SPECIÁLNÍ TYPY ROZDĚLENÍ

Definice: Uvažujme pravděpodobnostní prostor $(S, \mathcal{A}, P)$ (množina S nemusí být nutně konečná). Náhodnou veličinou na pravděpodobnostním prostoru $(S, \mathcal{A}, P)$ budeme rozumět reálnou f-ci $X : S \to \mathbb{R}$ takovou, že pro každou Borelovskou podmnožinu $B \subset \mathbb{R}$ platí:

$$\{e \in S : X(e) \in B\} \in \mathcal{A}.$$

Poznámka: Lze ukázat, že $X : S \to \mathbb{R}$ je náhodnou veličinou $\iff$ $\forall b \in \mathbb{R}$:

$$\{e \in S : X(e) \leq b\} \in \mathcal{A}$$

$$\iff \forall a, b \in \mathbb{R} : \{e \in S : a \leq X(e) \leq b\} \in \mathcal{A}.$$

Příklad: Předpokládejme, že jsme náhodně vybrali osobu z jistého vzorku populace a změřili jsme výšku a hmotnost této osoby. Nyní uvažujme základní prostor $S = \mathbb{R}^2 = \{(x, y) : x =$ „výška osoby", $y =$ „hmotnost osoby"$\}$. Dále $\mathcal{A} = \mathcal{B} =$ „sigma algebra Borelovských podmnožin $\mathbb{R}^2$". X_1 budeme definovat předpisem:

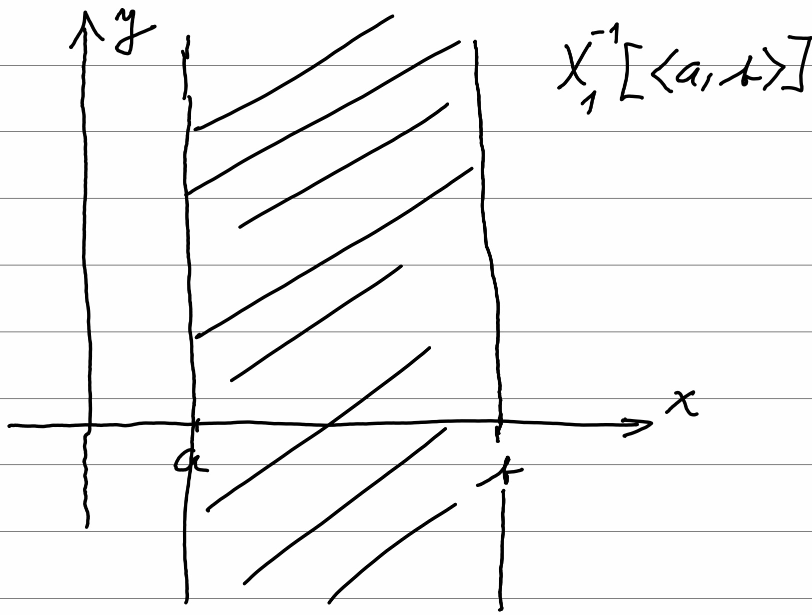
$$X_1(x, y) = x.$$

$$X_2(x,y) = y, \quad \forall (x,y) \in S = \mathbb{R}^2.$$

X_1 a X_2 jsou obě Borelovsky měřitelnými funkcemi. Např. $\forall a,b \in \mathbb{R}$ je $X_1^{-1}[\langle a,b\rangle] =$

$$= \{(x,y) : a \le X_1(x,y) \le b\} = \{(x,y) : a \le x \le b\}.$$

Tato množina je zřejmě Borelovskou množinou:



Klasifikace náhodných veličin

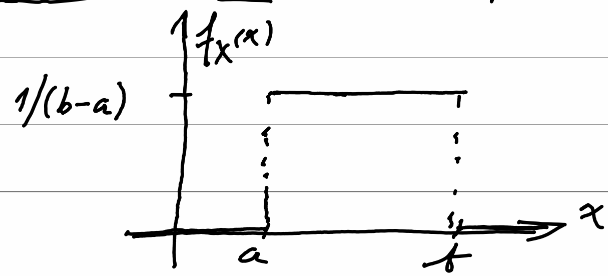
a) Řekneme, že náhodná veličina X je diskrétní náhodnou veličinou, je-li její obor hodnot buď konečnou nebo nekonečnou spočetnou množinou.

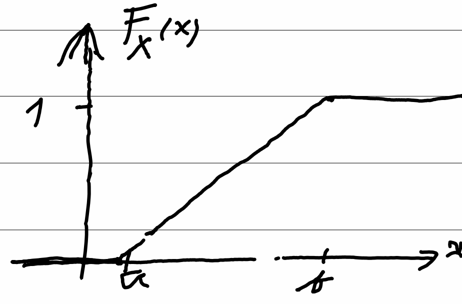
b) Řekneme, že náhodná veličina X je absolutně spojitá, existuje-li nezáporná funkce $f_X : \mathbb{R} \to \mathbb{R}$ taková, že platí:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt, \quad \forall x \in \mathbb{R},$$

kde F_X je distribuční funkcí náhodné veličiny X definované předpisem: $F_X(x) = P(\{X \leq x\})$.
Funkci f_X budeme nazývat hustotou pravděpodobnosti náhodné veličiny X.

Příklad na hustotu f_X:





DISTRIBUČNÍ FUNKCE NA $(\mathbb{R}, \mathcal{B}, \mathbb{P})$

Uvažujme pravděpodobnostní prostor $(\mathbb{R}, \mathcal{B}, \mathbb{P})$, kde $\mathcal{B}$ je σ-algebra Borelovských podmnožin $A \subset \mathbb{R}$. Pak $\forall x \in \mathbb{R}: \quad (-\infty, x\rangle \in \mathcal{B}$, potom funkce $F: x \longmapsto \mathbb{P}((-\infty, x\rangle)$ má vlastnosti:

(i) funkce $F(x)$ je neklesající;

(ii) $F(-\infty) = 0$, $F(+\infty)$, kde

$$F(-\infty) = \lim_{x \to -\infty} F(x) \quad \text{a}$$

$$F(+\infty) = \lim_{x \to +\infty} F(x);$$

(iii) funkce $F(x)$ je zprava spojitá v každém bodě $x \in \mathbb{R}$ a $\forall x \in \mathbb{R}$ existuje $\lim\limits_{t \to x-} F(t)$.

Definice: Každá funkce $F: \mathbb{R} \to \mathbb{R}$ splňující podmínky (i) – (iii) se nazývá **distribuční funkce** (na $\mathbb{R}$).

Věta : Nechť $F: \mathbb{R} \to \mathbb{R}$ je distribuční funkcí na reálné ose. Potom existuje na měřitelném prostoru $(\mathbb{R}, \mathcal{B})$ jediná pravděpodobnostní míra $\mathbb{P}: \mathcal{B} \to \mathbb{R}$ taková, že

$$\mathbb{P}((a,b\rangle) = F(b) - F(a),$$

kdykoliv je $a, b \in \mathbb{R}$, $-\infty \le a < b < \infty$.

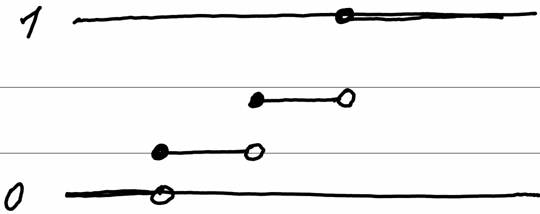
Důkaz : Důkaz nebude uvádět neboť bychom potřebovali mít k dispozici některé poznatky z teorie míry. $\square$

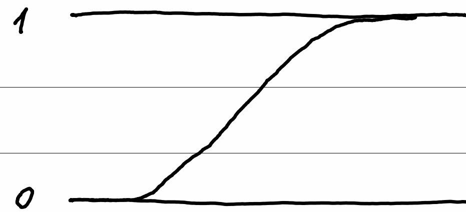
Z předchozí věty vyplývá, že máme-li dánu distribuční funkci $F: \mathbb{R} \to \mathbb{R}$, potom existuje pravděpodobnostní prostor $(S, \mathcal{A}, \mathbb{P})$ a náhodná veličina $X: S \to \mathbb{R}$ taková, že

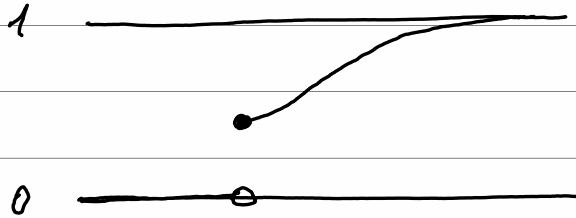
$$F = F_X.$$

Připomeňme, že $F_X: \mathbb{R} \to \mathbb{R}$ je definována předpisem : $F_X(x) = \mathbb{P}(\{X \le x\})$, $\forall x \in \mathbb{R}$.

Grafy distribučních funkcí mohou vypadat například takto:







Příklady pravděpodobnostních rozdělení

NÁZEV	PRAVDĚPODOBNOSTI p_k	PARAMETRY
Diskrétní stejnoměrné	$1/N,\ k=1,\dots,N$	$N=1,2,\dots$
Bernoulliho	$p_1=p,\ p_0=q$	$0\le p\le 1,\ q=1-p$
Binomické rozdělení	$\binom{n}{k} p^k q^{n-k}$ $k=0,1,2,\dots$	$0\le p\le 1,\ q=1-p$ $n=1,2,\dots$
Poissonovo rozdělení	$e^{-\lambda}\lambda^k/k!,\ k=0,1,\dots$	$\lambda>0$

geometrické rozdělení	$q^{k-1}p,\ k=1,2,\dots$	$0\le p\le 1,\ q=1-p$
Pascalovo rozdělení	$\binom{k-1}{r-1}p^r q^{k-r}$, $k=r,r+1,\dots$	$0\le p\le 1,\ q=1-p$, $r=1,2,\dots$

Příklady: 1) Uvažujme náhodný pokus při němž se dvakrát háže mincí přičemž hody jsou na sobě nezávislé. Zapišme základní prostor:

$$S=\{PP,\ PO,\ OP,\ OO\}.$$

Dále nechť $\mathcal{A}=\mathcal{P}(S)=$ potenční množina a nechť $P(PP)=0.36$, $P(PO)=P(OP)=0.24$, $P(OO)=0.16$. Nechť náhodná veličina X označuje počet hodů kdy padla na minci „panna". Obor hodnot je roven množině $\{x_1,x_2,x_3\}=\{0,1,2\}$. Pravděpodobnostní funkce p_X nabývá hodnot:

$$p_X(k)=\binom{n}{k}p^k\cdot q^{n-k},\quad n=2,\ k=0,1,2,\ p=0.6,\ q=0.4$$

Tedy dostáváme hodnoty:

$p_X(0)=\binom{2}{0}\times 0.6^0\times 0.4^{2-0}=1\times 1\times 0.4^2=0.16$;

$$p_X(1) = \binom{2}{1} \times 0.6^{1} \times 0.4^{2-1} = 2 \times 0.6 \times 0.4 = 0.48,$$

$$p_X(2) = \binom{2}{2} \times 0.6^{2} \times 0.4^{2-2} = 1 \times 0.36 \times 1 = 0.36.$$

Dále spočítejme funkční hodnoty distribuční funkce F_X:

$$F_X(0) = P(\{X \leq 0\}) = \sum_{i,\, x_i \leq 0} p_X(x_i) =$$

$$= p_X(0) = 0.16;$$

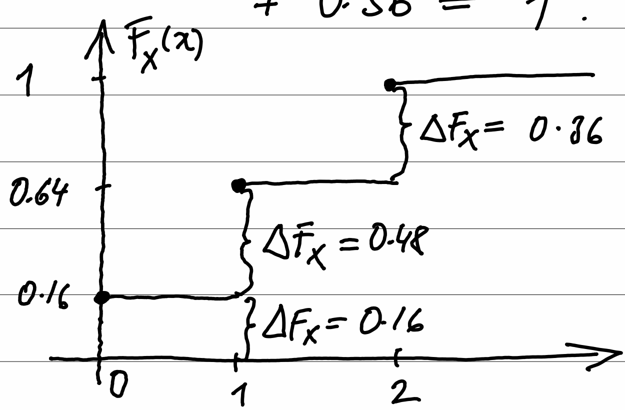
$$F_X(1) = P(\{X \leq 1\}) = \sum_{i,\, x_i \leq 1} p_X(x_i) = p_X(0) +$$

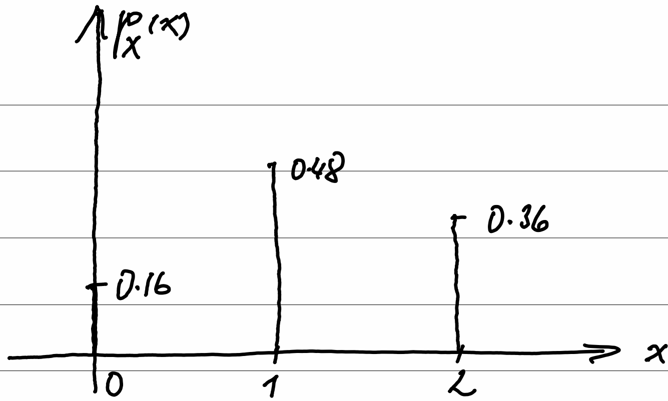
$$+ p_X(1) = 0.16 + 0.48 = 0.64,$$

$$F_X(2) = P(\{X \leq 2\}) = \sum_{i,\, x_i \leq 2} p_X(x_i) =$$

$$= p_X(0) + p_X(1) + p_X(2) = 0.16 + 0.48 +$$

$$+ 0.36 = 1.$$





2) Předpokládejme, že reálné číslo X je náhodně v intervalu $\langle a,b\rangle$. Dále nechť X má stejnoměrné rozdělení pravděpodobnosti. Položme $S = \mathbb{R}$, $\mathcal{A} = \mathcal{B}(\mathbb{R})$ = σ-algebra Borelovských podmnožin $\mathbb{R}$, $\forall e \in \mathbb{R}$: $X(e) = e$, $f(x) = f_X(x) =$

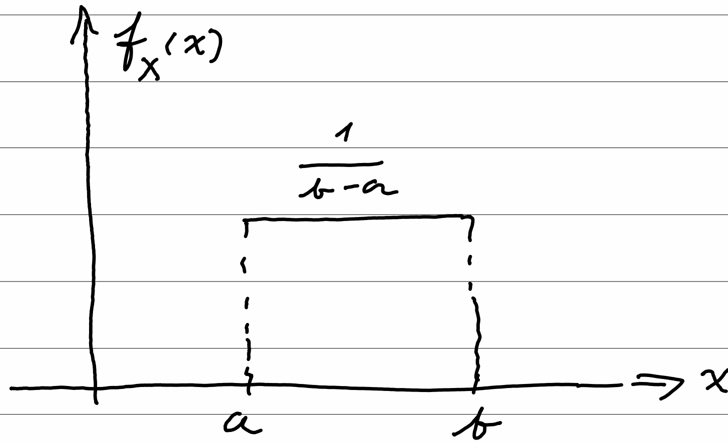
$$= \begin{cases} \dfrac{1}{b-a} & \text{pro } x \in \langle a,b\rangle \\ 0 & \text{pro } x \notin \langle a,b\rangle . \end{cases}$$

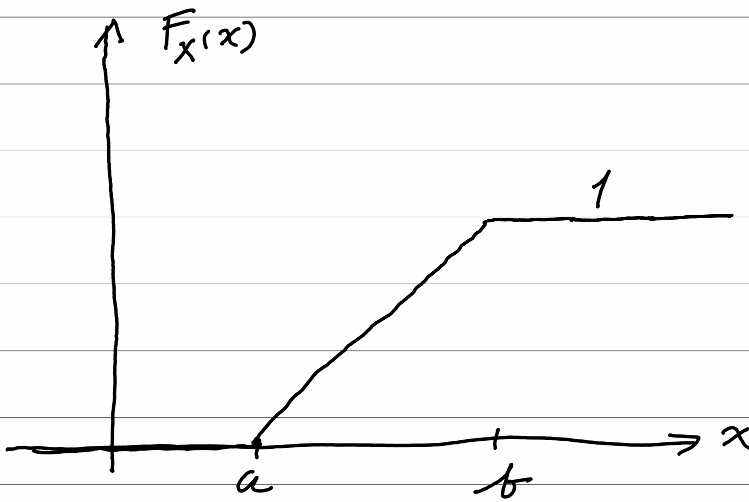
Pravděpodobnostní míra $\mathbb{P} : \mathcal{B}(\mathbb{R}) \to \mathbb{R}$ bude definovaná předpisem:

$$\mathbb{P}(B) = \int_B f(x)\,dx .$$

Pokud $B \subset \langle a,b\rangle$ a B je interval, pak $\mathbb{P}(B) = (\text{délka int. } B)/(b-a)$.

graf hustoty $f = f_X$ a distribuční funkce F_X vypadají takto:



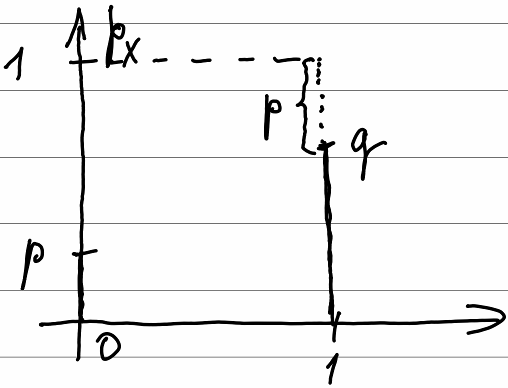


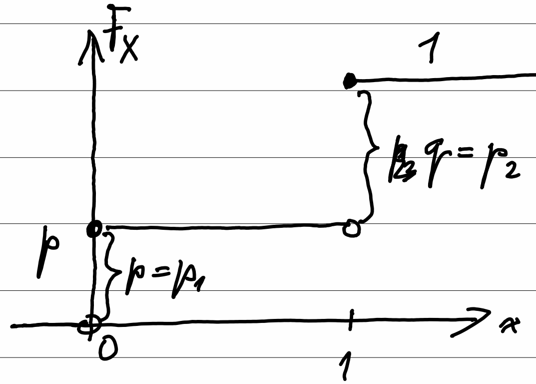
3) Alternativní rozdělení pravděpodobnosti

Náhodná veličina nabývá pouze dvou hodnot $X = 1 =$ „úspěch", $X = 0 =$ „neúspěch".
Pak

$$p_1 = \mathbb{P}(\{X = 1\}) = p \in (0,1) \quad \text{a}$$

$$p_0 = \mathbb{P}(\{X = 0\}) = q = 1 - p.$$





$$\boxed{EX} = \text{„střední hodnota"} = \sum x_i \cdot p_X(x_i) =$$

$$= 0 \cdot q + 1 \cdot p = \boxed{p.}$$

$$DX = \text{„rozptyl"} := \sum_i (x_i - EX)^2 p_X(x_i)$$

$$= (0-p)^2 \cdot p_0 + (1-p)^2 \cdot p_1 = p^2 \cdot (1-p) +$$

$$+ (1-p)^2 \cdot p = p^2 \cancel{- p^3} + p - 2p^2 \cancel{+ p^3} =$$

$$= p - p^2 = p(1-p) = p \cdot q \implies$$

$$\boxed{DX = pq}$$

4) <u>geometrické</u> <u>rozdělení</u> <u>pravděpodobnosti</u>:

Zde náhodná veličina X nabývá celých nezáporných hodnot. Zde pravděpodobnostní funkce p_X má předpis:

$$p_X(k) = \mathbb{P}(\{X=k\}) = p^k \cdot q,$$

kde $k = 0, 1, 2, \dots$ a $q = 1-p$, $p \in (0,1)$.

Spočtěme opět s tzv. střední hodnotu a rozptyl pro takovéto rozdělení pravděpodobnosti:

$$\boxed{\mathbb{E}X} = \sum_{i=0}^{\infty} x_i \cdot p_X(x_i) = \sum_{i=0}^{\infty} i \cdot p^i \cdot (1-p) =$$

$$= (1-p) \underbrace{\sum_{i=0}^{\infty} i \cdot p^i}_{\text{jedná se o tzv. aritmeticko-geometrickou řadu}} = \cancel{(1-p)} \frac{p}{(1-p)^{\cancel{2}}} = \frac{p}{1-p} = \boxed{\frac{p}{q}}$$

$$\mathbb{D}X = \sum_{i=0}^{\infty} (x_i - \mathbb{E}X)^2 p_X(x_i) = \sum_{i=0}^{\infty} \left(i - \frac{p}{q}\right)^2 p^i \cdot (1-p)$$

$$= \dots = \boxed{p/q^2}.$$

Toto rozdělení popisuje pravděpodobnost při opakovaném provádění Bernoulliova pokusu přičemž X označuje počet úspěšných pokusů předcházející prvnímu neúspěšnému pokusu. Pravděpodobnost úspěšného pokusu $= p \in (0,1)$ q označuje pravděpodobnost pokusu neúspěšného.

Sečteme nyní řadu $\sum_{i=0}^{\infty} i \cdot p^i$

$$S_n = p + 2p^2 + \ldots + np^n \quad / \cdot p$$

$$pS_n = \quad + p^2 + 2p^3 + \ldots + (n-1)p^n + np^{n+1}$$

$$S_n(1-p) = \underbrace{p + p^2 + p^3 + \ldots + p^n}_{\text{suma geometrické řady}} - np^{n+1} =$$

$$= p\frac{1-p^n}{1-p} - np^{n+1} \Rightarrow$$

$$\Rightarrow S_n = p\frac{1-p^n}{(1-p)^2} - \frac{np^{n+1}}{1-p} = \frac{p - p^{n+1} - np^{n+1}(1-p)}{(1-p)^2}$$

$$= \frac{p - \overset{\to 0}{p^{n+1}} - \overset{\to 0}{np^{n+1}} + \overset{\to 0}{np^{n+2}}}{(1-p)^2} \longrightarrow \frac{p}{(1-p)^2}$$

5) Hypergeometrické rozdělení pravděpodobnosti

Předpokládejme, že náhodná veličina X nabývá celých nezáporných hodnot $0, 1, 2, \ldots, n$. Pak je pravděpodobnostní funkce dána vztahem:

$$p_X(m) = \mathbb{P}(\{X = m\}) =$$

$$= \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}, \quad m = 0, 1, 2, \ldots, n.$$

Zde $M, N, n \in \mathbb{N}$ a platí podmínky:

$N > M$, $N > n$, $M > m$ a

$N - M > n - m$.

Aplikovat lze na situaci, kdy máme vzorek mající N prvků z nichž M má sledovanou vlastnost. Pak náhodně vybereme n prvků přičemž vybrané prvky zpět nevracíme. X pak bude označovat počet vybraných prvků, které mají sledovanou vlastnost.

Pak lze výpočtem zjistit, že:

$$\boxed{\mathbb{E}X = \sum{}' x_i \cdot p_X(x_i) = n\frac{M}{N}} \text{ a } \boxed{\mathbb{D}X = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}}$$

<u>Poznámka</u>: Uvažujme nezápornou Riemannovsky integrabilní funkci $f: \mathbb{R} \to \mathbb{R}_0^+$ takovou, že

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Položme $S = \mathbb{R}$, $\mathcal{A} = \mathcal{B}(\mathbb{R}) = \sigma$-algebra Borelovských množin. Platí následující věta:

(*) <u>Věta</u>. Nechť $f: \mathbb{R} \to \mathbb{R}_0^+$ je Riemannovsky integrabilní fce takovou, že $\int_{-\infty}^{\infty} f(x)\,dx = 1$. Potom existuje jediná pravděpodobnostní míra $\mathbb{P}: \mathcal{B}(\mathbb{R}) \to \mathbb{R}_0^+$ taková, že

$$P((a,b\rangle) = \int_a^b f(x)\,dx \quad \text{pro}$$

všechny intervaly $(a,b\rangle \subset \mathbb{R}$.

Tedy je-li dána nezáporná R-integrabilní fce $f: \mathbb{R} \to \mathbb{R}_0^+$ taková, že $\int_{-\infty}^{\infty} f(x)\,dx = 1$, pak díky předchozí větě definujeme $\forall B \in \mathcal{B}(\mathbb{R})$:

$$\int_B f(x)\,dx := \mathbb{P}(B),$$

kde $\mathbb{P}$ je jednoznačně určená pravděpodobnostní míra.

Nyní předpokládejme, že X je absolutně spojitá náhodná veličina. Pak $\forall x \in \mathbb{R}$:

$$F_X(x) = P(\{X \le x\}) = \int_{-\infty}^{x} f_X(t)\,dt.$$

Pak platí: $F_X(x) \to 1$ pro $x \to \infty \Rightarrow$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1.$$

Dále platí: $\mathbb{P}(\{a < X \le b\}) = \int_a^b f_X(x)\,dx.$

Pokud $\forall B \in \mathcal{B}(\mathbb{R})$ položíme $Q(B) := \mathbb{P}(\{X \in B\})$, pak $Q : \mathcal{B}(\mathbb{R}) \to \mathbb{R}_0^+$ je pravděpodobnostní mírou na prostoru $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ a pro každý interval $(a, b\rangle \subset \mathbb{R}$ platí:

$$Q((a,b\rangle) = \int_a^b f_X(x)\,dx.$$

Pak z Věty (*) plyne, že $\forall B \in \mathcal{B}(\mathbb{R})$:

$$\boxed{\mathbb{P}(\{X \in B\}) = Q(B) = \int_B f_X(x)\,dx.}$$

Je-li tedy X absolutně spojitou náhodnou veličinou a $c \in \mathbb{R}$, potom:

$$\mathbb{P}(\{X = c\}) = \mathbb{P}(\{c \leq X \leq c\}) = \int_c^c f_X(x)\,dx$$
$$= 0.$$

Odtud máme: $\mathbb{P}(\{a \leq X \leq b\}) =$
$$= \mathbb{P}(\{a < X \leq b\}) = \mathbb{P}(\{a \leq X < b\})$$
$$= \mathbb{P}(\{a < X < b\}) =$$
$$= \int_a^b f_X(x)\,dx = F_X(b) - F_X(a).$$

Je-li hustota pravděpodobnosti $f_X(x)$ spojitou funkcí, pak $\forall x \in \mathbb{R}$ platí:

$$\boxed{\frac{d}{dx}\left(\int_{-\infty}^{x} f_X(t)\,dt\right) = F_X'(x) = f_X(x).}$$

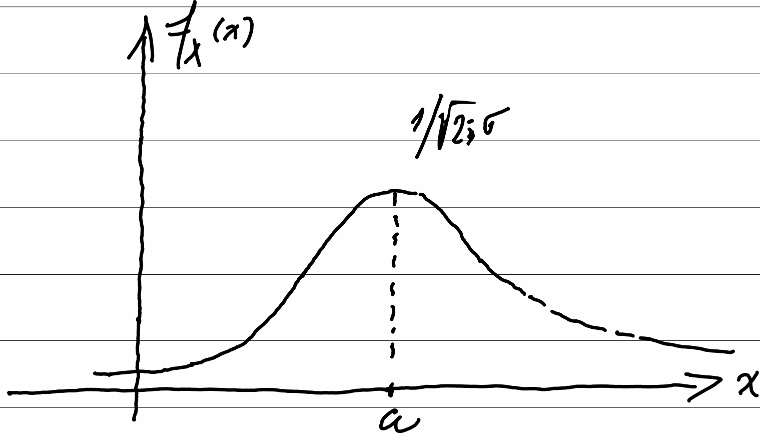
<u>VĚTA</u>: Je-li X absolutně spojitá náhodná veličina, potom je distribuční funkce F_X spojitou funkcí.

Příklad (Náhodná veličina s normálním rozdělením pravděpodobnosti).

Je-li X náhodná veličina s hustotou pravděpodobnosti:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-\mu)^2/2\sigma^2}, \quad x \in \mathbb{R}$$

$(b > 0,\ a \in \mathbb{R})$, pak říkáme, že náhodná veličina X má normální rozdělení pravděpodobnosti.



Lze spočítat, že $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$. Odtud pak plyne, že

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-\mu)^2/2\sigma^2}\,dx = \left| y = \frac{x-\mu}{\sqrt{2}\,\sigma} \right| = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}}\, e^{-y^2}\,dy = 1.$$