. Veličiny X_1, X_2 tedy nejsou stochasticky nezávislé.

Příklad

10.5. Příklad:

Nechť spojitý vektor (X_1, X_2) má simultánní hustotu pravděpodobnosti

$\varphi(x_1, x_2) = \begin{cases} 24\,x_1^2 x_2 (1 - x_1) \text{ pro } 0 \le x_1 < 1, 0 \le x_2 < 1 \\ 0 \text{ jinak} \end{cases}$. Dokažte, že náhodné veličiny X_1, X_2 jsou stochasticky nezávislé.

Řešení:

Vypočítáme obě marginální hustoty a ověříme platnost multiplikativního vztahu

$\forall (x_1, ..., x_n) \in R^n$: $\varphi(x_1, ..., x_n) = \varphi_1(x_1) \dots \varphi_n(x_n)$ s případnou výjimkou na množině bodů neovlivňujících integraci.

$$\varphi_1(x_1) = \int_0^1 24x_1^2 x_2 (1 - x_1)\, dx_2 = 24x_1^2 (1-x_1) \left[\frac{x_2^2}{2}\right]_0^1 = 12x_1^2 (1-x_1) \text{ pro } 0 \le x_1 < 1,$$

$\varphi_1(x_1) = 0$ jinak.

$$\varphi_2(x_2) = \int_0^1 24x_1^2 x_2 (1 - x_1)\, dx_1 = 24x_2 \left[\frac{x_1^3}{3} - \frac{x_1^4}{4}\right]_0^1 = 2x_2 \text{ pro } 0 \le x_2 < 1,$$

$\varphi_2(x_2) = 0$ jinak.

Vidíme, že multiplikativní vztah je splněn, tudíž veličiny X_1, X_2 jsou stochasticky nezávislé.

Stochasticky nezávislé náhodné vektory

10.6. Definice:

Nechť $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor, $\mathbf{X}_1 = (X_{11}, \dots, X_{p_1 1}), \dots, \mathbf{X}_n = (X_{1n}, \dots, X_{p_n n})$ náhodné vektory definované na $(\Omega, \mathcal{A}, P)$. Řekneme, že tyto náhodné vektory jsou **stochasticky nezávislé**, právě když každá složka náhodného vektoru $\mathbf{X}_i$ je stochasticky nezávislá se všemi složkami náhodného vektoru $\mathbf{X}_k$ pro $\forall i \neq k$.

10.7. Věta:

Nechť $X_1, \dots, X_n$ jsou stochasticky nezávislé náhodné veličiny, $g_1, \dots, g_n$ borelovské funkce. Pak transformované náhodné veličiny $Y_1 = g_1(X_1), \dots, Y_n = g_n(X_n)$ jsou opět stochasticky nezávislé náhodné veličiny.

(Tvrzení lze zobecnit i pro transformované náhodné vektory.)

Příklad

10.8. Příklad:

Nechť $X_1,\dots,X_n$ jsou stochasticky nezávislé náhodné veličiny s distribučními funkcemi $\Phi_1(x_1),\dots,\Phi_n(x_n)$. Zavedeme transformované náhodné veličiny $Y=\max\{X_1,\dots,X_n\}$, $Z=\min\{X_1,\dots,X_n\}$. Odvoďte jejich distribuční funkce $\Phi_{\max}(y),\Phi_{\min}(z)$.

Řešení:

$$\Phi_{\max}(y)=P(Y\le y)=P(\max\{X_1,\dots,X_n\}\le y)=P(X_1\le y\wedge\dots\wedge X_n\le y)=P(X_1\le y)\cdot\ldots\cdot P(X_n\le y)=\Phi_1(y)\cdot\ldots\cdot\Phi_n(y)$$

$$\Phi_{\min}(z)=P(Z\le z)=P(\min\{X_1,\dots,X_n\}\le z)=P(X_1\le z\vee\dots\vee X_n\le z)=1-P(X_1>z\wedge\dots\wedge X_n>z)=1-P(X_1>z)\cdot\ldots\cdot P(X_n>z)=$$
$$=1-[1-P(X_1\le z)]\cdot\ldots\cdot[1-P(X_n\le z)]=1-[1-\Phi_1(z)]\cdot\ldots\cdot[1-\Phi_n(z)]$$

Příklad

10.9. Příklad:

Na automatické lince jsou láhve plněny mlékem. Je známo, že množství mléka v láhvích kolísá od od 0,98 l do 1,02 l. V tomto intervalu považujeme každé množství mléka za stejně možné. Za 1 s se naplní 3 láhve. Jaká je pravděpodobnost, že

a) nejméně naplněná láhev obsahuje aspoň 1 l mléka,

b) v nejvíce naplněné láhvi není víc než 1,01 l mléka?

Řešení:

Náhodná veličina X_i udává množství mléka v i-té láhvi, i = 1, 2, 3. Je to spojitá náhodná veličina, její hustota pravděpodobnosti je konstantní na intervalu (0,98; 1,02). Z podmínky normovanosti S2 dostaneme, že hustota

$$\varphi(x)=\begin{cases}\frac{1}{40} & \text{pro } x\in(980,1020)\\ 0 & \text{jinak}\end{cases}. \quad \text{Distribuční funkce: } \Phi(x)=\begin{cases}0 & \text{pro } x\in(-\infty,980\rangle\\ \int_{980}^{x}\frac{1}{40}\,dt=\frac{1}{40}[x]_{980}^{x}=\frac{x-980}{40} & \text{pro } x\in(980,1020)\\ 1 & \text{pro } x\in\langle 1020,\infty)\end{cases}$$

ad a) $P(Z\geq 1000)=1-P(Z\leq 1000)=1-\Phi_{min}(1000)=[1-\Phi(1000)]^3=\left[1-\frac{1000-980}{40}\right]^3=\left(\frac{1}{2}\right)^3=\frac{1}{8}=0{,}125$

ad b) $P(Y\leq 1010)=\Phi_{max}(1010)=[\Phi(1010)]^3=\left(\frac{3}{4}\right)^3=\frac{27}{64}=0{,}42$

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 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 je nejčastěji reálné číslo nebo reálný vektor.

Zápis $\boxed{X \sim L(\vartheta)}$ čteme: <u>náhodná veličina X má rozložení L s parametrem ϑ</u>.

<u>Na webu:</u>

<u>http://en.wikipedia.org/wiki/List_of_probability_distributions</u>

Vybraná rozložení diskrétních náhodných veličin

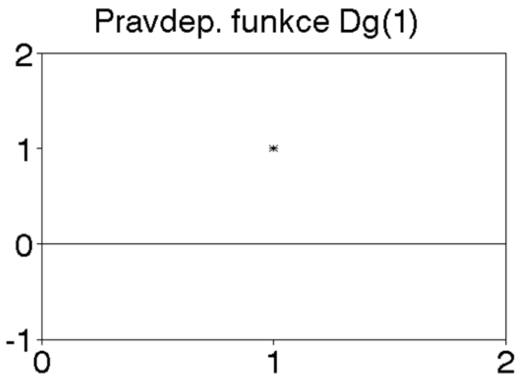
Důležitá diskrétní rozdělení:

- Degenerované rozložení
- Alternativní (Bernoulliho) rozdělení
- Binomické rozdělení
- Multinomické rozdělení
- Poissonovo rozdělení
- Negativně binomické (Pascalovo) rozdělení
- Geometrické rozdělení (zvláštní případ negativně binomického rozdělení)
- Hypergeometrické rozdělení
- Rovnoměrné rozdělení

Degenerované rozložení

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

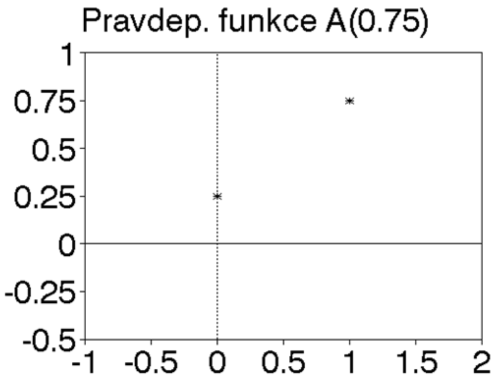
$$\pi(x) = \begin{cases} 1 \text{ pro } x = \mu \\ 0 \text{ jinak} \end{cases}$$



Alternativní rozložení

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

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



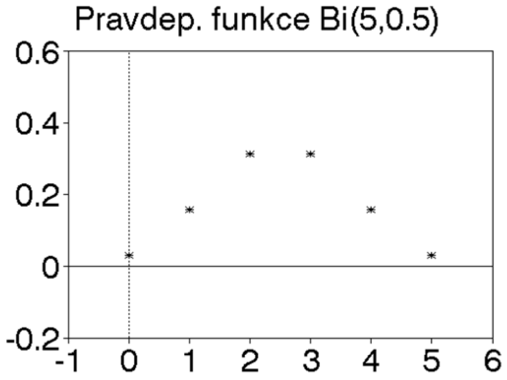
Binomické rozložení

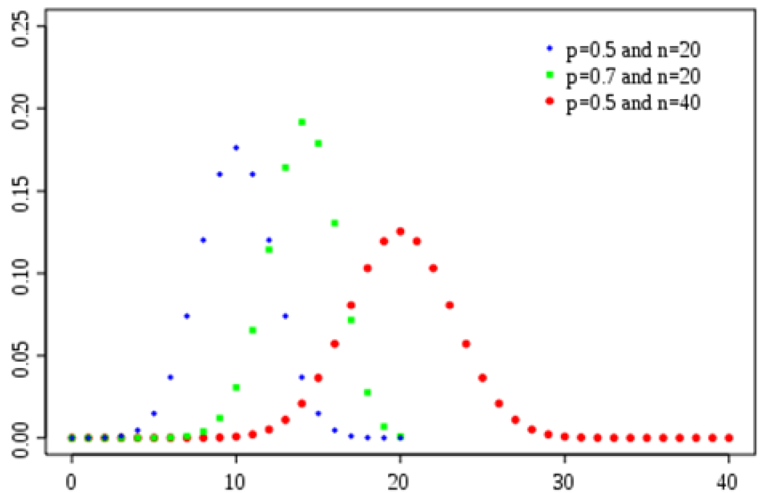
Binomické rozložení: Náhodná veličina X udává počet úspěchů v posloupnosti n nezávislých opakovaných pokusů, přičemž pravděpodobnost úspěchu je v každém pokusu ϑ. Píšeme $X \sim Bi(n, \vartheta)$.

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

(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 $X = \sum_{i=1}^{n} X_i \sim Bi(n, \vartheta)$.)





Příklad

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 ~ Bi(4; 0,7)

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

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

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

Příklad

Pravděpodobnost narození chlapce je 0,515. Určete takový počet dětí, aby pravděpodobnost, že mezi nimi bude aspoň jeden chlapec, byla vetší než 0,99.

Řešení:

Označme jako X veličinu udávající počet chlapců mezi n dětmi, je X~Bi(n,0,515). Hledáme takové n, aby P(X > 0) > 0,99, přitom platí

$$P(X>0)=1-P(X\leq 0)=1-P(X=0)=1-\binom{n}{0}\cdot 0{,}515^{0}\cdot(1-0{,}515)^{n-0}$$

$\Rightarrow$ $1-(0{,}485)^{n}>0{,}99$

$\Rightarrow$ $n>\dfrac{\ln 0{,}01}{\ln 0{,}485}\cong 6{,}36$

$\Rightarrow$ $n\geq 7$

Multinomické rozložení

Multinomické rozložení: Zobecnění binomického rozložení. Složky náhodný vektoru $(X_1, \ldots X_k)$ udávají počty úspěchů (nastane jev $A_1, \ldots A_k$) v posloupnosti *n* nezávislých opakovaných pokusů, přičemž pravděpodobnosti úspěchů jsou $\vartheta_1, \ldots, \vartheta_k$. Předpokládáme, že při každém pokusu nastane právě jeden z jevů $A_1, \ldots A_k$, přičemž platí $\vartheta_1 + \ldots + \vartheta_k = 1$. Píšeme $X \sim Mu(n, \vartheta_1, \ldots, \vartheta_k)$

$$\pi(x_1, \ldots, x_k) = \frac{n!}{x_1! \ldots x_k!} \vartheta_1^{x_1} \cdot \ldots \cdot \vartheta_k^{x_k}, \quad x_1, \ldots, x_k \in \{1, \ldots, n\}, \sum_{i=1}^{k} x_i = n$$

$$= 0 \qquad\qquad jinak$$

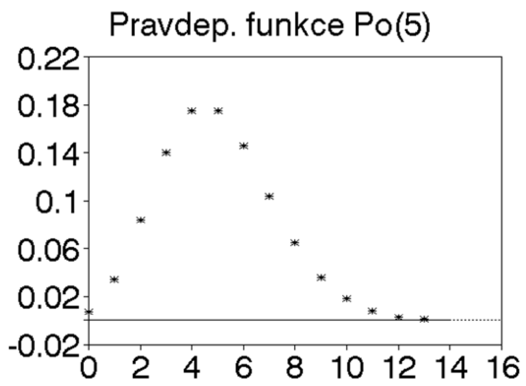
Platí: $X_j \sim Bi(n, \vartheta_j)$

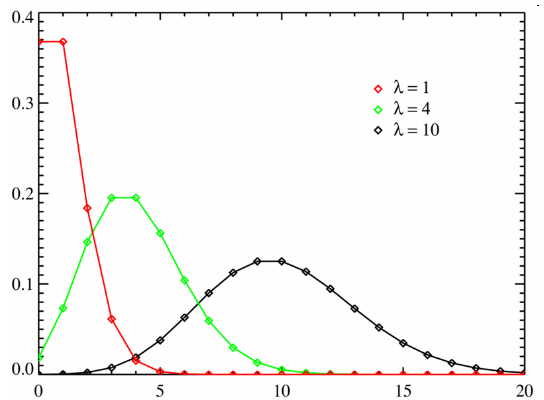
Multinomické rozložení – příklady využití

- Předvolební průzkum:
  - n – počet tázaných
  - ϑ_j – skutečný podíl voličů j-té strany v populaci
  - X_j – počet (četnost) voličů j-té strany ve výběru
- Hody hrací kostkou:
  - n – počet hodů
  - $\vartheta_1,\dots,\vartheta_6$ – pravděpodobnost jednotlivých stran kostky
  - $X_1,\dots X_6$ – absolutní četnosti jednotlivých stran kostky
- Krevní skupiny:
  - n=4 (skupiny 0,A,B,AB)
  - $\vartheta_0, \vartheta_A, \vartheta_B, \vartheta_{AB}$ – pravděpodobnosti skupin 0, A, B, AB
  - X_0, X_A, X_B, X_{AB} – počty osob se skupinami 0, A, B, AB

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 $X \sim Po(\lambda)$.
(Poissonovým rozložením se řídí např. počet výzev, které dojdou na telefonní ústřednu během určitého časového intervalu nebo počet mikroorganizmů v zorném poli mikroskopu. Jde o tzv. řídce se vyskytující jevy.)

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





Příklad

Vztah mezi pravděpodobnostní funkcí binomického a Poissonova rozložení:
Nechť náhodná veličina $X \sim Po(\lambda)$ a náhodná veličina $Y \sim Bi(n, \vartheta_n)$. Nechť $\vartheta_n \to 0$ pro $n \to \infty$ a přitom $n\vartheta_n \to \lambda$. Pak pravděpodobnostní funkce náhodné veličiny Y konverguje k pravděpodobnostní funkci náhodné veličiny X, tj.

$$\lim_{n\to\infty}\binom{n}{y}\vartheta_n^{\,y}(1-\vartheta_n)^{n-y} = \frac{\lambda^y}{y!}e^{-\lambda}$$

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

Příklad na Poissonovo rozložení: Dělnice v přádelně 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, $Y \sim Bi(800;0{,}005)$.

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

Aproximativní výpočet: podmínky dobré aproximace jsou splněny, parametr

$\lambda = n\vartheta = 800.0{,}005 = 4,\ P(Y \le 10) = \sum_{y=0}^{10}\frac{4^y}{y!}e^{-4} = 0{,}9971602$

Příklad

1) Průměrný telefonní hovor trvá 1,5 min. Má-li ústředna 10 linek a dochází-li průměrně k 120 hovorům za hodinu, jaká je pravděpodobnost ztráty volání?

Řešení: X udává počet volajících, X~Po(2*1,5). Ke ztrátě volání dojde, pokud chce současně volat více než 10 volajících (tj. není volná linka). Tedy

$$P(X > 10) = 1 - P(X \leq 10) = 1 - \sum_{x=0}^{10} \frac{3^x}{x!} e^{-3} \cong 0{,}001.$$

2) Průměrný telefonní hovor trvá 1,5 min. Kolik linek musí ústředna mít, dochází-li průměrně k 240 hovorům za hodinu a pravděpodobnost ztráty volání nemá překročit a) 0,01, b) 0,001?

Řešení: X udává počet volajících, X~Po(240/60*1,5). Hledáme n tak aby $P(X > n) \leq 0{,}01$

tj. $P(X \leq n) \geq 0{,}99 \Rightarrow \sum_{x=0}^{n} \frac{6^x}{x!} e^{-6} \geq 0{,}99 \Rightarrow n = 12.$

Pro případ b) chceme $\sum_{x=0}^{n} \frac{6^x}{x!} e^{-6} \geq 0{,}999 \Rightarrow n = 15.$

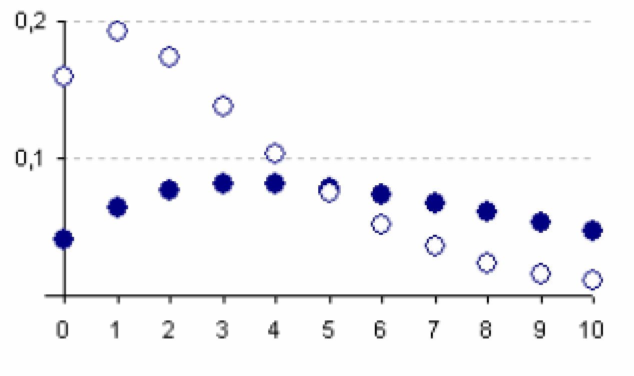
Negativní binomické rozložení (Pascalovo):

Náhodná veličina X udává počet neúspěchů před n-tým úspěchem v posloupnosti n nezávislých opakovaných pokusů, přičemž pravděpodobnost úspěchu je v každém pokusu ϑ. Píšeme X ~ NB(n, ϑ).

$$\pi(x) = \binom{n+x-1}{x} \vartheta^n (1-\vartheta)^x, \quad x = 0,1,\dots, \quad 0 < \vartheta < 1$$

$$= 0 \qquad\qquad jinak$$

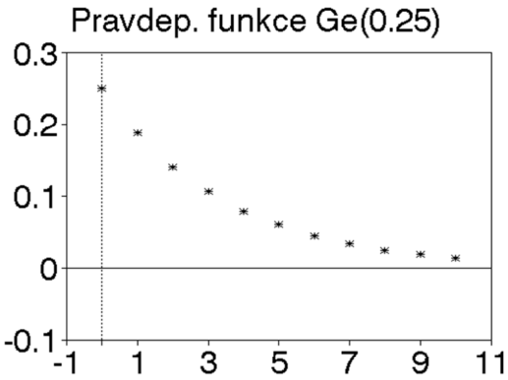
$NB(2,0.2) \quad a \quad NB(2,0.4)$

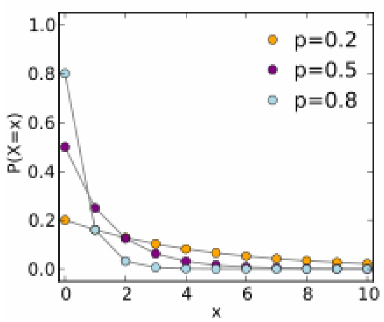


➢ Negativně binomické rozdělení lze definovat obecněji. Tak jak je zde uvedeno jde o rozdělení Pascalovo.

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 $X \sim Ge(\vartheta)$

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





Příklad

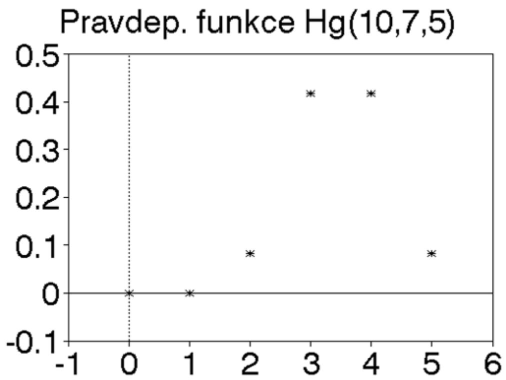
Dva hráči střídavě házejí kostkou. Vyhrává ten, kdo první hodí šestku. Jaká je pravděpodobnost, že vyhraje ten, který začínal?

Řešení: X udává počet nehození šestky (neúspěch) před prvním hozením šestky (úspěch), X ~ Ge(1/6). Hledáme tedy pravděpodobnost jevu A: 1. úspěch po sudém počtu neúspěchů.

$$P(A)=\sum_{k=0}^{\infty}\frac{1}{6}\left(\frac{5}{6}\right)^{2k}=\frac{1}{6}\frac{1}{1-\frac{25}{36}}=\frac{6}{11}=0{,}545$$

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 ~ Hg(N, M, n)

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



Příklad

V klobouku jsou 3 černé a 4 bílé koule. Určete pravděpodobnost, že při vytažení 3 koulí budou aspoň 2 černé.

Řešení: X udává počet vytažených černých koulí, X ~ HG(7,3,3). Hledaná pravděpodobnost je

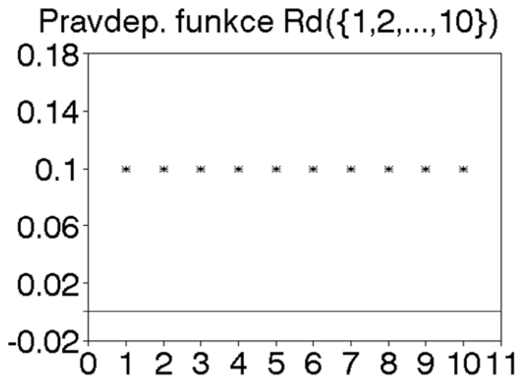
$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \frac{\binom{3}{0}\binom{4}{3} + \binom{3}{1}\binom{4}{2}}{\binom{7}{3}} =$$

$$1 - \frac{1 \cdot 4 + 3 \cdot 6}{35} = 1 - \frac{22}{35} = \frac{13}{35} = 0{,}371.$$

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

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

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



Vybraná rozložení spojitých náhodných veličin

Důležitá spojitá rozdělení:

- Rovnoměrné rozdělení
- Normální rozdělení (označované také jako Gaussovo rozdělení)
- Logaritmicko-normální rozdělení (také log-normální rozdělení)
- Studentovo rozdělení
- Fischerovo-Snedecorovo rozdělení
- χ^2 rozdělení (Chí kvadrát)
- Cauchyho rozdělení
- Exponenciální rozdělení
- Laplaceovo rozdělení (nebo také dvojitě exponenciální rozdělení)
- Weibullovo rozdělení

Rovnoměrné spojité rozložení: Předpokládejme, že veličina X

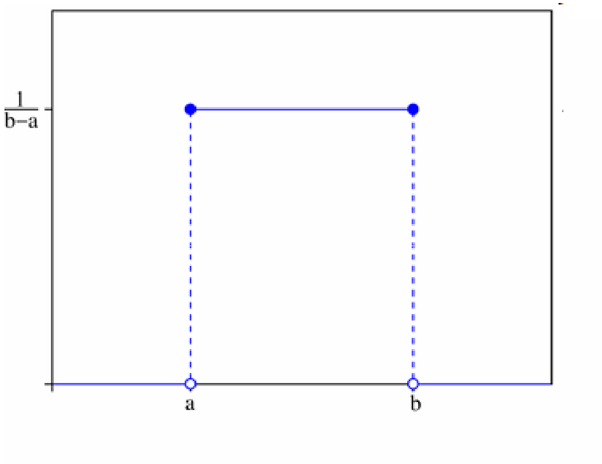
- může nabýt jakékoliv hodnoty mezi čísly a, b
- pravděpodobnost, že nabude hodnoty z jakéhokoliv intervalu v tomto rozmezí je stejná jako pravděpodobnost, že nabude hodnoty z jakéhokoliv jiného intervalu stejné délky.

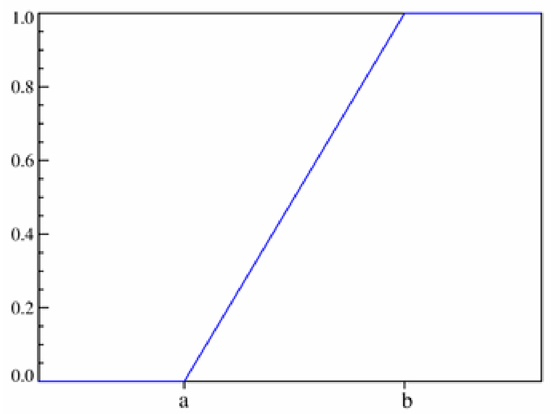
Jsou-li tyto podmínky splněny, pak X má rovnoměrné spojité rozložení na intervalu (a, b). Hustota pravděpodobnosti náhodné veličiny X je konstantní na intervalu (a, b) a plocha pod křivkou hustoty tvoří obdélník.

Píšeme X ~ Rs(a, b).

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

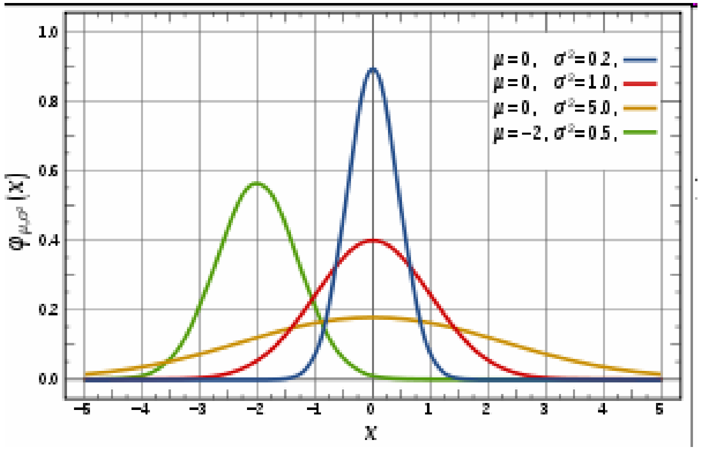
$$\Phi(x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & x \in (a,b) \\ 1 & x \ge b \end{cases}$$

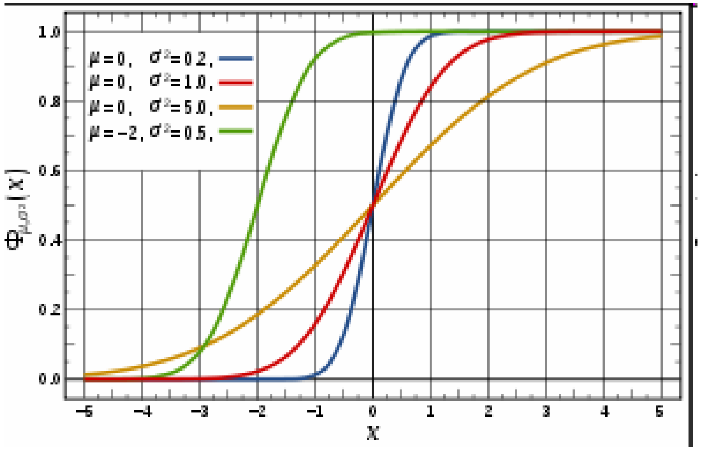




Normální rozložení: Tato náhodná veličina vzniká např. tak, že ke konstantě μ se přičítá velké množství nezávislých náhodných vlivů mírně kolísajících kolem nuly. Proměnlivost těchto vlivů je vyjádřena konstantou $\sigma > 0$.

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



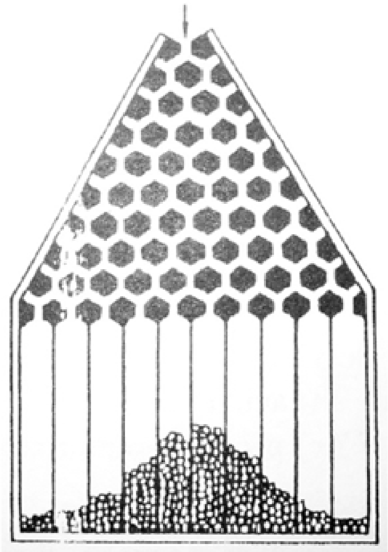


Ilustrace vzniku normálního rozložení pomocí Galtonovy desky:

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 shromaždují. 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.

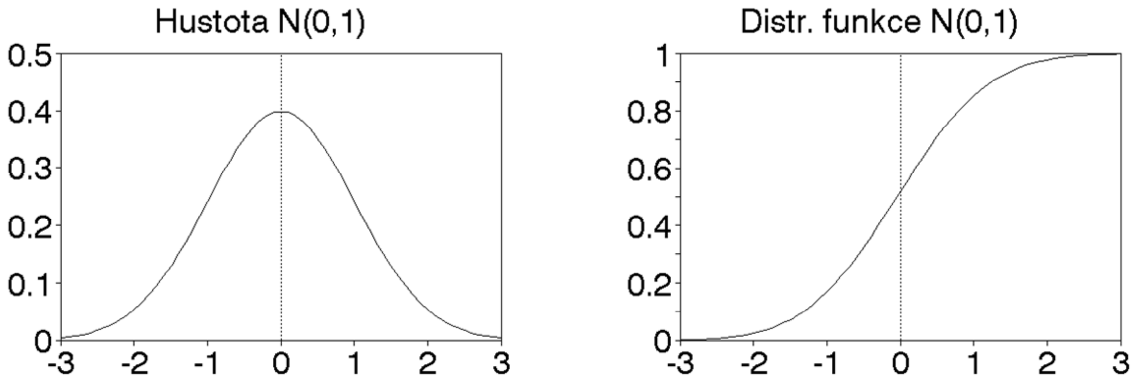
Náhodné vychylování kuliček jednotlivými řadami překážek je možno chápat jako speciální případ velkého množství chybových faktorů, náhodně působících na nějaký proces, jako působení mnoha blíže nespecifikovatelných vlivů, které ovlivňují zcela náhodně rozložení jeho výsledku.

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 $\Phi(-u) = 1 - \Phi(u)$.

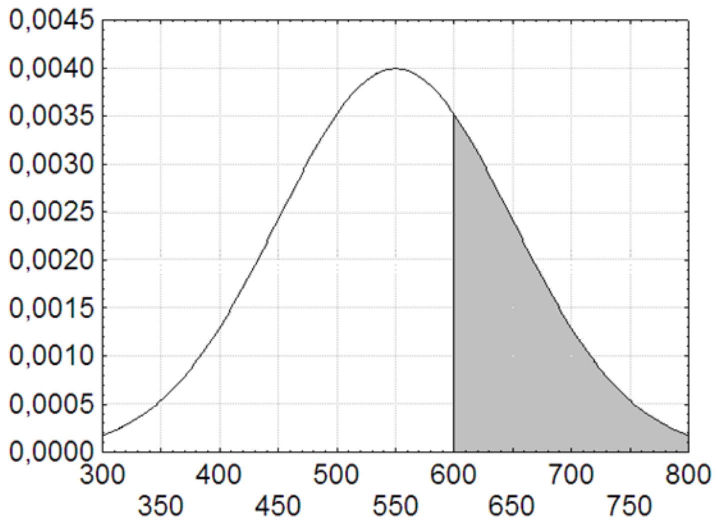
Příklad na normální rozložení: Výsledky u přijímacích zkoušek na jistou VŠ jsou normálně rozloženy s parametry $\mu = 550$ bodů, $\sigma = 100$ bodů. S jakou pravděpodobností bude mít náhodně vybraný uchazeč aspoň 600 bodů?

Řešení:

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

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

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



Některé vlastnosti normálního rozložení:

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

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

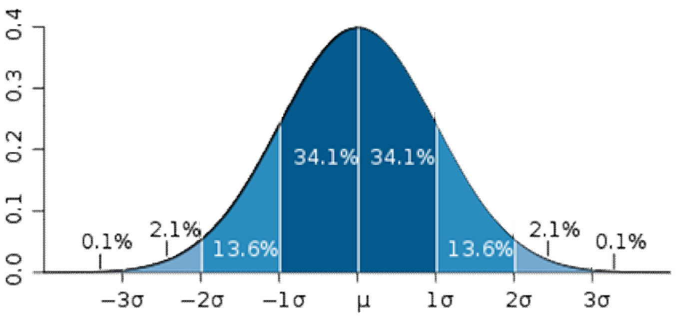
Jestliže $X_1, \ldots, X_n$ jsou stochasticky nezávislé náhodné veličiny, $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$, $Y = \sum_{i=1}^{n} X_i$, pak $Y \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$.

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 týmž rozložením (viz centrální limitní věta).

„koncentrace hodnot" normální NV:

Přes 68% hodnot „leží" v intervalu (μ-σ, μ+σ).
Přes 95% hodnot „leží" v intervalu (μ-2σ, μ+2σ).
Přes 99% hodnot „leží" v intervalu (μ-3σ, μ+3σ).

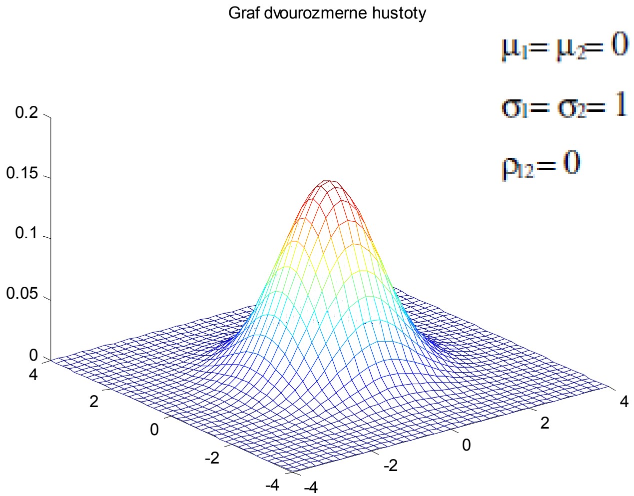


Definice:
O spojitém náhodném vektoru $\mathbf{X} = \binom{X_1}{X_2}$ říkáme, že má dvojrozměrné normální rozložení s parametry $\mu = \binom{\mu_1}{\mu_2}$ a $\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$, když jeho hustota je dána vzorcem

$$\varphi(\mathbf{x}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{x_1-\mu_1}{\sigma_1}\cdot\frac{x_2-\mu_2}{\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}, \quad \mathbf{x} \in \mathbf{R}^2.$$

Zkráceně píšeme $\mathbf{X} = \binom{X_1}{X_2} \sim N_2(\mu, \mathbf{\Sigma})$.
Pro $\mu = \binom{0}{0}$ a $\mathbf{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ mluvíme o standardizovaném dvojrozměrném normálním rozložení.

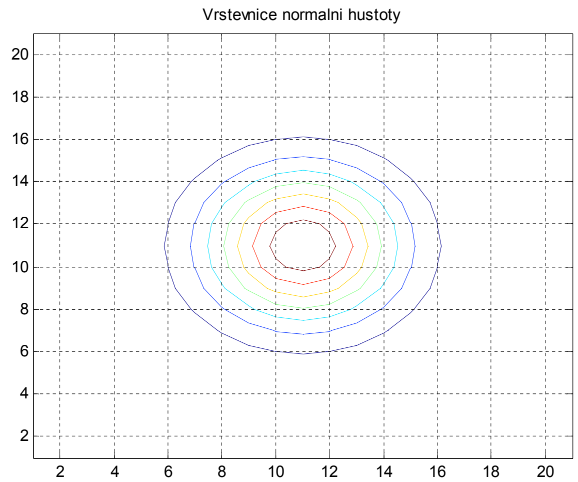
Poznámka:
Význam parametrů je následující:
$\mu_1 = E(X_1),\ \mu_2 = E(X_2),\ \sigma_1^2 = D(X_1),\ \sigma_2^2 = D(X_2),\ \rho = R(X_1, X_2)$

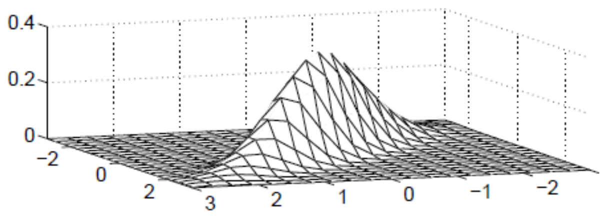


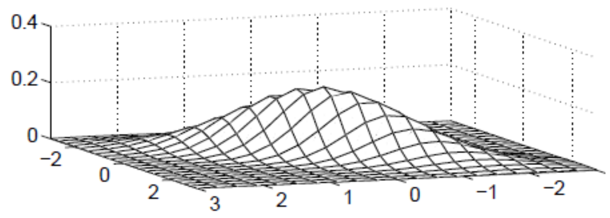
$\mu_1 = \mu_2 = 0$

$\sigma_1 = \sigma_2 = 1$

$\rho_{12} = 0$



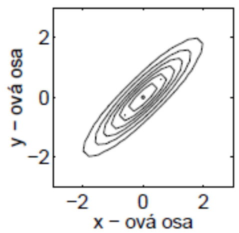




$\mu_1 = \mu_2 = 0$

$\sigma_1 = \sigma_2 = 1$

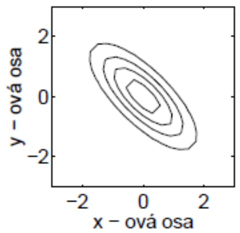
$\rho_{12} = 0.9$



$\mu_1 = \mu_2 = 0$

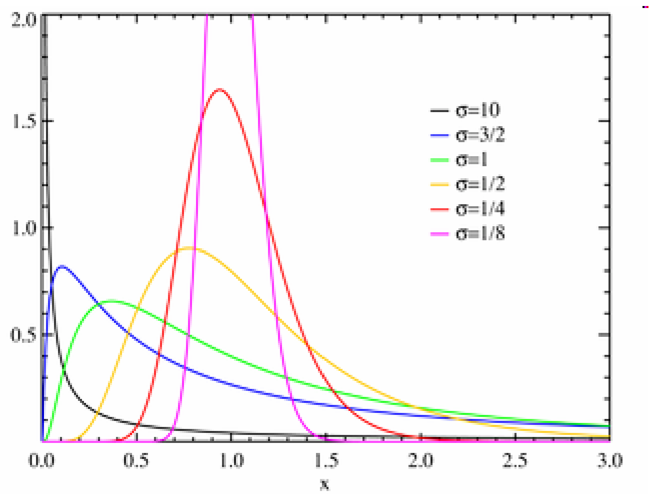
$\sigma_1 = \sigma_2 = 1$

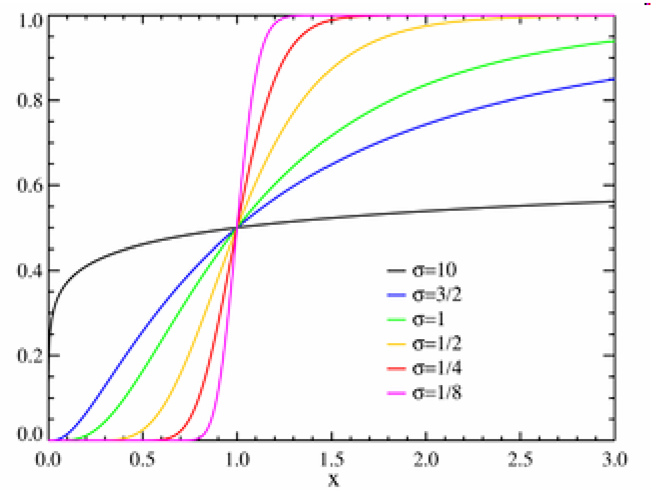
$\rho_{12} = -0.75$



Logaritmicko normální rozložení: Náhodná veličina X ~ LN(μ, σ) vzniká v situacích, kdy kladná konstanta logaritmu μ je násobena velkým množstvím nezávislých náhodných veličin, kolísajících mírně kolem jedníčky. Variabilita jejich logaritmů je charakterizována parametrem σ. Logaritmicko normální rozdělení má hustotu

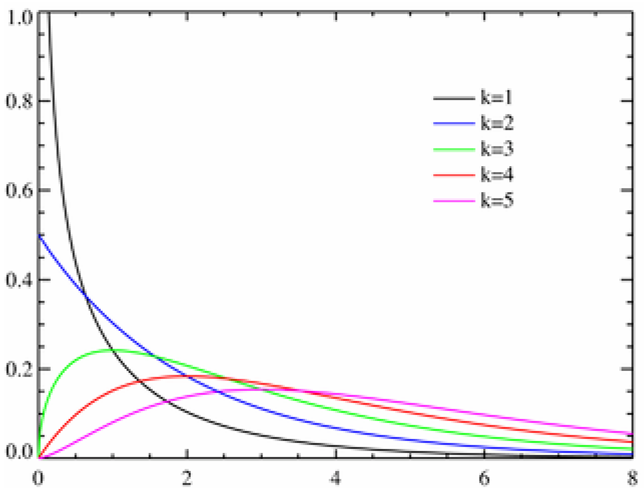
$$\varphi(x;\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \quad x > 0$$

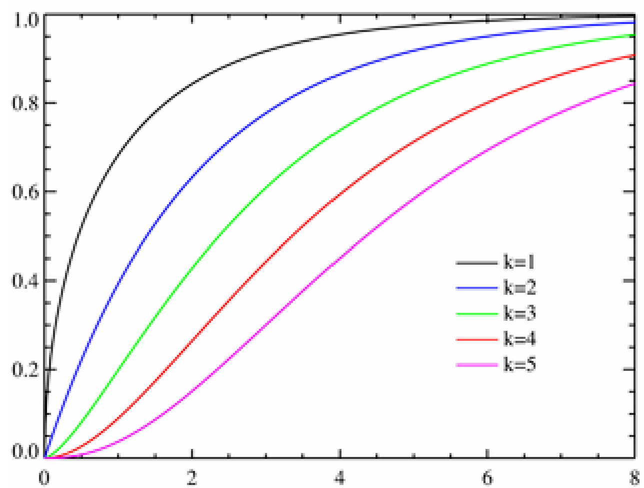




Pearsonovo rozložení chí-kvadrát s n stupni volnosti: Nechť $X_1, \ldots, X_k$ jsou stochasticky nezávislé náhodné veličiny, $X_i \sim N(0, 1)$, $i = 1, \ldots, k$.
Pak náhodná veličina $X = X_1^2 + \ldots + X_k^2 \sim \chi^2(k)$.

$$\varphi(x,k) = \begin{cases} \dfrac{1}{2^{k/2} \cdot \Gamma(k/2)} \cdot x^{k/2-1} \cdot e^{-x/2} & x > 0 \\ 0 & jinak \end{cases} \qquad \Gamma(s) = \int_0^\infty e^{-t} \cdot t^{s-1}\, dt$$

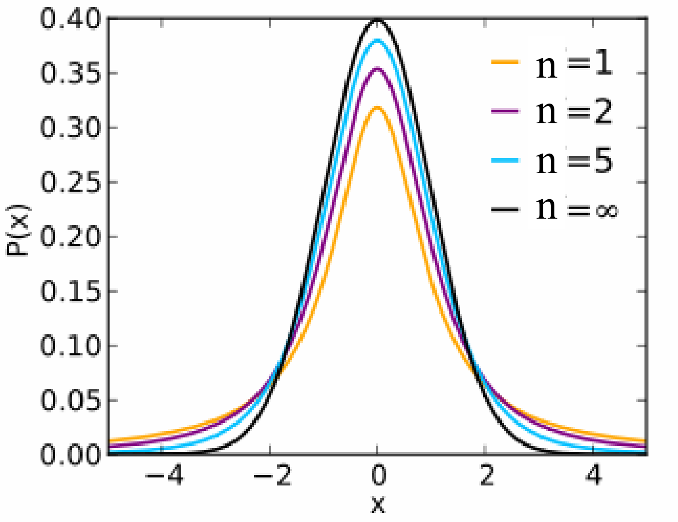


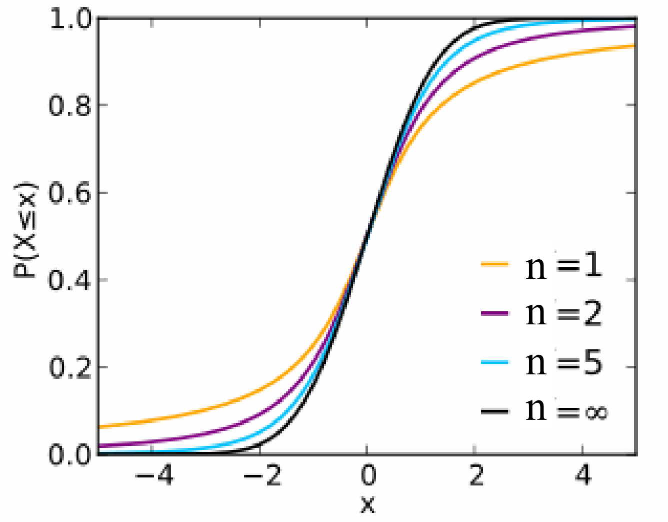


Studentovo rozložení s n stupni volnosti: Nechť X_1, X_2 jsou stochasticky nezávislé náhodné veličiny, $X_1 \sim N(0, 1)$, $X_2 \sim \chi^2(n)$.

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

$$\varphi(x,n) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma(n/2)} \cdot \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} \qquad x \in (-\infty, \infty)$$

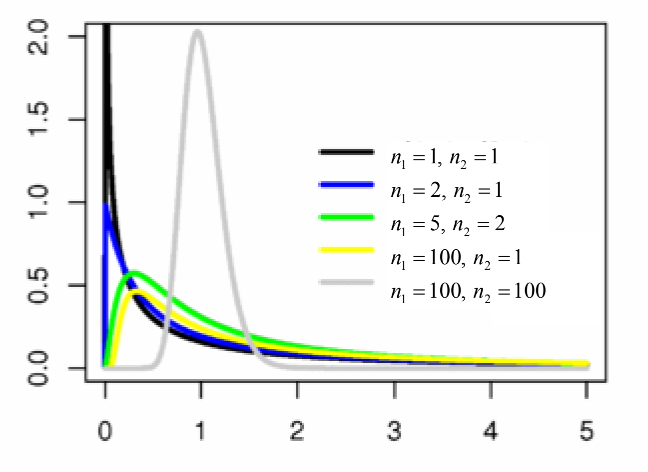


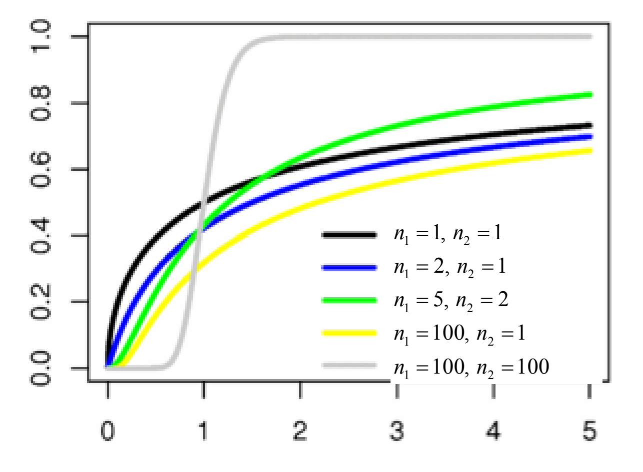


Fisherovo-Snedecorovo rozložení s n_1 a n_2 stupni volnosti: Nechť X_1, X_2 jsou stochasticky nezávislé náhodné veličiny, $X_i \sim \chi^2(n_i)$, i = 1, 2.

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

$$\varphi(x, n_1, n_2) = \frac{\Gamma\left(\dfrac{n_1+n_2}{2}\right) n_1^{n_1/2} n_2^{n_2/2}}{\Gamma(n_1/2)\Gamma(n_2/2)} \cdot \left(\frac{x^{(n_1-2)/2}}{(n_2+n_1 x)^{(n_1+n_2)/2}}\right) \quad \text{pro } x > 0$$





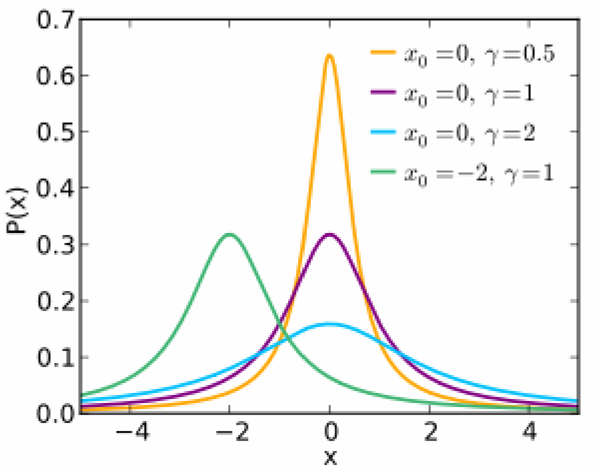
Cauchyho rozložení pravděpodobnosti s parametry x_0 a γ, pro $-\infty < x_0 < \infty$ a $\gamma > 0$, je definováno hustotou pravděpodobnosti ve tvaru

$$\varphi(x;x_0,\gamma) = \frac{1}{\pi\gamma\left[1+\left(\frac{x-x_0}{\gamma}\right)^2\right]} = \frac{1}{\pi}\left[\frac{\gamma}{(x-x_0)^2+\gamma^2}\right]$$

kde x_0 je parametr, určující umístění největší hodnoty rozdělení.
Zvláštní případ, kdy $x_0 = 0$ a $\gamma = 1$ se nazývá **standardní Cauchyho rozdělení** s hustotou pravděpodobnosti vyjádřenou vztahem

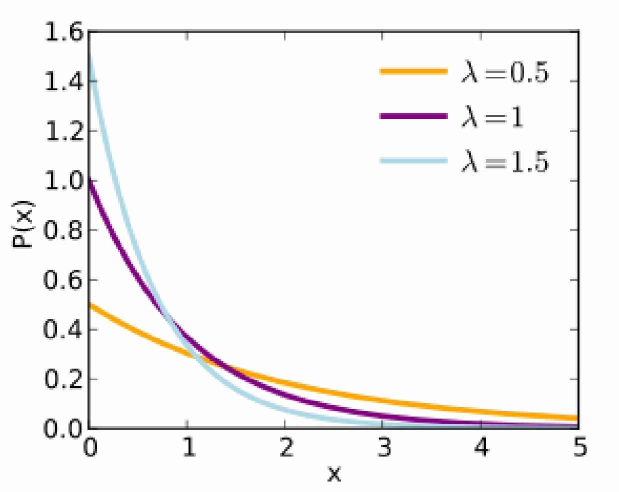
$$\varphi(x;0,1) = \frac{1}{\pi(1+x^2)}.$$

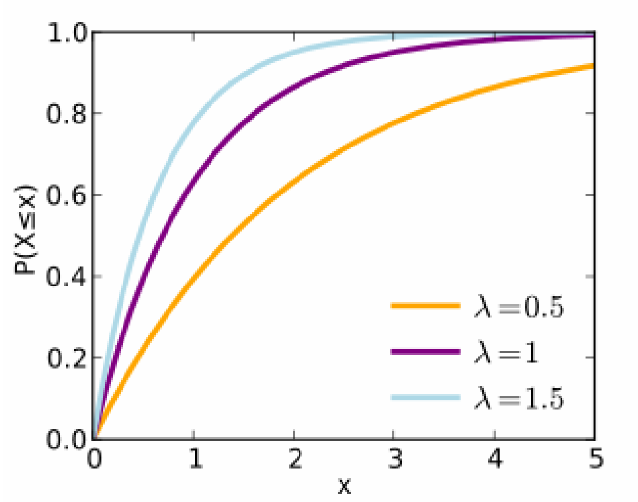
Standardní Cauchyho rozdělení je speciální případ Studentova rozdělení (pro n = 1).



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 $X \sim Ex(\lambda)$

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





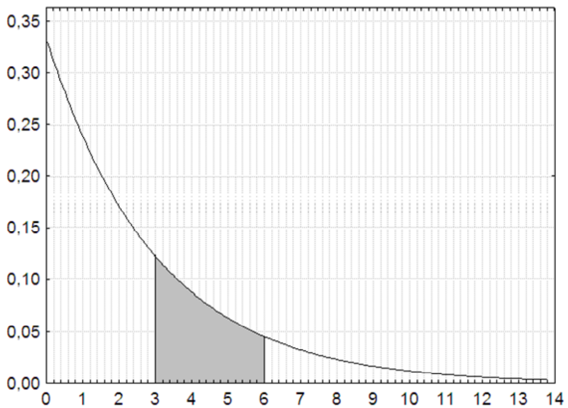
Příklad na exponenciální rozložení: Doba (v minutách) potřebná k obsloužení zákazníka v prodejně potravin je náhodná veličina, která se řídí rozložením $Ex\left(\frac{1}{3}\right)$. Jaká je pravděpodobnost, že doba potřebná k obsloužení náhodně vybraného zákazníka v této prodejně bude v rozmezí od 3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

Řešení:

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

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

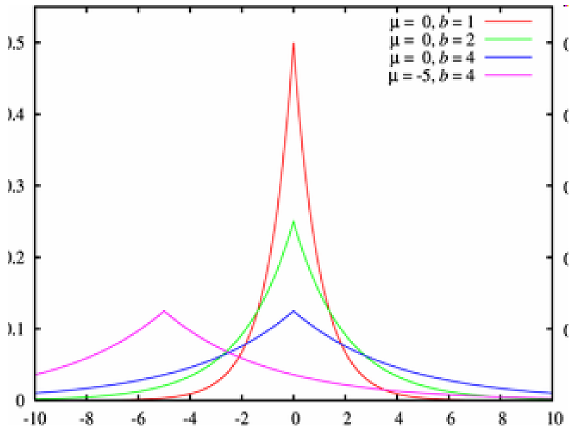
S pravděpodobností 0,233 bude zákazník obsloužen v době od 3 do 6 minut.



Laplaceovo rozložení: Náhodná veličina, která vznikne rozdílem dvou NV z exponenciálního rozložení, se řídí tímto rozložením. Využití ve fyzice, ekonomii – Brownův pohyb. Hustota je dána vzorcem

$$\varphi(x;\mu,b) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right)$$

$$= \frac{1}{2b}\begin{cases}\exp\left(-\frac{\mu-x}{b}\right) & x < \mu\\ \exp\left(-\frac{x-\mu}{b}\right) & x \geq \mu\end{cases}$$



Platí např.:

$$X \sim Laplace(0,b) \Rightarrow |X| \sim Ex\left(\frac{1}{b}\right)$$

$$X_1 \sim Ex(\lambda_1),\ X_2 \sim Ex(\lambda_2) \Rightarrow \lambda_1 X_1 - \lambda_2 X_2 \sim Laplace(0,1)$$

Weibullovo rozdělení: Náhodná veličina X ~ Wb(δ, ε) vyjad5uje dobu čekání na nějakou událost, která se každým okamžikem může dostavit se šancí úměrnou mocninné funkci pročekané doby. Přitom čísla $\delta > 0$ a $\varepsilon > 0$ se nazývají parametry měřítka a formy.

$$\varphi(x;\delta,\varepsilon) = \begin{cases} \varepsilon \cdot \delta \cdot \mathrm{x}^{\varepsilon\text{-}1} e^{-\delta \cdot x^{\varepsilon}} & \text{pro } x > 0 \\ 0 & \text{pro } x \le 0 \end{cases}$$

Jiná forma zápisu:

$$\varphi(x;\lambda,k) = \begin{cases} \dfrac{k}{\lambda}\left(\dfrac{\mathrm{x}}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^{k}} & \text{pro } x > 0 \\ 0 & \text{pro } x \le 0 \end{cases}$$