4. cvičení z STP

11. března 2021

Příklad 4.1 *Do obchodu dodávají čipy tři výrobci, po řadě 50 %, 30 % a 20 % zásoby obchodu. Pravděpodobnosti výroby funkčního čipu od jednotlivých výrobců jsou po řadě $p_1 = 0.98$, $p_2 = 0.95$ a $p_3 = 0.99$.*

(a) Určete pravděpodobnosti, že zakoupený náhodně vybraný čip je vadný.

(b) Určete pravděpodobnost, že čip je od druhého výrobce, za předpokladu, že je funkční.

Řešení:
Označme si jevy:

A_i = *"zakoupený čip byl od i-tého výrobce"*,
B = *"zakoupený čip byl funkční"*.

Ze zadání plyne, že

$$P(A_1) = 0.5 \quad P(A_2) = 0.3 \quad P(A_3) = 0.2$$

$$P(B|A_1) = 0.98 \quad P(B|A_2) = 0.95 \quad P(B|A_3) = 0.99$$

(a) Počítáme pravděpodobnost jevu B^c. Z věty o úplné pravděpodobnosti máme:

$$P(B) = \sum_{i=1}^{3} P(B|A_i) \cdot P(A_i) = 0.98 \cdot 0.5 + 0.95 \cdot 0.3 + 0.99 \cdot 0.2 = 0.973$$

a

$$P(B^c) = 1 - P(B) = 1 - 0.973 = 0.027 \ .$$

Všimněme si ještě, že díky výpočtu $P(B)$ platí, že

$$0.95 = \min\{0.98, 0.95, 0.99\} \leq \underbrace{P(B)}_{0.973} \leq \max\{0.98, 0.95, 0.99\} = 0.99 \ .$$

Pokud nám tedy takovéto omezení nebude vycházet, někde jsme museli udělat chybu.

$P(B^c)$ jsme mohli spočítat i přímo jako

$$P(B^c) = \sum_{i=1}^{3} \underbrace{P(B^c|A_i)}_{1-P(B|A_i)} \cdot P(A_i) = 0.02 \cdot 0.5 + 0.05 \cdot 0.3 + 0.01 \cdot 0.2 = 0.027 \ .$$

(b) Zajímá nás $P(A_2|B)$. Z Bayesovy věty máme:

$$P(A_2|B) = \frac{P(B|A_2) \cdot P(A_2)}{P(B)} = \frac{0.95 \cdot 0.3}{0.973} = 0.293 \ .$$

Často se Bayesova věta používá pro úplný disjunktní systém jevů A a A^c, kde $0 < P(A) < 1$. Pak pro B máme

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)}$$

pokud $P(B) \neq 0$.

Příklad 4.2 *Požití alkoholu bylo prokázáno u 1% všech řidičů a u 10% řidičů, kteří způsobili dopravní nehodu. Kolikrát se požitím alkoholu zvyšuje riziko nehody?*

Řešení:
Označme si jevy

$A =$ *"požil alkohol"*,

$H =$ *"způsobil nehodu"*.

Pak máme $P(A) = 0.01$ a $P(A|H) = 0.1$. Hledáme hodnotu $\frac{P(H|A)}{P(H|A^c)}$. Tudíž podle Bayesovy věty máme

$$\frac{P(H|A)}{P(H|A^c)} = \frac{P(A|H) \cdot P(H)}{P(A)} \cdot \frac{P(A^c)}{P(A^c|H) \cdot P(H)} =$$
$$= \frac{P(A|H) \cdot (1 - P(A))}{P(A) \cdot (1 - P(A|H))} = \frac{0.1 \cdot 0.99}{0.01 \cdot 0.9} = 11\ .$$

Také na to můžeme jít následujícím způsobem:

$$0.1 = P(A|H) = \frac{P(H|A) \cdot P(A)}{P(H|A) \cdot P(A) + P(H|A^c) \cdot P(A^c)} = \frac{P(H|A) \cdot 0.01}{P(H|A) \cdot 0.01 + P(H|A^c) \cdot 0.99}$$

a odsud je

$$P(H|A) \cdot 0.001 + P(H|A^c) \cdot 0.099 = P(H|A) \cdot 0.01$$
$$P(H|A^c) \cdot 0.099 = P(H|A) \cdot 0.009\ .$$

opět s výsledkem

$$\frac{P(H|A)}{P(H|A^c)} = \frac{0.099}{0.009} = 11\ .$$

Co je náhodná veličina:
Náhodná veličina je funkce, která náhodnému výsledku (tj. elementárnímu jevu) přiřadí konkrétní hodnotu. Např. vybranému člověku přiřadí jeho tělesnou výšku. Jak je vidět, náhodná veličina vůbec není náhodná, co do hodnoty, kterou přiřazuje (ta je určena naprosto jasně). Náhodnost výstupu veličiny je dána náhodností jejího vstupu!

Abychom mohli s náhodnou veličinou X vůbec pracovat, potřebujeme umět určovat pravděpodobnosti toho, že hodnoty X budou např. v intervalu $\langle 1.5,\ 7.2) \subseteq \mathbb{R}$, což znamená, že množině $X^{-1}\langle 1.5,\ 7.2) = \{\omega \in \Omega \mid 1.5 \leq X(\omega) < 7.2\}$ musíme umět přiřadit její pravděpodobnost. To půjde jen tehdy, jestliže to bude množina měřitelná v daném systému jevů. Proto následující definice:

Definice náhodné veličiny: Náhodná veličina X v pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$ je takové zobrazení

$$X : \Omega \to \mathbb{R}$$

že vzor každého intervalu I v $\mathbb{R}$ je "přípustná" množina (neboli jev), tj. $X^{-1}(I) \in \mathcal{A}$. Tuto vlastnost stačí ověřit jen pro určité typy intervalů v $\mathbb{R}$:

$$X \text{ je náhodná veličina} \quad \Leftrightarrow \quad (\forall t \in \mathbb{R}) \;\; X^{-1}\Big((-\infty, t\rangle\Big) \in \mathcal{A}$$

Důležité vlastnosti distribuční funkce: Pro distribuční funkci F_X veličiny X platí, že

- má hodnoty v intervalu $\langle 0, 1\rangle$,
- je neklesající,
- je zprava spojitá,
- $\lim\limits_{t\to -\infty} F_X(t) = 0$ a $\lim\limits_{t\to +\infty} F_X(t) = 1$.

(Tyto vlastnosti dokonce už distribuční funkce úplně charakterizují v tom smyslu, že každá funkce splňující výše uvedené vlastnosti je distribuční funkcí nějaké náhodné veličiny.)

Co umožňuje náhodná veličina:
Náhodná veličina $X : \Omega \to \mathbb{R}$ umožňuje jednu podstatnou věc - zapomenout na původní prostor Ω všech výsledků (i s celou jeho strukturou jevů a jejich pravděpodobnostmi) a přitom stále umět vyčíslovat pravděpodobnosti pro X, které potřebujeme znát.

Veličina X totiž umí přenést a vytvořit rozdělení pravděpodobnosti tam, kde má své hodnoty - konkrétně vytvoří nový pravděpodobnostní prostor na reálné přímce $\mathbb{R}$. A to tak, že každý interval $I \subseteq \mathbb{R}$ bude mít prostě pravděpodobnost $P_X(I) := P\Big(X^{-1}(I)\Big)$.

Pravděpodobnosti P_X se říká rozdělení pravděpodobnosti veličiny X (na $\mathbb{R}$). Toto rozdělení můžeme úplně popsat pokud opět známe pravděpodobnosti jen některých speciálních intervalů: opět jsou to intervaly $(-\infty, t\rangle$ pro $t \in \mathbb{R}$. Jejich pravděpodobnosti pak přirozeně definují tzv. *distribuční funkci* $F_X : \mathbb{R} \to \langle 0, 1\rangle$ jako

$$F_X(t) := P(X \leq t) \qquad \Big(= P\big(X^{-1}(-\infty, t\rangle\big)\Big)$$

Náhodná veličina a její distribuční funkce jsou jedny z nejdůležitějších pojmů v celém kurzu! Proto se je snažte pochopit.

Příklad 4.3 *Uvažujme hod mincí s následujícími výsledky*

- $\omega_1 =$ *"padl rub" (s pravděpodobností* 0.49*)*
- $\omega_2 =$ *"padl líc" (s pravděpodobností* 0.49*)*
- $\omega_3 =$ *"nastala výjimečná situace" (hrana, zakutálení mince apod.)(s pravděpodobností* 0.02*).*

Sestrojte dvě různé náhodné veličiny a nakreslete jejich distribuční funkce.

Řešení:
Množina všech možných výsledků je tedy $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Každá náhodná veličina na tomto prostoru Ω má maximálně tři různé hodnoty, takže určitě bude *diskrétní*

(tj. má nejvýše spočetně mnoho hodnot $A = \{x_1, x_2, \dots\} \subseteq \mathbb{R}$ takových, že $\sum\limits_{u\in A} P(X = u) = 1$.)

Pro distribuční funkci diskrétní veličiny X pak platí

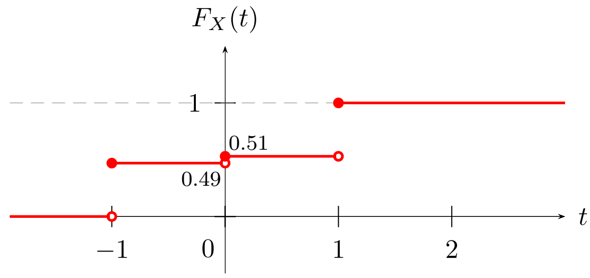
$$F_X(t) = P(X \leq t) = \sum_{u\leq t} P(X = u) \ .$$

Veličiny mohou být např.

- $X: \begin{array}{ccc} \omega_1 & \mapsto & 1 \\ \omega_2 & \mapsto & -1 \\ \omega_3 & \mapsto & 0 \end{array}$

  Distribuční funkce je skokovitá se skoky v bodech -1, 0 a 1 o velikostech 0.49, 0.02 a 0.49, tj.

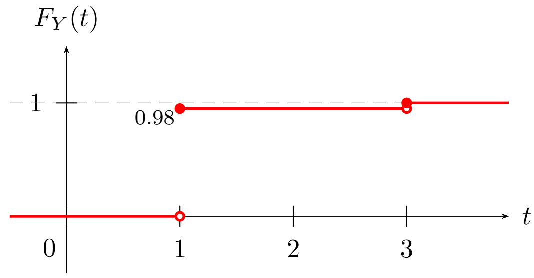
$$F_X(t) = \sum_{u \le t} P(X = u) = \left\{ \begin{array}{llll} & 0 & \text{pro} & t \in (-\infty, -1) \\ P(X=-1) = & 0.49 & \text{pro} & t \in \langle -1, 0) \\ P(X=-1) + P(X=0) = & 0.51 & \text{pro} & t \in \langle 0, 1) \\ P(X=-1) + P(X=0) + P(X=1) = & 1 & \text{pro} & t \in \langle 1, \infty) \end{array} \right.$$



- $Y: \begin{array}{ccc} \omega_1 & \mapsto & 1 \\ \omega_2 & \mapsto & 1 \\ \omega_3 & \mapsto & 3 \end{array}$

  Distribuční funkce je skokovitá se skoky v bodech 1 a 3 o velikostech 0.98 a 0.02 (protože obraz roven 1 mají dva elementární jevy ω_1 a ω_2, jejichž souhrnná pravděpodobnost je $0.49+0.49 = 0.98$), tj.

$$F_Y(t) = \sum_{u \le t} P(Y = u) = \left\{ \begin{array}{llll} & 0 & \text{pro} & t \in (\infty, 1) \\ P(Y=1) = P(\{\omega_1, \omega_2\}) = & 0.98 & \text{pro} & t \in \langle 1, 3) \\ P(Y=1) + P(Y=3) = & 1 & \text{pro} & t \in \langle 3, \infty) \ . \end{array} \right.$$



Co je to spojité rozdělení:
Veličina $X : \Omega \to \mathbb{R}$ má tzv. *spojité rozdělení* $\Leftrightarrow$ její distribuční funkce F_X je spojitá.

Často je v tomto případě F_X tzv. *absolutně spojitá* $\Leftrightarrow$ ex. $f_X : \mathbb{R} \to \langle 0, +\infty)$, která je integrabilní a platí

$$F_X(t) = \int_{-\infty}^{t} f_X(u) \, du$$

pro každé $t \in \mathbb{R}$.
Funkce f_X se nazývá *hustotou pravděpodobnosti* veličiny X.

Poznámka: Stojí za to ještě poznamenat, že vlastnost absolutní spojitosti není samozřejmá pro spojité distribuční funkce - příkladem je tzv. *Cantorova funkce* c (viz např. `https://en.wikipedia.org/wiki/Cantor_function`). Tato funkce (rozšířená z intervalu $\langle 0, 1\rangle$ přirozeně na celé $\mathbb{R}$) má nulovou derivaci c' až na (z hlediska integrálu) nepodstatnou množinu bodů. Přesto však Cantorova funkce není konstantní (a tudíž nelze získat integrováním své derivace - příslušné integrály $\int_{-\infty}^{t} c'(t)\ \mathrm{d}t = 0$ jsou všechny nulové). Pro veličinu s touto distribuční funkcí neexistuje hustota pravděpodobnosti.

Příklad 4.4 *Určete konstantu $c \in \mathbb{R}$ tak, aby funkce*

$$f(x) = \begin{cases} c \cdot xe^{-2x} & , x \in (0, 1) \\ 0 & , jinak. \end{cases}$$

byla hustotou nějaké náhodné veličiny.

Řešení:
Máme charakterizační větu:

Nezáporná integrabilní funkce $f : \mathbb{R} \to \mathbb{R}$ je hustotou pravděpodobnosti nějaké náhodné veličiny X právě když $\int\limits_{\mathbb{R}} f(x)\ dx = 1$.

Použitím integrace per partes dostaneme

$$1 = \int\limits_{-\infty}^{\infty} f(x)\ dx = \int\limits_{-\infty}^{0} 0\ dx + \int\limits_{0}^{1} c \cdot xe^{-2x}\ dx + \int\limits_{1}^{\infty} 0\ dx =$$

$$= c\left[x \cdot \frac{e^{-2x}}{-2}\right]_0^1 + c\int\limits_0^1 \frac{e^{-2x}}{2}\ dx = -\frac{c}{2}e^{-2} + c\left[\frac{e^{-2x}}{-4}\right]_0^1 = c \cdot \frac{1 - 3e^{-2}}{4},$$

a tudíž $c = \frac{4}{1-3e^{-2}}$. Protože $c > 0$, je splněna i nezápornost funkce f.

K významu hustoty pravděpodobnosti:

Nechť náhodná veličina X má hustotu pravděpodobnosti veličiny f_X, tj. $F_X(t) = \int\limits_{-\infty}^{t} f_X(u)\ \mathrm{d}u$ pro každé $t \in \mathbb{R}$. Pak pro interval $I \subseteq \mathbb{R}$ (nebo i nějakou množinu poskládanou z intervalů) platí, že

$$P(X \in I) = \int\limits_I f_X(u)\ \mathrm{d}u$$

tedy pravděpodobnost, že hodnoty veličiny X padnou do I, zjistíme prostě zintegrováním hustoty přes I (podobně zjišťujeme např. hmotnost nějaké křivky, když zintegrujeme hustotu (hmotnosti) přes danou křivku).

Poznamenejme ještě důležitou věc a sice, že hustota f_X NENÍ zdaleka určena jednoznačně jako funkce (např. její změnou v konečně mnoha bodech se nezmění příslušné integrály, takže i změněná funkce bude také hustotou).

Co je to střední hodnota náhodné veličiny:

Jestliže pro veličinu X provedeme n měření, ve kterých se budou vyskytovat číselné hodnoty $a_1, \ldots, a_k$ (ne nutně navzájem různé), kde každá hodnota a_i se bude opakovat n_i-krát, pak za jejich průměrnou hodnotu $\overline{\boldsymbol{x}}$ (v rámci tohoto našeho měření) budeme považovat jejich aritmetický průměr (kde se započítá i opakování dané hodnoty a_i), tj.

$$\overline{\boldsymbol{x}} = \frac{\sum_i n_i \cdot a_i}{\sum_i n_i} = \sum_i a_i \cdot \frac{n_i}{n}$$

kde $n = \sum_i n_i$ je celkový počet měření. (Ve statistice se tato hodnota $\overline{\boldsymbol{x}}$ nazývá *výběrový průměr* - viz později). Výraz $\frac{n_i}{n}$ přitom vyjadřuje podíl výskytu hodnoty a_i při daném počtu měření. Uvědomme si ještě, že k výpočtu $\overline{\boldsymbol{x}}$ musíme mít k dispozici konkrétní soubor naměřených hodnot!

Tento intuitivní přístup vede přirozeně k definici *střední hodnoty* $E(X)$ pro *diskrétní* veličinu X (tj. pro takovou veličinu X, že existuje *nejvýše spočetná* množina $A \subseteq \mathbb{R}$, že $P(X \in A) = 1$) v podobě

$$E(X) = \sum_{k \in A} k \cdot P(X = k)$$

samozřejmě pokud suma vůbec existuje. Spočetnost nebo konečnost množiny A tu je ovšem podstatná - jestliže sčítáme nespočetně mnoho nenulových čísel, pak tento součet buď neexistuje nebo bude vždy nekonečný. Tedy šanci na to, abychom dostali konečně číslo mají jen nejvýše spočetně součty nenulových čísel. No a protože pro hodnoty $k \in \mathbb{R} \setminus A$ je vždy $P(X = k) = 0$ (pro naši diskrétní veličinu X), tak můžeme použít i vyjádření ve tvaru:

$$E(X) = \sum_{k \in \mathbb{R}} k \cdot P(X = k) \ .$$

Střední hodnotu pro náhodnou veličinu X se *spojitým* rozdělením a s hustotou pravděpodobnosti f_X si definujeme analogicky jako pro diskrétní případ

$$E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t) \ \mathrm{d}t$$

(pokud tento integrál existuje.)

Názornou interpretací střední hodnoty $E(X)$ pak je, že hodnota $E(X)$ je vodorovná souřadnice těžiště plochy, která je určena grafem hustoty f_X (a vodorovnou osou).

Poznámka: Ve všech případech (diskrétním i spojitém) platí pro střední hodnotu následující vztah

$$E(X) = -\int_{-\infty}^{0} F_X(t) \ dt + \int_{0}^{\infty} \big(1 - F_X(t)\big) \ dt \ .$$

Co je rozptyl náhodné veličiny:

Rozptyl $\mathrm{var}(X)$ (libovolné) náhodné veličiny X určuje, jak moc se hodnoty odchylují od střední hodnoty. Rozptyl je proto definován jako střední hodnota veličiny $\big(X - E(X)\big)^2$, která vyjadřuje kvadrát odchylky X od své střední hodnoty, tedy předpisem:

$$\mathrm{var}(X) = E\Big(\big(X - E(X)\big)^2\Big) = \cdots = E\big(X^2\big) - \Big(E\big(X\big)\Big)^2$$

(ovšem opět jen pokud uvedené střední hodnoty existují.)

Věta: Pro (borelovsky) měřitelnou funkci $h : \mathbb{R} \to \mathbb{R}$ (v praxi to je téměř libovolná funkce, např. po částech spojitá) a diskrétní veličinu X platí, že

$$E\big(h(X)\big) = \sum_{k \in \mathbb{R}} h(k) \cdot P(X = k) \ .$$

a analogický vzorec platí pro veličinu X s spojitým rozdělením a hustotou pravděpodobnosti f_X:

$$E\big(h(X)\big) = \int_{-\infty}^{\infty} h(t) \cdot f_X(t) \ \mathrm{d}t \ .$$

Příklad 4.5 *Mějme funkci*

$$f(x) = \begin{cases} 3e^{-3x} & , \ x \in \langle 0, \infty) \\ 0 & , \ jinak. \end{cases}$$

(a) Ověřte, že f je hustota pravděpodobnosti nějaké náhodné veličiny X.

(b) Určete distribuční funkci F_X veličiny X příslušnou této hustotě.

(c) Spočtěte pravděpodobnost $P(-1 \le X \le 1)$.

(d) Spočtěte střední hodnotu $E(X)$ a rozptyl $\operatorname{var}(X)$.

Řešení:

(a) Funkce f je nezáporná a platí, že

$$\int_{-\infty}^{\infty} f(x)\ dx = \int_{-\infty}^{0} 0\ dx + \int_{0}^{\infty} 3e^{-3x}\ dx = \left[e^{-3x}\right]_0^\infty = 1,$$

tudíž vlastnost $\int_{\mathbb{R}} f(x)\ dx = 1$ je splněna a f tedy je hustotou nějaké veličiny X. Graf hustoty f viz níže.

(b) Příslušná distribuční funkce F_X pro veličinu X je

pro $t \in (0, \infty)$: $F_X(t) = \int_{-\infty}^{t} f(x)\ dx = \int_{-\infty}^{0} 0\ dx + \int_{0}^{t} 3e^{-3x}\ dx = \left[e^{-3x}\right]_0^t = 1 - e^{-3t}$.

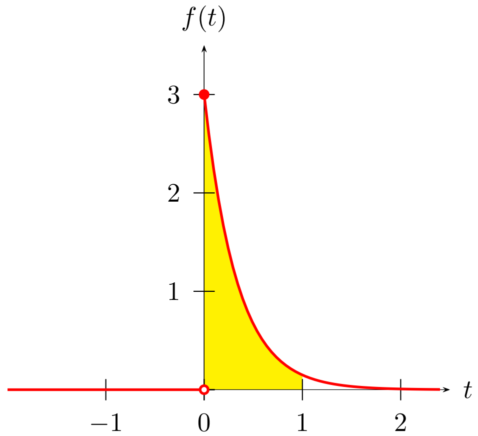
pro $t \in (-\infty, 0\rangle$: $F_X(t) = \int_{-\infty}^{t} 0\ dx = 0$.

Graf distribuční funkce F_X viz níže.

(c) Hledaná pravděpodobnost je

$$P(-1 \le X \le 1) = \int_{-1}^{1} f(x)dx = \int_0^1 3e^{-3x}dx = 1 - e^{-3}.$$

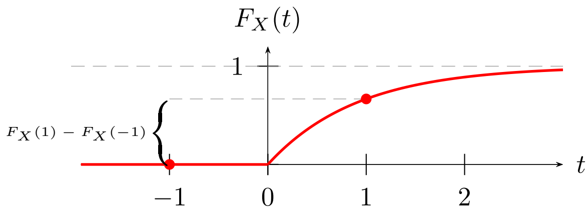
Geometrická interpretace této hodnoty je plocha pod grafem hustoty v intervalu $\langle -1, 1\rangle$:



Lze využít také distribuční funkce:

$$P(-1 \le X \le 1) = \underbrace{P(X \le 1)}_{F_X(1)} - \underbrace{P(X < -1)}_{\substack{\text{limita zleva } F_X(t) \\ \text{v bodě } -1}} \overset{(\text{spojitost } F_X)}{=} F_X(1) - F_X(-1) = 1 - e^{-3\cdot 1} - 0 = 1 - e^{-3}.$$

Geometrická interpretace v tomto případě je rozdíl funkčních hodnot distribuční funkce:



(d) Použitím integrace per partes dostaneme

$$E(X) = \int_{-\infty}^{\infty} x f(x)\ dx = \int_{0}^{\infty} 3xe^{-3x} dx = \frac{1}{3},$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\ dx = \int_{0}^{\infty} 3x^2 e^{-3x} dx = \frac{2}{9}.$$

Střední hodnota je tudíž $E(X) = 1/3$ a rozptyl

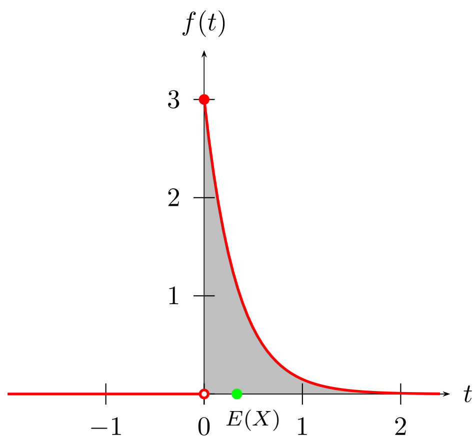
$$\mathrm{var}(X) = E\Big(\big(X - E(X)\big)^2\Big) = E\big(X^2\big) - \big(E(X)\big)^2 = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}.$$

Poznamenejme ještě, že veličina X má tzv. *exponenciální* rozdělení, tj.

$$F_X(t) = \begin{cases} 0 & , \ t < 0 \\ 1 - e^{-\frac{t}{\tau}} & , \ t \geq 0 \end{cases}$$

kde $\tau > 0$ je parametr, jehož význam je, že $E(X) = \tau$. Dále ještě platí, že $\mathrm{var}(X) = \tau^2$.

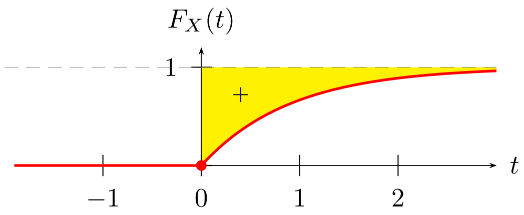
Geometrická interpretace hodnoty $E(X)$ - je to vodorovná souřadnice (zelený bod) těžiště plochy pod grafem hustoty (šedá plocha):



Ještě jedna geometrická interpretace hodnoty $E(X)$ - je to rozdíl velikosti plochy *nad* grafem distr. funkce F_X *v intervalu* $(0, +\infty)$ a velikosti plochy *pod* grafem distr. funkce F_X *v intervalu* $(-\infty, 0)$. Konkrétně:

$$E(X) = -\int\limits_{-\infty}^{0} F_X(t)\ dt + \int\limits_{0}^{\infty} \big(1 - F_X(t)\big)\ dt$$

Tento vztah platí dokonce pro jakoukoliv náhodnou veličinu X (diskrétní, spojitou i smíšenou).



V tomto případě je velikost plochy pod grafem F_X na intervalu $(-\infty, 0)$ nulová, takže v obrázku nejde zvýraznit jako plocha se záporným znaménkem.