přednáška 06:
spojitá náhodná veličina

Po *diskrétní náhodné veličině projdeme ještě základní pojmy u spojité náhodné veličiny*

Následující výklad viz Budíková, Králová, Maroš: Průvodce základními statistickými metodami, str. 69-76

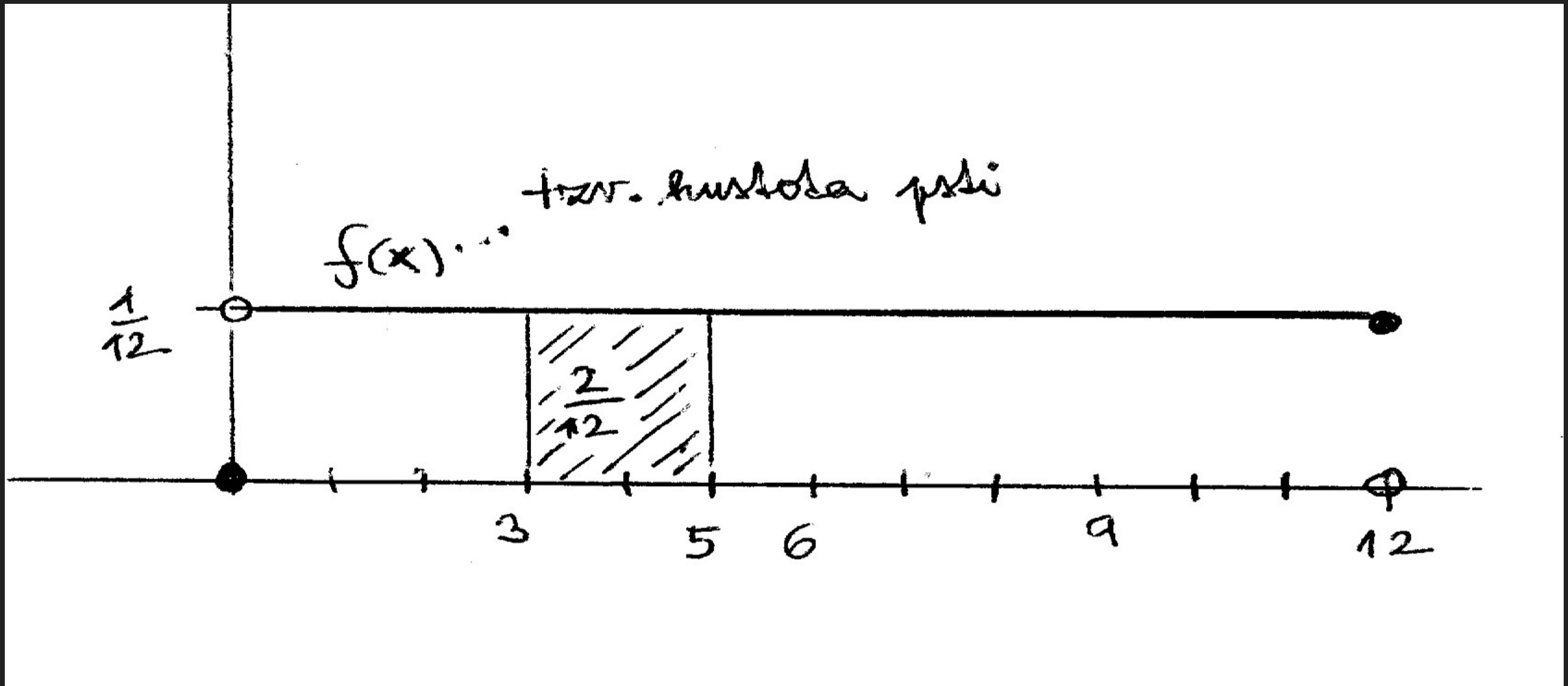
výklad pro EX, DX doplněn z el.textu Matematika 3, kapitola 10.

Příklad 3: představte si hodinový ciferník a malou hodinovou ručičku, která se náhodně zastaví v určité poloze

Poloh, ve kterých se ručička může zastavit, je nekonečně mnoho.

Pak pst, že se ručička zastaví v určité konkrétní poloze, musí být rovna nule. Jak tedy počítat a „měřit“ pst?

Pst, že se ručička zastaví např. mezi 3. a 5. hodinou, je nenulová a její hodnota je vyznačena na obrázku:



Spojitá náhodná veličina X tedy může nabývat hodnot z intervalu nebo z celé množiny reálných čísel.

Další příklady spojitých veličin: teplota, doba mezi příchody dvou emailů nebo dvou zákazníků, životnost výrobku, výška stromu, výška člověka, atd.

Klasické značení: platí i pro spojitou veličinu

X … označení veličiny

x … konkrétní hodnota veličiny

$X \leq x$ … náhodná veličina X nabývá hodnoty menší než nebo rovné x

$X \leq 2$ … náhodná veličina X nabývá hodnoty menší než nebo rovné 2

Definice: Distribuční funkce F(x) náhodné veličiny X je pst, že veličina X nabývá hodnoty menší než nebo rovné x ..... $F(x) := P(X \leq x)$

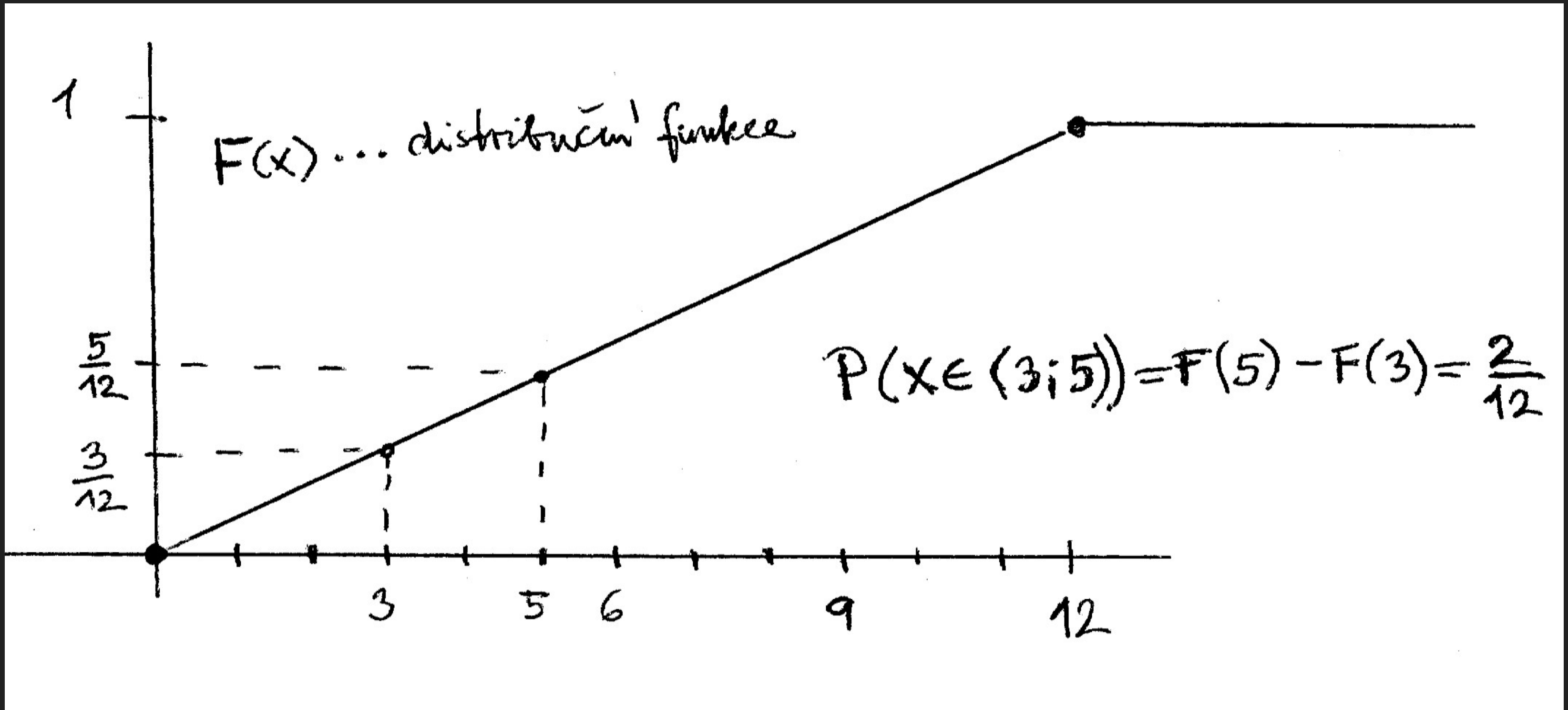
Definice distribuční funkce je tedy stejná jako u diskrétní veličiny

Pokračujme v příkladu 3: najdeme graf distribuční funkce F(x) místa, kde se točící ručička náhodně zastaví

Distribuční funkce $F(x) := P(X \leq x)$

Pokračujme v příkladu 3: distribuční funkce funguje podobně jako kumulativní relativní četnost u měření hodnot – s tím rozdílem že nyní u spojité veličiny relativní četnost narůstá spojitě

Distribuční funkce $F(x) := P(X \leq x)$ *u příkladu 3:*



Jedná se vlastně o ot.10: model psti 04 – spojitá pst … pouze jsme dodali další pojmy (distrib fce, střední hodnota, rozptyl a směrodatná odchylka)

Z axiomů psti plyne pro hustotu psti f(x):

1. $\int_{-\infty}^{+\infty} f(x)dx = 1$ … axiom normovanosti
2. $p(\omega_i) = \int_{\omega_i}^{\omega_i} f(x)dx = 0$ (pst měříme pomocí obsahů, nikoli pomocí funkčních hodnot);

   P((a;b))= $\int_a^b f(x)dx \geq 0$ (protože hustota f je nezáporná, tj. obsah podgrafu je nezáporný)
3. P((a;b))= $\int_a^b f(x)dx$ (součet má formu integrálu)

V přednášce 2 jsme počítali příklad:

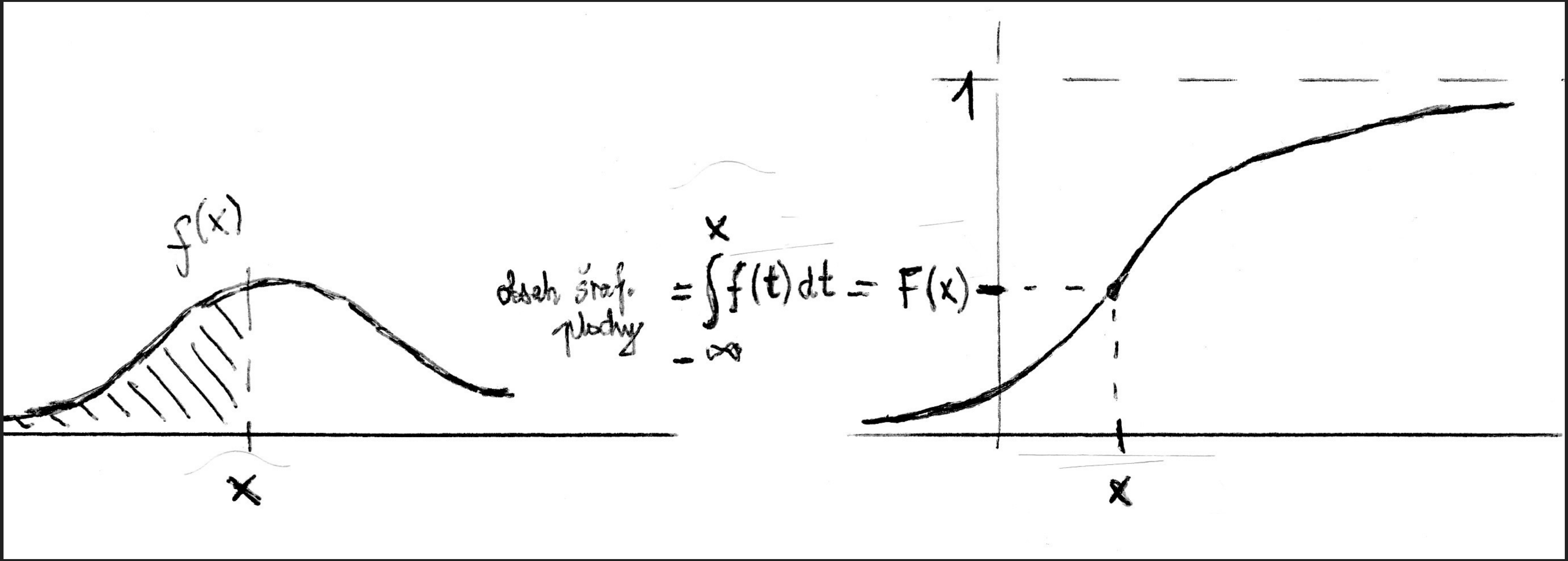
Př: mobily jisté firmy mají životnost v průměru 15 let; jaká je pst, že náhodně koupený mobil vydrží více než 10 let? X= životnost mobilu

vlastnosti distribuční funkce $F(x) := P(X \leq x)$ jsou stejné jako u diskrétní veličiny … přitom specifické vlastnosti jsou vyznačeny žlutě

1. F je neklesající.
2. F je spojitá na celé množině reálných čísel
3. F je normovaná, tj. $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$
4. $P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x)dx$
5. $P(X = x_0) = 0$ … (díky vlastnosti (2) nemá F body nespojitosti)
6. $P(X > x) = 1 - P(X \leq x) = 1 - F(x)$ … důležité pro výpočty pstí

Specifická vlastnost 01: F získáme z f integrací

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$



Specifická vlastnost 02: hustotu psti f získáme z distribuční funkce F derivací

V bodech, ve kterých existuje derivace: $F'(x) = f(x)$

Střední hodnota EX spojité veličiny X:

Definice střední hodnoty vychází z pojmů vážený aritmetický průměr měření:

$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{k} n_i x_i = \bar{x} = \sum_{i=1}^{k} \frac{n_i}{n} x_i = \sum_{i=1}^{k} p_i x_i,$

- Kde p_i jsou relativní četnosti měření dané hodnoty

Přesně stejnou strukturu má i vzorec pro střední hodnotu veličiny X, jen místo empirických (= naměřených) četností p_i ve vzorci vystupují teoretické psti p_i a místo součtu píšeme integrál

Střední hodnota EX spojité veličiny X:
EX= $\int_{-\infty}^{\infty} x \cdot f(x)\, dx$, kde meze případně měníme na interval (a;b), na kterém je hustota f(x) různá od nuly

Střední hodnota EX náhodné veličiny X je tedy jakýsi teoretický průměr, který bychom vypočetli při velkém množství měření veličiny X, pokud by se veličina X chovala přesně podle daného teoretického popisu (popsaného hustotou f)

Na rozdíl od diskrétní veličiny sumu u spojité velič nahradíme integrálem … ad příklad 3: u tabule

Pozor na konflikt s analýzou 1:

V analýze reálné funkce existuje pojem střední hodnota funkce jako $\frac{1}{b-a} \cdot \int_a^b f(x)dx$ výška obdélníka (na ose y), která po vynásobení dílkou interval (b-a) dá obsah podgrafu f

- Střední hodnota v psti je něco zcela jiného – je to hodnota na ose x, která by se pro větší a větší množství měření veličiny X stále více blížila průměru těchto měření

(výška ideálního obdélníka zde nemá žádný význam, protože psti u spojité veličiny neměříme podle fukčních hodnot f(x), ale podle obsahu plochy podgrafu f(x), nemluvě o intervalu nekonečné délky, kde je výška ideálního obd = 0, kdežto EX ne)

Rozptyl DX a směrodatná odchylka $\sqrt{DX}$ spojité veličiny

Podobně jako u střední hodnoty: rozptyl náhodné veličiny X vychází z pojmu rozptyl hodnot měření:

$s^2 = \frac{1}{n} \cdot \sum_{i=1}^{k} n_i (x_i - \bar{x})^2 = \sum_{i=1}^{k} \frac{n_i}{n} (x_i - \bar{x})^2 = \sum_{i=1}^{k} p_i (x_i - \bar{x})^2,$

- Kde p_i jsou relativní četnosti měření dané hodnoty;

Úpravou lze převést na tvar $s^2 = \left(\sum_{i=1}^{k} p_i \cdot x_i^2\right) - \bar{x}^2$

(vážený průměr čtverců MINUS čtverec průměru)

stejnou strukturu má vzorec pro rozptyl DX spojité náhodné veličiny: $DX = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx - (EX)^2$

Rozptyl DX náhodné veličiny X je tedy jakýsi teoretický rozptyl od střední hodnoty EX, který bychom vypočetli při velkém množství měření veličiny X, pokud by se veličina X chovala přesně podle daného teoretického popisu (popsaného hustotou f)

Jen s tím rozdílem, že místo sumy píšeme integrál

Ad př. 3: výpočet DX u tabule

Přesná definice DX diskrétní i spojité veličiny:

$DX = E\,(X\text{-}EX)^2$ … rozptyl je střední hodnota čtverce odchylek veličiny X od střední hodnoty = „teoretický průměr čtverců odchylek veličiny X od střední hodnoty EX“

- Výpočtem: $DX = E\,(X-EX)^2 = E(X^2 - 2X \cdot EX + (EX)^2) =$
$= EX^2 - 2 \cdot EX \cdot EX + (EX)^2 = EX^2 - (EX)^2$ … střední hodnota čtverců veličiny X minus čtverec střední hodnoty veličiny X

Konkrétní výpočet: $DX = EX^2 - (EX)^2$

Protože ovšem DX je jakýsi „teoretický“ průměr čtverců veličiny, nemá rozměr stejný jako veličina X

Abychom viděli rozptýlení hodnot analogické rozměru veličiny X, často upravujeme rozptyl do tvaru směrodatné odchylky = $\sqrt{DX}$

- Ad př.3: DX= 12, tj. $\sqrt{DX}$ = 3,4641 ... většina měření veličiny X leží v intervalu

  EX ± $\sqrt{DX}$ = ... 6 ± 3,4641 = (2,5359; 9,4641) ... většina měření hodnoty času u náhodně zastavené ručičky bude ležet v tomto intervalu

Dopočtení F(x), ED, DX pro příklad 4

Označme jej jako př. 4:

Př: mobily jisté firmy mají životnost v průměru 15 let; jaká je pst, že náhodně koupený mobil vydrží více než 10 let?

mohli bychom si označit X = životnost mobilu

Určete F(x), EX, DX veličiny X

Rekapitulace otázek:

18: distribuční funkce, střední hodnota a rozptyl spojité náhodné veličiny