6. cvičení z STP

23. - 27. březen 2020

Připomenutí: Nechť $a \in \mathbb{R}$ je bod *spojitosti distribuční funkce* F_X náhodné veličiny X. Pak máme

$$P(X = a) = P(X \leq a) - P(X < a) = F_X(a) - \lim_{x \to a_-} F_X(t) = 0$$

a tedy

$$P(X \leq a) = P(X < a)$$

neboli v bodech *spojitosti* nezáleží na typu nerovnosti (neostré vs. ostré).

Příklad 6.1 *Při numerickém výpočtu se reálná čísla zaokrouhlují na jedno desetinné místo. Jaká je pravděpodobnost, že vzdálenost skutečného čísla od zaokrouhleného bude větší než* 0.04*?*

Řešení:
Podobně jako v příkladu o házení mince na nekonečnou mřížku (**Příklad 2.2**) budeme předpokládat, že vstupní hodnoty (tj. čísla, která budeme zaokrouhlovat) pocházejí z nějakého referenčního intervalu $\Omega \subseteq \mathbb{R}$ délky 0.1 (zaokrouhlujeme na jedno desetinné místo), na kterém máme geometrickou pravděpodobnost, např. $\Omega = \langle 0,\ 0.1 \rangle$.

Označme si teď náhodnou veličinu $X : \Omega \to \mathbb{R}$ jako

$$X = \text{“\textit{skutečná hodnota}” - “\textit{zaokrouhlená hodnota}”}$$

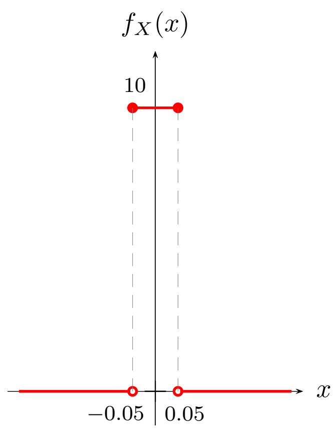
tj. pro $\omega \in \Omega = \langle 0,\ 0.1 \rangle$ to funguje jako

$$X(\omega) = \begin{cases} \omega & \text{pro} \quad 0 \leq \omega < 0.05, \\ \omega - 0.1 & \text{pro} \quad 0.05 \leq \omega \leq 0.1\ . \end{cases}$$

A nás teď zajímá pravděpodobnost $P\big(|X| > 0.04\big)$. K tomu potřebujeme znát rozdělení veličiny X. Intuitivně tušíme, že X bude mít rovnoměrné rozdělení na své množině hodnot, tj. v intervalu $\langle -0.05,\ 0.05)$. Předpokládejme tedy, že $X \sim \mathrm{Ro}(-0.05,\ 0.05)$ (níže si to pak zdůvodníme). Veličina X má pak spojité rozdělení s hustotou pravděpodobnosti

$$f_X(x) = \begin{cases} \frac{1}{0.05-(-0.05)} = 10 & , \ -0.05 \leq x \leq 0.05 \\ 0 & , \ \text{jinak} \end{cases}$$

s grafem



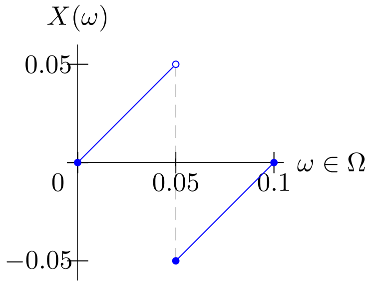
(Hodnoty f_X v krajních bodech intervalu nejsou podstatné.)

Hledaná pravděpodobnost je tudíž

$$P\Big(|X| > 0.04\Big) = 1 - P\Big(|X| \leq 0.04\Big) = 1 - P(-0.04 \leq X \leq 0.04) =$$
$$= 1 - \int_{-0.04}^{0.04} 10\; dx = 1 - 10 \cdot 0.08 = 0.2\ .$$

- -

Z cvičných důvodu si teď výše zmíněné rovnoměrné rozdělení odvodíme. Veličinu X si můžeme znázornit tímto grafem:



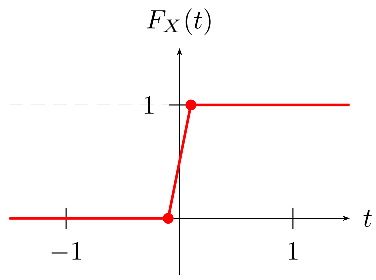
Odsud snadno vidíme, že pro $t \in \langle -0.05,\ 0.05)$ je

$$vol\Big(\{\omega \in \Omega \mid X(\omega) \leq t\}\Big) = \begin{cases} vol\Big(\langle 0, t\rangle \cup \langle 0.05,\ 0.1\rangle\Big) = t + 0.05 & ,\ 0 \leq t < 0.05 \\ vol\Big(\langle 0.05,\ 0.1 + t\rangle\Big) = t + 0.05 & ,\ -0.05 \leq t < 0. \end{cases}$$

takže distribuční funkce je

$$F_X(t) = P(X \leq t) = \begin{cases} 0 & , t < -0.05 \\ \frac{vol\{\omega\in\Omega \mid X(\omega)\leq t\}}{vol(\Omega)} = \frac{t+0.05}{0.1} = 10t + 0.5 & , t \in \langle -0.05,\ 0.05) \\ 1 & , t \geq 0.05 \end{cases}$$

s grafem



Distribuční funkce F_X je tedy spojitá a podle věty o existenci hustoty (viz poznámky v 5.cvičení) vidíme, že funkce

$$f(t) = \begin{cases} (F_X)'(t) & \text{pokud tato derivace v bodě } t \text{ existuje a je konečná} \\ 0 & \text{jinak} \end{cases}$$

která odpovídá funkci f_X (až na nepodstatnou množinu bodů), je skutečně hustota X.

Poznámka: Pro ověření toho, že výše uvedená funkce f_X je skutečně hustota, tj. že $F_X(t) = \int\limits_{-\infty}^{t} f_X(x)\ dx$, se nemusíme opírat o větu o existenci, ale můžeme si to také snadno ověřit sami. Pro $t \in \langle -0.05,\ 0.05\rangle$ máme

$$\int\limits_{-\infty}^{t} f_X(x)\ dx = \int\limits_{-0.05}^{t} 10\ dx = \big[10x\big]_{-0.05}^{t} = 10t + 0.5 = F_X(t)$$

a zbytek je zřejmý.

Připomenutí: 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).

Pro (borelovsky) měřitelnou funkci $h : \mathbb{R} \to \mathbb{R}$ (tj. v praxi "téměř" každou, např. po částech spojitou) pak opět máme vzorec analogický tomu pro diskrétní rozdělení:

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

Příklad 6.2 *Prodejna očekává dodávku nového zboží v určitý den buď dopoledne v době od* 8 *do* 10 *hodin nebo odpoledne od* 14 *do* 15 *hodin. Uskutečnění dodávky je stejně možné kdykoliv během těchto časových úseků. Veličina X bude představovat čas dodání zboží.*

(a) Určete rozdělení veličiny X a její distribuční funkci F_X.

(b) Jaká je pravděpodobnost, že zboží bude dodáno v době od 8:30 *do* 8:45*? Jaká bude pravděpodobnost, že zboží bude dodáno až po* 9:00*?*

(c) Průměrně v kolik hodin bude zboží dodáno?

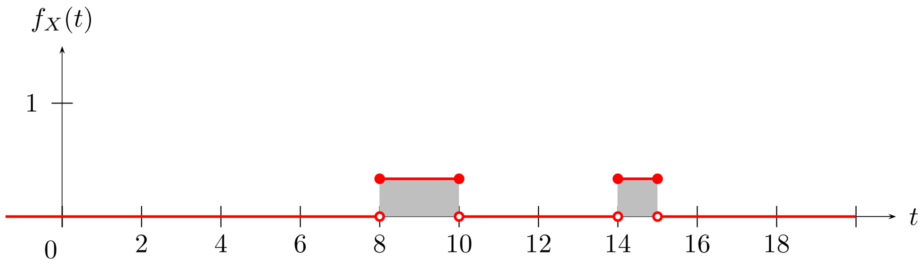
Řešení:
(a) Veličina X (měřená v hodinách) má rovnoměrné rozdělení na sjednocení intervalů $\langle 8, 10\rangle$ a $\langle 14, 15\rangle$. Její hustota f_X je tedy dána jako

$$f_X(t) = \begin{cases} c & , t \in \langle 8,\ 10\rangle \cup \langle 14,\ 15\rangle \\ 0 & , \text{jinak} \end{cases}$$

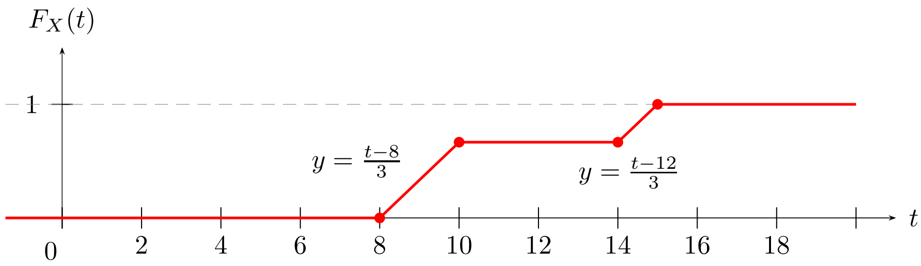
pro vhodné c, které určíme z podmínky, že (šedá) plocha pod hustotou je rovna 1:

$$1 = P(X \in \mathbb{R}) = \int_{-\infty}^{\infty} f_X(t)\ \mathrm{d}t = \int_{8}^{10} c\ \mathrm{d}t + \int_{14}^{15} c\ \mathrm{d}t = 3c$$

tedy $c = \frac{1}{3}$.



Funkci F_X nyní už snadno napočítáme pomocí integrování hustoty f_X (anebo ještě jednodušeji z obrázku hustoty: F_X poroste s lineárním přírůstkem o směrnici $\frac{1}{3}$ tam, kde je nenulová hustota, a jinde bude funkce F_X konstantní):



$$F_X(t) = \int_{-\infty}^{t} f_X(u)\ \mathrm{d}u = \begin{cases} \int_{-\infty}^{t} 0\ \mathrm{d}u = \quad 0 & , t \leq 8 \\ \int_{8}^{t} \frac{1}{3}\ \mathrm{d}u = \quad \frac{t-8}{3} & , t \in \langle 8,\, 10 \rangle \\ \frac{2}{3} + \int_{10}^{t} 0\ \mathrm{d}u = \quad \frac{2}{3} & , t \in \langle 10,\, 14 \rangle \\ \frac{2}{3} + \int_{14}^{t} \frac{1}{3}\ \mathrm{d}u = \quad \frac{t-12}{3} & , t \in \langle 14, 15 \rangle \\ \quad 1 & , t \geq 15 \end{cases}$$

(b) Zajímají nás pravděpodobnosti $P\big(8.5 \leq X \leq 8.75\big)$ a $P\big(X \geq 9\big)$. Ty můžeme zjistit buď jednoduše z geometrické pravděpodobnosti (která je jen jiným vyjádřením rovnoměrného rozdělení):

$$P\big(8.5 \leq X \leq 8.75\big) = \frac{vol\big(\langle 8.5,\, 8.75 \rangle\big)}{vol\big(\langle 8,\, 10 \rangle \cup \langle 14,\, 15 \rangle\big)} = \frac{0.25}{2+1} = \frac{1}{12} \doteq 0.083$$

$$P\big(X \geq 9\big) = \frac{vol\big(\langle 9,\, 10 \rangle \cup \langle 14,\, 15 \rangle\big)}{vol\big(\langle 8,\, 10 \rangle \cup \langle 14,\, 15 \rangle\big)} = \frac{1+1}{2+1} = \frac{2}{3} \doteq 0.667$$

nebo pomocí hustoty f_X a (ještě lépe) pomocí distribuční funkce F_X, kterou už máme určenou:

$$P\big(8.5 \leq X \leq 8.75\big) = P(X \leq 8.75) - P(X < 8.5) \stackrel{\text{spoj. rozd.}}{=} P(X \leq 8.75) - P(X \leq 8.5) =$$

$$= F_X(8.75) - F_X(8.5) = \frac{8.75-8}{3} - \frac{8.5-8}{3} = \frac{1}{12}$$

$$P\big(X \geq 9\big) = 1 - P(X < 9) \overset{\text{spoj. rozd.}}{=} 1 - P(X \leq 9) = 1 - \frac{9-8}{3} = \frac{2}{3} \ .$$

(c) Zde chceme zjistit $E(X)$. Máme

$$E\big(X\big) = \int_{-\infty}^{\infty} t \cdot f_X(t) \ \mathrm{d}t = \int_{8}^{10} t \cdot \tfrac{1}{3} \ \mathrm{d}t + \int_{14}^{15} t \cdot \tfrac{1}{3} \ \mathrm{d}t =$$

$$= \left[\tfrac{t^2}{6}\right]_8^{10} + \left[\tfrac{t^2}{6}\right]_{14}^{15} = \frac{100 - 64 + 225 - 196}{6} = \frac{65}{6} = 10 \, \text{hod} \, 50 \, \text{min} \ .$$

Průměrná doba, kdy bude zboží dodáno tedy je v 10:50. To se může zdát na první pohled divné, když v tento čas přece zboží vůbec nemůže být (fyzicky) dodáno, ale tento výsledek jen vyjadřuje, že střední čas dodání bude blíž k dopolednímu rozmezí než k odpolednímu (a jak moc blízko to bude). Podobně, u alternativního rozdělení $\mathrm{Alt}(p)$ je střední hodnota také rovna $p \in (0,1)$, přestože hodnoty alternativního rozdělení jsou jen 0 a 1.

Příklad 6.3 *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 $P(-1 \leq X \leq 1)$.

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

Řešení:

(a) Funkce f je nezáporná a platí, že

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

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 $t \in (0, \infty)$: $F_X(t) = \int\limits_{-\infty}^{t} f(x) \ dx = \int\limits_{-\infty}^{0} 0 \ dx + \int\limits_{0}^{t} 3e^{-3x} \ dx = \big[e^{-3x}\big]_0^t = 1 - e^{-3t}$.

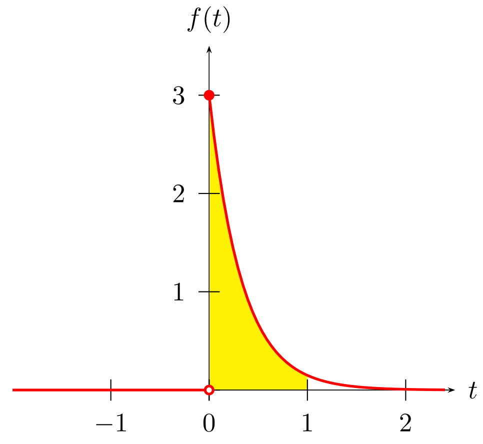
pro $t \in (-\infty, 0\rangle$: $F_X(t) = \int\limits_{-\infty}^{t} 0 \ dx = 0$.

Graf distribuční funkce F_X viz níže.

(c) Hledaná pravděpodobnost je

$$P(-1 \leq X \leq 1) = \int_{-1}^{1} f(x) dx = \int_{0}^{1} 3e^{-3x} dx = 1 - e^{-3}.$$

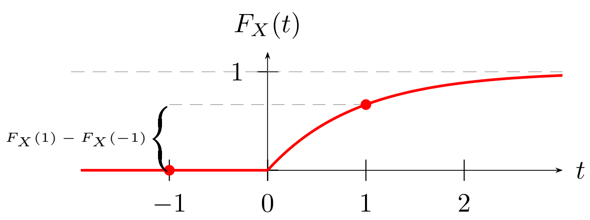
Geometrická interpretace této hodnoty je plocha pod grafem hustoty v intervalu $\langle -1, 1\rangle$:



Lze využít také distribuční funkce:

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

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

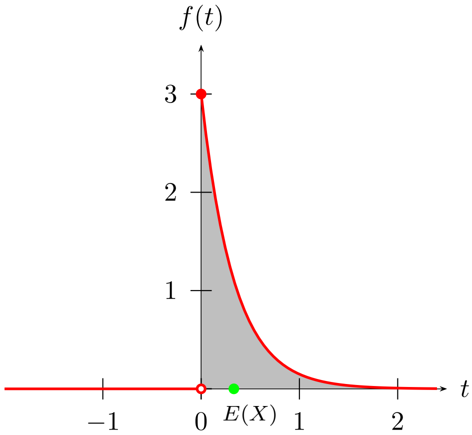
$$\operatorname{var}(X) = E\Big(\big(X - E(X)\big)^2\Big) = E\big(X^2\big) - \big(E(X)\big)^2 = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}.$$

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 $\mathrm{var}(X) = \tau^2$.

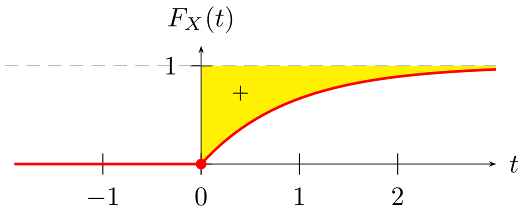
Geometrická interpretace hodnoty $E(X)$ - je to vodorovná souřadnice (zelený bod) těžiště plochy pod grafem hustoty (šedá plocha):



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:

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

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.

Poznámky k Poissonovu rozdělení: Pro veličinu

$$X = \text{„} \textit{počet událostí během intervalu délky } T \text{''}$$

kde interval je obvykle časový (ale může být i délkový nebo měřený nějakou jinou jednotkou), můžeme předpokládat *Poissonovo* rozdělení

$$P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \quad \text{kde} \quad k = 0, 1, 2, \dots$$

s (bezrozměrným) parametrem $\lambda > 0$, pokud jsou splněny následující podmínky

- počet událostí může nabývat libovolných (konečných) hodnot.
- jednotlivé události jsou nezávislé a nenastávají současně (lze je časově oddělit),
- *průměrný* počet událostí v libovolném časovém podintervalu je úměrný pouze časové délce tohoto podintervalu a ne jeho umístění v původním intervalu (tj. lze říct, že střední četnosti událostí za jednotku času se s průběhem doby nemění),

V praxi půjde např. o příchod zákazníka do fronty, chytání ryb, průjezd aut atd. a to během nějaké předem určené doby.

Parametr λ pak představuje střední hodnotu (tj. $E(X) = \lambda$) protože:

$$E(X) = \sum_{k=0}^{\infty} k \cdot P(X = k) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} \cdot e^{-\lambda} = \lambda e^{-\lambda} \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{=e^{\lambda}} = \lambda$$

Poissonovo rozdělení je většinou spíše limitní případ a používá se jako aproximace binomického rozdělení $\mathrm{Bi}(n, p)$, u kterého sice neznáme n a p, ale víme, že n je (dostatečně) velké a známe střední hodnotu dané veličiny (viz níže).

V praxi tedy můžeme podmínky Poissonova rozdělení přibližně zajistit pokud budou události pocházet z velkého počtu nezávislých zdrojů (z každého jen jednou) a podmínky se během měření nebudou měnit (tj. nebude se náhle měnit "okamžitá střední četnost události").

Poznámka: Ke tvaru Poissonova rozdělení se můžeme dostat pomocí binomického rozdělení (s využitím výše uvedených předpokladů) takto:

Časový interval délky T si rozdělíme na n dílků, a budeme předpokládat, že v každém se může stát maximálně jedna událost se stejnou pravděpodobností p_n. Dostaneme tak binomické rozdělení veličiny

$$X_n = \text{\textit{"počet událostí v daném časovém úseku délky } T \text{ \textit{rozděleném na} } n \text{ \textit{dílků"}}}$$

se střední hodnotou $\lambda = E(X_n) = n \cdot p_n$, kterou si vezmeme jako pevnou (neboli vlastně položíme $p_n := \frac{\lambda}{n}$). Tedy X_n má binomické rozdělení $\mathrm{Bi}(n, p_n)$. Spočítáme si teď limitu (pro pevně zvolené k):

$$\lim_{n\to\infty} P(X_n = k) = \lim_{n\to\infty} \binom{n}{k} (p_n)^k (1 - p_n)^{n-k} = \lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\to\, e^{-\lambda}} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\to\, 1} \underbrace{\prod_{i=0}^{k-1} \left(\frac{n-i}{n}\right)}_{\to\, 1} = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \ .$$

Pro lepší pochopení si ještě uvědomme, že (při naší diskretizaci času T na n dílků) pro $n \to \infty$ musí hodnota p_n nutně jít k nule, protože představuje pravděpodobnost, že událost nastane během krátkého časového úseku Δt délky $\frac{T}{n}$. Dále, počet událostí v tomto časovém úseku Δt popisuje (při naší diskretizaci) alternativní rozdělení $\mathrm{Alt}(p_n)$. Vidíme tedy, že podíl

$$\frac{\text{střední počet událostí v časovém intervalu } \Delta t}{\text{délka časového intervalu } \Delta t} = \frac{E(\mathrm{Alt}(p_n))}{\frac{T}{n}} = \frac{p_n}{\frac{T}{n}} = \frac{\lambda}{T}$$

(který bychom mohli nazvat "okamžitou střední četností událostí") zůstává vlastně konstatní v každém časovém okamžiku.

Aproximace binomického rozdělení pomocí Poissonova: Hodnotu $P(Y = k)$ pro $Y \sim \mathrm{Bi}(n, p)$ můžeme aproximovat pomocí Poissonova rozdělení jako

$$P(Y = k) \doteq \frac{\lambda^k}{k!} e^{-\lambda}$$

když je $k << n$ a $\lambda := E(Y) << n$.

Příklad 6.4 *Na látce (pevné šířky) je průměrně jeden kaz na* 10 m *délky. Předpokládáme, že počet kazů se řídí Poissonovým rozdělením. Jaká je pravděpodobnost, že na* 50 m *délky látky bude*

(a) přesně 10 *kazů,*

(b) maximálně 3 *kazy,*

(c) přesně 5 *kazů, z toho* 4 *na prvních* 20 m*?*

Řešení:
Označme si náhodnou veličinu

$$X = \text{“ \textit{počet kazů na} 50 m \textit{délky látky}”}$$

Pak $X \sim \text{Poiss}(\lambda)$ a pravděpodobnosti hodnot jsou

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{pro} \quad k = 0, 1, \ldots$$

kde $E(X) = \lambda$. Pro veličinu $\widetilde{X}$, označující počet kazů na 10 m délky, předpokládáme také Poissonovo rozdělení se střední hodnotou $E(\widetilde{X}) = 1$. Protože střední hodnota má být úměrná délce intervalu dostaneme

$$\frac{E(X)}{E(\widetilde{X})} = \frac{50\,\text{m}}{10\,\text{m}} = 5 \quad \Rightarrow \quad \lambda = E(X) = 5$$

Tudíž

(a) $P(X = 10) = \frac{5^{10}}{10!}\, e^{-5} \doteq 0.0181.$

(b) $P(X \leq 3) = e^{-5} \left(\frac{5^0}{0!} + \frac{5^1}{1!} + \frac{5^2}{2!} + \frac{5^3}{3!} \right) = e^{-5}(1 + 5 + \frac{25}{2} + \frac{125}{6}) = \frac{118}{3} e^{-5} \doteq 0.265.$

(c) Označme si veličiny

$$X_1 = \text{“ \textit{počet kazů na prvních} 20 m \textit{délky látky}”}$$

$$X_2 = \text{“ \textit{počet kazů na zbylých} 30 m \textit{délky látky}” .}$$

Ty budou nezávislé. Analogicky k předešlému budeme mít

$$P(X_1 = k) = \frac{\lambda_1^k}{k!} e^{-\lambda_1} \qquad \text{a} \qquad P(X_2 = k) = \frac{\lambda_2^k}{k!} e^{-\lambda_2},$$

kde $\lambda_1 = 2$ a $\lambda_2 = 3$. Hledaná pravděpodobnost je (díky nezávislosti X_1 a X_2) tedy

$$P(X_1 = 4,\ X_2 = 1) \overset{(\text{nezáv. } X_i)}{=} P(X_1 = 4) \cdot P(X_2 = 1) = \frac{2^4}{4!} e^{-2} \cdot \frac{3^1}{1!} e^{-3} = 2e^{-5} \doteq 0.01348.$$

Poznámky k exponenciálnímu rozdělení:
Exponenciální rozdělení popisuje pravděpodobnost veličiny

$$Y = \text{“\textit{doba mezi dvěma následnými výskyty událostí}”},$$

v systému, který nemá paměť na předchozí události. Tedy to, co se stane od určitého okamžiku, nezávisí na tom, co bylo předtím. V praxi jde např. o dobu, za kterou se porouchá zařízení, které se "neopotřebovává" (např.

polovodičové součástky), nebo o dobu radioaktivního rozpadu atd. Exponenciální rozdělení je jediné, které splňuje následující rovnici:

$$P(Y > s + t | Y > s) = P(Y > t)$$

pro všechna $s, t > 0$. Rovnice vyjadřuje to, že pravděpodobnost, že zařízení bude bez poruchy pracovat alespoň t hodin, je stejná v případě, že jsme jej právě zapnuli (pravá strana rovnice), jako za předpokladu, že předtím už bez poruchy pracovalo s hodin (levá strana rovnice).

Exponenciální rozdělení $\mathrm{Exp}(\frac{1}{\tau})$ je charakterizováno parametrem $\tau > 0$ (s fyzikálním rozměrem času), který představuje střední dobu čekání, tedy $E(Y) = \tau$ a dále ještě platí $\mathrm{var}(Y) = \tau^2$. Distribuční funkce pro $Y \sim \mathrm{Exp}(\frac{1}{\tau})$ je

$$F_Y(t) = \begin{cases} 0 & , \ t \leq 0 \\ 1 - e^{-\frac{t}{\tau}} & , \ t > 0 \ . \end{cases}$$

Doplnění: Diskrétní analogií exponenciálního rozdělení je *geometrické* rozdělení, které neměří čas spojitě ale pouze diskrétně. Jak už víme, je to rozdělení veličiny

$$\widetilde{Y} = \text{"}\textit{počet neúspěšných pokusů než nastane první úspěch}\text{"},$$

např. v situaci, že se chceme trefit míčem do koše atd. Hodnoty $\widetilde{Y}$ jsou $\{0, 1, 2, \dots\}$. Je to opět jediné takové diskrétní rozdělení splňující analogickou rovnici:

$$P(\widetilde{Y} > k + n | \widetilde{Y} > n) = P(\widetilde{Y} > k)$$

pro všechna $k, n \in \mathbb{N}_0$ s podobným významem jako u exponenciálního rozdělení.

Souvislost mezi exponenciálním a Poissonovým rozdělením:
Nechť

$Y =$ *"doba čekání na událost"*

je veličina s exponenciálním rozdělením. Pak veličina

$X =$ *"počet událostí během doby T"*

má Poissonovo rozdělení a platí

$$E(Y) = \frac{T}{E(X)}$$

kde doba T je vyjádřená ve stejných jednotkách, jaké má veličina Y. Neboli

$$\text{"střední doba čekání"} = \frac{\text{"délka intervalu"}}{\text{"střední počet událostí v tomto intervalu"}} \ .$$

Příklad 6.5 *Na zákaznickou linku přichází průměrně 12 hovorů za hodinu. Doba čekání na hovor má exponenciální rozdělení.*

(a) *Jaká je pravděpodobnost, že nejbližší hovor přijde nejdříve za 10 minut?*

(b) *Určete čas t takový, že nejbližší hovor přijde nejdříve za t minut s pravděpodobností 0.7.*

Řešení:
Tuto úlohu můžeme řešit jak s využitím exponenciálního rozdělení, tak Poissonova rozdělení.

(a)

- *Pomocí exponenciálního:* Podle věty má veličina

$$X = \text{“}\textit{počet hovorů během doby } 60 \textit{ min}\text{”}$$

Poissonovo rozdělení s parametrem $\lambda = E(X) = 12$. Tudíž náhodná veličina

$$Y = \text{“}\textit{doba čekání na hovor}\text{”} \ [\textit{v minutách}]$$

má exponenciální rozdělení se střední hodnotou $\tau = E(Y) = \frac{60 \ min}{12} = 5 \ min$. Její distribuční funkce je tedy

$$F_Y(t) = \begin{cases} 1 - e^{-t/5} & , \ t > 0 \\ 0 & , \ t \leq 0. \end{cases}$$

Hledaná pravděpodobnost pak je

$$P(Y > 10 \ min) = 1 - P(Y \leq 10 \ min) = 1 - F_Y(10) = 1 - \left(1 - e^{-\frac{10}{5}}\right) = e^{-2} \doteq 0.1353 \ .$$

Pravděpodobnost také můžeme spočítat pomocí hustoty f_Y veličiny Y:

$$f_Y(t) = \begin{cases} \frac{1}{5} e^{-t/5} & , \ t > 0 \\ 0 & , \ t \leq 0. \end{cases}$$

jako

$$P(Y > 10 \ min) = \int_{10}^{\infty} f_Y(t) \ dt = \int_{10}^{\infty} \tfrac{1}{5} e^{-t/5} \ dt = e^{-2} \ .$$

- *Pomocí Poissonova:* Podle věty opět víme, že veličina

$$X' = \text{“}\textit{počet hovorů během doby } 10 \textit{ min}\text{”}$$

má Poissonovo rozdělení s parametrem λ'. Hledaná pravděpodobnost je dána jako

$$P(X' = 0) = \frac{(\lambda')^0}{0!} \cdot e^{-\lambda'} = e^{-\lambda'} \ .$$

K určení parametru λ' použijeme vztah mezi veličinami Y a X' a podobně mezi veličinami Y a X, tj.

$$\frac{E(X')}{10 \ min} = \frac{1}{E(Y)} = \frac{E(X)}{60 \ min}$$

což odpovídá i požadavku, že průměrný počet událostí v časovém intervalu je úměrný jeho délce. Konkrétně tedy $\lambda' = E(X') = \frac{10 \ min}{60 \ min} \cdot 12 = 2$ a hledaná pravděpodobnost je pak

$$P(X' = 0) = e^{-2} \ .$$

(b)

- *Pomocí exponenciálního:* Z předchozího víme, že

$$Y = \text{“}\textit{doba čekání na hovor}\text{”} \ [\textit{v minutách}]$$

má exponenciální rozdělení se střední hodnotou $\tau = E(Y) = 5 \ min$. Pak pro hledaný čas $t > 0$ máme

$$0.7 = P(Y > t) = 1 - P(Y \leq t) = 1 - F_Y(t) = 1 - (1 - e^{-t/5}) = e^{-t/5}$$

tedy

$$t = -5 \cdot \ln 0.7 \doteq 1.78 \ min \ .$$

- *Pomocí Poissonova:* Vezmeme veličinu (závislou na t - zde je to parametr)

$$\widetilde{X} = \text{“}počet\ hovorů\ během\ doby\ t\ minut\text{”}$$

která bude mít Poissonovo rozdělení s parametrem $\widetilde{\lambda}$, pro který máme vztah

$$\widetilde{\lambda} = E(\widetilde{X}) = \frac{t}{\tau} = \frac{t}{5}$$

Hodnotu t pak dostaneme z rovnice

$$0.7 = P\left(\widetilde{X} = 0\right) = \frac{\left(\widetilde{\lambda}\right)^0}{0!} e^{-\widetilde{\lambda}} = e^{-t/5}$$

což vede pochopitelně na stejný výsledek

$$t = -5 \cdot \ln 0.7 \doteq 1.78\ min\ .$$

Příklad 6.6 *Do pojišťovny přijdou průměrně 2 hlášení škody denně.*

(a) Jaká je pravděpodobnost, že za den přijdou alespoň 4?

(b) Jaká je pravděpodobnost, že do pojišťovny přijde nejbližší hlášení škody nejdříve třetí den?

Řešení:

(a) Máme tedy veličinu

$$X = \text{“}počet\ hlášení\ za\ 1\ den\text{”}\ ,$$

s Poissonovým rozdělením Poiss(λ), kde $\lambda = E(X) = 2$ (jde o bezrozměrné číslo).

$$P(X \geq 4) = 1 - P(X < 4) = 1 - e^{-\lambda}\Big(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!}\Big) = 1 - e^{-2}\Big(1 + 2 + \frac{2^2}{2} + \frac{2^3}{3!}\Big) =$$

$$= 1 - \tfrac{19}{3}e^{-2} \doteq 1 - 0.857 = 0.143\ .$$

(b) Tuto úlohu můžeme řešit jak s využitím exponenciálního rozdělení, tak Poissonova rozdělení. Exponenciální rozdělení bude mít veličina

$$Y = \text{“}doba\ čekání\ na\ příchod\ dalšího\ hlášení\text{”}.$$

- *Pomocí exponenciálního:* Podle zadání má veličina X Poissonovo rozdělení s parametrem $\lambda = 2$. Tudíž pro náhodnou veličinu Y s exponenciálním rozdělením a střední hodnotou $E(Y) = \tau$ platí, že $\tau = \frac{T}{\lambda} = \frac{1\ den}{2} = 0.5\ dne$, kde $T = 1\ den$. Hledaná pravděpodobnost je pak

$$P(Y > 2\ dny) = 1 - P(Y \leq 2\ dny) = 1 - F_Y(2) = 1 - \left(1 - e^{-\frac{2}{0.5}}\right) = e^{-4} \doteq 0.018.$$

- *Pomocí Poissonova: Pomocí Poissonova:* Uvažujme veličinu

  X'= “*počet hlášení během 2 dnů*”

Hledaná pravděpodobnost je dána jako

$$P(X' = 0) = \frac{(\lambda')^0}{0!} \cdot e^{-\lambda'} = e^{-\lambda'}$$

kde λ' je parametr veličiny X' s Poissonovým rozdělením. K určení parametru λ' použijeme vztah mezi veličinami Y a X' a už známou hodnotu parametru veličiny X, tj. platí $\lambda' = \frac{T'}{\tau} = \lambda \cdot \frac{T'}{T}$ neboli

$$\frac{E(X')}{E(X)} = \frac{T'}{T}$$

což odpovídá i výše zmíněnému požadavku, že průměrný počet událostí v časovém intervalu je úměrný jeho délce. Konkrétně tedy $\lambda' = 2 \cdot \frac{2\ dny}{1\ den} = 4$, kde $T' = 2\ dny$, a hledaná pravděpodobnost je pak

$$P(X' = 0) = e^{-4} \doteq 0.018.$$