

Příklad 1

Uvažujme dvě absolutně spojitě nezávislé náhodné veličiny X a Y se stejným rozdělením. Spočtěte a dokažte $P(X > Y)$.

Klíčovou informací je, že veličiny X a Y jsou spojitě, nezávislé a stejně rozdělené.

$$P(X > Y) = P(Y < X) = a$$

(Toto víme, protože jsou veličiny **stejně rozdělené a nezávislé.**)

Dále víme

$$P(X = Y) = 0$$

(Toto víme, protože jsou veličiny **spojité a nezávislé.**)

Platí tedy, že

$$\begin{aligned} P(X = Y) + P(X > Y) + P(Y < X) &= 1 \\ 0 + a + a &= 1 \\ 2a &= 1 \\ a &= \frac{1}{2} \\ &= \underline{\underline{\frac{1}{2}}} \end{aligned}$$

Příklad 2

Vyjádřete výraz $P(X_{n+2} = j, X_{n+1} = k, X_n = i) | P(X_n = i)$ pomocí přechodových pravděpodobností $P_{i,j}$.

Příklad 3

Dokažte, že geometrické rozdělení nemá paměť.

- Nemá paměť = nezáleží na historii nějakého jevu.

Předpokládejme dobu čekání $T \sim \text{Geom}(p)$. Pokud jsme již čekali s jednotek času, zbývajících čas čekání $T - s$ má opět $\text{Geom}(p)$, tedy stejný, jako kdybychom zatím vůbec nečekali. Snažíme se tedy dokázat

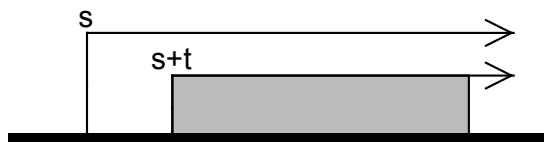
$$P(T - s > k | T > s) \stackrel{?}{=} P(T > k) \text{ pro } k, s > 0.$$

Upravme nerovnici a použijme vzorec pro podmíněnou pravděpodobnost a vzorec funkce přežití geometrického rozdělení

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)}, P(B) \neq 0 \\ P(T > n) &= (1-p)^n \end{aligned}$$

$$P(T - s > k | T > s) = P(T > k + s | T > s) = \frac{P(T > k + s \cap T > s)}{P(T > s)} = \dots$$

Průnik jevu v čitateli je následující:



celý výraz pokračuje vyjádřením pomocí funkce přežití:

$$\dots = \frac{P(T > k + s)}{P(T > s)} = \frac{(1-p)^{k+s}}{(1-p)^s} = \frac{\cancel{(1-p)^s} (1-p)^k}{\cancel{(1-p)^s}} = \boxed{(1-p)^k = P(T > k)}$$

Geometrické rozdělení tedy skutečně **nemá paměť**.

Příklad 4

Dokažte, že exponenciální rozdělení nemá paměť.

Struktura tohoto příkladu je stejná jako u důkazu, že geometrické rozdělení nemá paměť.

$$X \sim \text{Exp}(\lambda)$$

$$P(X > u + t | X > u) = \frac{P(X > u + t \cap X > u)}{P(X > u)} = \frac{P(X > u + t)}{P(X > u)} = \dots$$

Funkce přežití exponenciálního rozdělení je

$$P(X > k) = e^{-\lambda k}$$

$$\dots = \frac{e^{-\lambda(u+t)}}{e^{-\lambda u}} = \frac{\cancel{e^{-\lambda u}} e^{-\lambda t}}{\cancel{e^{-\lambda u}}} = \boxed{e^{-\lambda t} = P(X > t)}$$

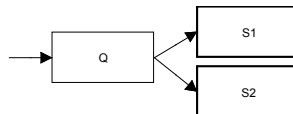
Exponenciální rozdělení tedy skutečně nemá paměť.

Příklad 5

Mějme server, který dokáže zpracovat paralelně až dva požadavky. Pokud je server zahlcen, ukládají se příchozí požadavky do fronty o kapacitě 5. Zpracování požadavků probíhá náhodně a nezávisle. Střední doba zpracování požadavku na serveru je $25ms$.

- (A) Server zpracovává dva požadavky. Jaká je průměrná doba odbavení třetího požadavku.
- (B) Jaká je průměrná doba zpracování všech tří požadavků.
- (C) Jaká je pravděpodobnost, že třetí požadavek bude zpracován jako poslední?

Graf exponenciálních závodů:



Důležité vzorce

$$\begin{aligned} E_{\min}\{S, T\} &= \frac{1}{\lambda + \mu} \\ E_{\max}\{S, T\} &= E(S) + E(T) - E_{\min}\{S, T\} \end{aligned}$$

Za předpokladu

$$S \sim \text{Exp}(\lambda), T \sim \text{Exp}(\mu)$$

$$S_1 \sim \text{Exp}(\lambda), E(S_1) = 25ms, \lambda = \frac{1}{E(S_1)} \rightarrow \lambda = \frac{1}{25}$$

$$S_2 \sim \text{Exp}(\mu), E(S_2) = 25ms, \mu = \frac{1}{E(S_2)} \rightarrow \mu = \frac{1}{25}$$

A)

S_1 a S_2 jsou obsazené, čekáme tedy, až se jeden z nich uvolní (ten nejrychlejší).

$$E(C_A) = E_{\min}\{S_1, S_2\} + E(C) = \frac{1}{\mu + \lambda} + E(C) = \frac{1}{\frac{1}{25} + \frac{1}{25}} + 25 = \dots = \underline{\underline{37,5ms}}$$

B)

Všechny tři požadavky budou zpracovány, když bude zpracován nejpomalejší z nich.

$$\begin{aligned} E(C_B) &= E_{\min}\{S_1, S_2\} + E_{\max}\{S_1, S_2\} = E_{\min}\{S_1, S_2\} + E(S_1) + E(S_2) - E_{\min}\{S_1, S_2\} = \\ &= E(S_1) + E(S_2) = 25 + 25 = \underline{\underline{50ms}} \end{aligned}$$

C)

$$\begin{aligned} P(C_L) &= P(C_L \cap A_F) + P(C_L \cap B_F) = \frac{\mu}{\mu + \lambda} * \frac{\lambda}{\mu + \lambda} + \frac{\mu}{\mu + \lambda} * \frac{\lambda}{\mu + \lambda} = \\ &= \frac{\frac{1}{25}}{\frac{1}{25} + \frac{1}{25}} * \frac{\frac{1}{25}}{\frac{1}{25} + \frac{1}{25}} + \frac{\frac{1}{25}}{\frac{1}{25} + \frac{1}{25}} * \frac{\frac{1}{25}}{\frac{1}{25} + \frac{1}{25}} = \dots = \underline{\underline{\frac{1}{2}}} \end{aligned}$$

Vzorec vychází z

$$P(S < T) * P(S > T) + \dots$$

Příklad 6

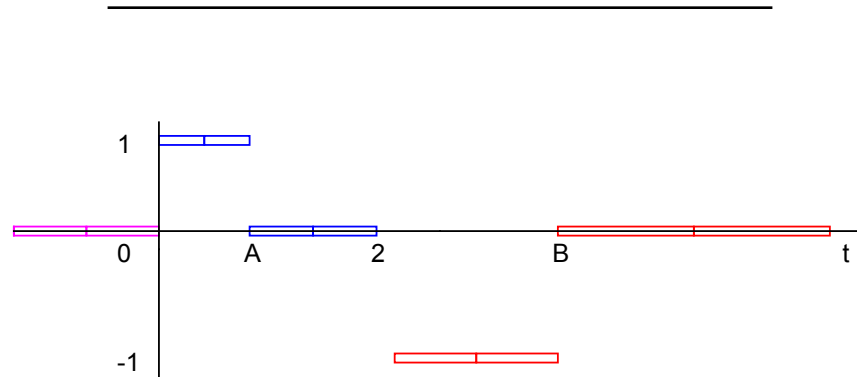
Spočtete entropii geometrického rozdělení, $p = \frac{1}{2}$.

Příklad 7

Náhodný proces $X(t)$ je definován jako: $X(t) = 1$ pro $0 \leq t < A$ a $X(t) = -1$ pro $2 \leq t \leq B$. A je rovnoměrně rozdělená veličina na intervalu $(0, 2)$. B je náhodná veličina s exponenciálním rozdělením s intenzitou $\lambda = 2$.

- (A) Nakreslete graf náhodného procesu.
 (B) Spočítejte $E(X(t))$.
 (C) Určete, zda je proces $X(t)$ stacionární ve střední hodnotě.

A)

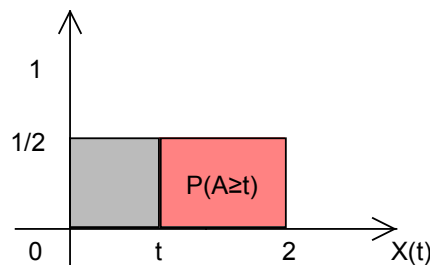


B)

Veličiny A a B jsou jinak rozdělené. Výpočet $E(X)$ tedy rozdělíme na tři části (první interval je jen pro formálnost).

- I. $t \leq 0$ $E(X(t)) = 0$
 II. $t \in (0, 2)$ $E(X(t)) = 1 * P(0 \leq t \leq A) + 0 * P(A \leq t \leq 2) = P(0 \leq t \leq A) = P(A \geq t)$
 III. $t \geq 2$ $E(X(t)) = (-1) * P(t \leq B) + 0 * P(t \geq B) = -P(B \geq t)$

Interval II.



Chceme spočítat pravděpodobnost (plocha pod křivkou), že $P(A \geq t)$. Jedná se o jednoduchý výpočet obsahu:

$$P(A \geq t) = (2 - t) * \frac{1}{2} = \frac{2 - t}{2}$$

Interval III.

Funkce přežití exponenciálního rozdělení je

$$P(X > k) = e^{-\lambda k}$$

$$-P(B \geq t) = -(e^{-\lambda t}) [\lambda = 2] = -e^{-2t}$$

Celkový výsledek je tedy:

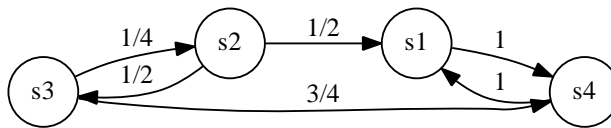
$$E(X(t)) = \frac{2 - t}{2} - e^{-2t}$$

C)

Střední hodnota není ustálená, závisí na čase t . Proces tedy **není stacionární ve střední hodnotě**.

Příklad 8

U markovského řetězce



- (A) Definujte rekurentní stav.
- (B) Najděte stacionární rozdělení.
- (C) Vypište rekurentní stavy.

A)

Stav j se nazývá rekurentní, jestliže řetězec, který z něj vychází, se do j vrátí s pravděpodobností 1 po konečně mnoha krocích. Tj. $P(T_j < \infty) = 1$.

B)

Sestrojme matici přechodů Q

$$Q = \begin{matrix} & \begin{matrix} s_1 & s_2 & s_3 & s_4 \end{matrix} \\ \begin{matrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

Protože je markovským řetězcem popsán jev diskrétní, platí pro stacionární rozdělení:

$$\pi : \pi * Q = \pi$$

Sestavme soustavu rovnic

$$\begin{aligned} \frac{1}{2}\pi_2 + \pi_4 &= \pi_1 \\ \frac{1}{4}\pi_3 &= \pi_2 \\ \frac{1}{2}\pi_2 &= \pi_3 \\ \pi_1 + \frac{3}{4}\pi_3 &= \pi_4 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1 \end{aligned}$$

Nyní vyřešme soustavu rovnic, pro tento případ použijme GEM.

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & \frac{1}{2} & 0 & 1 & 0 \\ 0 & -1 & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{2} & -1 & 0 & 0 \\ 1 & 0 & \frac{3}{4} & -1 & 0 \end{pmatrix} \sim \dots \sim \dots \Rightarrow \pi = \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right)$$

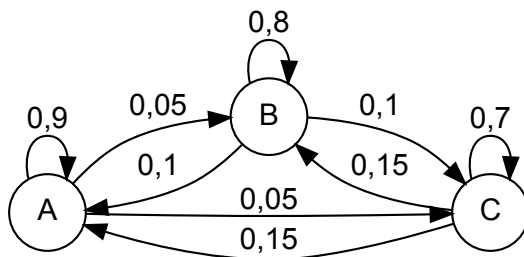
C)

Rekurentní stavy jsou s_1 a s_2 . Přijít na to můžeme pomocí *silně souvislých komponent* z teorie grafů.

Def.: Pro všechny dvojice vrcholů u a v existuje cesta z u do v a zároveň z v do u .

Příklad 9

Tři výrobky A , B , C mají následující pravděpodobnosti nákupu: $A = 0,9$, $B = 0,8$ a $C = 0,7$. Pro jiný než oblíbený výrobek se zákazníci rozhodnou ve zbytku procent rovnoměrně mezi zbylými dvěma. Spočítejte oblíbenost jednotlivých výrobků z dlouhodobého hlediska.



Sestavme matici přechodů Q :

$$Q = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} \frac{9}{10} & \frac{1}{20} & \frac{1}{20} \\ \frac{1}{10} & \frac{8}{10} & \frac{1}{10} \\ \frac{3}{20} & \frac{3}{20} & \frac{7}{10} \end{pmatrix} \end{matrix}$$

Pro stacionární rozdělení (*tedy oblíbenost z dlouhodobého hlediska*) platí:

$$\pi * Q = \pi \quad (1)$$

$$\pi_1 + \pi_2 + \dots + \pi_n = 1 \quad (2)$$

Sestavme rovnice pro výpočet stacionárního rozdělení (po sloupcích):

$$\begin{aligned} \frac{9}{10}\pi_1 + \frac{1}{20}\pi_2 + \frac{1}{20}\pi_3 &= \pi_1 \\ \frac{1}{10}\pi_1 + \frac{8}{10}\pi_2 + \frac{1}{10}\pi_3 &= \pi_2 \\ \frac{3}{20}\pi_1 + \frac{3}{20}\pi_2 + \frac{7}{10}\pi_3 &= \pi_3 \\ \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned}$$

Soustavu upravíme pro použití v GEM:

$$\begin{aligned} -2\pi_1 + 2\pi_2 + 3\pi_3 &= 0 \\ \pi_1 - 4\pi_2 + 3\pi_3 &= 0 \\ \pi_1 + 2\pi_2 - 6\pi_3 &= 0 \\ \pi_1 + \pi_2 + \pi_3 &= 1 \end{aligned}$$

A vyřešíme (nesmíme zapomenout počítat s pravou stranou matice a s rovnicí $\pi_1 + \pi_2 + \pi_3 = 1$):

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -6 & 0 \\ 1 & -4 & 3 & 0 \\ -2 & 2 & 3 & 0 \end{pmatrix} \sim \dots \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -7 & -1 \\ 0 & 0 & -33 & -6 \\ 0 & 0 & 33 & 6 \end{pmatrix}$$

$$\boxed{\pi = \left(\frac{6}{11}, \frac{3}{11}, \frac{2}{11} \right)}$$

Příklad 10

Mějme zadány tabulky rozdělení X a Y a *codeword*.

X, Y	-12	5	X, Y	-12	5
-12	$\frac{1}{8}$	$\frac{1}{8}$	-12	110	110
4	$\frac{1}{2}$	$\frac{1}{2}$	5	10	0

- (A) Najděte nejmenší možnou očekávanou délku kódového slova jakéhokoliv binárního kódu pro náhodný vektor X, Y s výše uvedeným rozdělením.
- (B) Dosahuje výše uvedené kódování tohoto minima? Proč?
- (C) Jsou náhodné veličiny X a Y nezávislé? Proč?

A)

Délku kódového slova spočteme jako

$$L(C) = \mathbb{E} \ell(X) = \sum_i \ell(x_i) * \mathbb{P}(X = x_i)$$

$$L(C) \geq H_D(X)$$

Tedy:

$$L(C) = \frac{1}{8} * 3 + \frac{1}{8} * 3 + \frac{1}{2} * 2 + \frac{1}{2} * 1 = 2$$

B)

Kódování je optimální tehdy, když platí $L(C) = H(X)$.

$$H(X, Y) = - \sum p_i \log_b p_i = - \left(\frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{8} \log_2 \frac{1}{8} + \frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{4} \log_2 \frac{1}{4} \right) = -\frac{7}{4}$$

$L(C) \geq H(X)$ kódování tedy **není optimální** (délka slova není optimální)

C)

Pro nezávislost jevů obecně platí

$$P(A \cap B) = P(A) * P(B)$$

V tabulce jsou uvedené sdružené pravděpodobnosti ($P(A \cap B)$), spočteme tedy samostatné pravděpodobnosti a ty porovnáme s tabulkou.

$$P(X = -12) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4} \text{ (součet řádku)}$$

$$P(Y = -12) = \frac{1}{8} + \frac{1}{2} = \frac{5}{8} \text{ (součet sloupce)}$$

$$P(X = -12 \wedge Y = -12) = \frac{1}{4} * \frac{5}{8} = \frac{5}{32}$$

$$\frac{5}{32} \neq \underbrace{\frac{1}{8}}_{\text{z tabulky}}$$

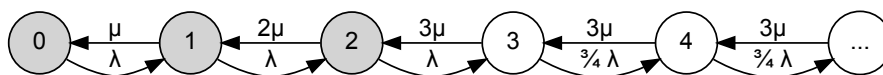
Veličiny jsou tedy závislé, tedy nejsou nezávislé.

Příklad 11

Systém serverů přijímá požadavky s intenzitou $\lambda = 20$ požadavků za sekundu. V systému jsou 3 servery, jeden požadavek se zpracovává $E(T) = 50ms$. Pokud jsou servery vytížené, s pravděpodobností $P = 25\%$ jsou požadavky zahozeny, jinak se ukládají do fronty.

- (A) Nakreslete diagram systému.
 (B) Nalezněte stacionární rozdělení.
 (C) Jaká je pravděpodobnost, že ve frontě není žádný požadavek?

A)



B)

Řešení tohoto příkladu má **tři fáze**: I. rekurentní předpis, II. explicitní předpis a III. obecný předpis.

$$E(T) = \frac{1}{\mu} \rightarrow 50 = \frac{1}{\mu} \rightarrow \mu = \frac{1}{50} \text{ požadavků na ms}$$

μ musíme převést na stejné jednotky jako λ . $\frac{1}{50} * 1000 \rightarrow \boxed{\mu = 20}$ požadavků za s.

Stacionární rozdělení nalezneme pomocí *detailní rovnováhy* (řetězec je typu „birth-and-death“).

I. rekurentní předpis:

$$\begin{aligned} \lambda\pi_0 &= \mu\pi_1 \rightarrow \pi_1 = \frac{\lambda}{\mu}\pi_0 \\ \lambda\pi_1 &= 2\mu\pi_2 \rightarrow \pi_2 = \frac{\lambda}{2\mu}\pi_1 \\ \lambda\pi_2 &= 3\mu\pi_3 \rightarrow \pi_3 = \frac{\lambda}{3\mu}\pi_2 \\ \frac{3}{4}\lambda\pi_3 &= 3\mu\pi_4 \rightarrow \pi_4 = \frac{\lambda}{4\mu}\pi_3 \\ \vdots & \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \vdots & \quad \vdots \quad \vdots \rightarrow \pi_{n+1} = \frac{\lambda}{4\mu}\pi_n \quad \text{pro } \forall n \geq 3 \end{aligned}$$

Pomocí zpětného dosazování (π_0) vytvoříme **II. explicitní předpis**:

$$\begin{aligned} \pi_0 &= \pi_0 \\ \pi_1 &= \frac{\lambda}{\mu}\pi_0 \\ \pi_2 &= \frac{\lambda^2}{2\mu^2}\pi_0 \\ \pi_3 &= \boxed{\frac{\lambda^3}{6\mu^3}\pi_0} \\ \pi_4 &= \frac{\lambda^3}{24\mu^4}\pi_0 \\ \vdots & \quad \vdots \quad \vdots \end{aligned}$$

Ve stavu π_3 dojde k ustálení systému a graf bude vypadat až do ∞ stejně. Vyjádříme tedy obecně π_n (**III. obecný předpis**):

$$\pi_n = \pi_3 * \left(\frac{\frac{3}{4}\lambda}{3\mu}\right)^{n-3} = \left(\frac{\lambda^3}{6\mu^3} * \pi_0\right) * \left(\frac{\frac{3}{4}\lambda}{3\mu}\right)^{n-3} = \boxed{\frac{\lambda^3}{6\mu^3} * \pi_0 * \left(\frac{\lambda}{4\mu}\right)^{n-3}} \quad \text{pro } n \geq 3$$

Dále musí platit, že $\sum_{i=0}^{\infty} \pi_i = 1$ nebo-li $\pi_0 + \pi_1 + \dots + \pi_n = 1$, tedy

$$\pi_0 + \frac{\lambda}{\mu} \pi_0 + \frac{\lambda^2}{2\mu^2} \pi_0 + \sum_{n=3}^{\infty} \left[\frac{\lambda^3}{6\mu^3} * \pi_0 * \left(\frac{\lambda}{4\mu} \right)^{n-3} \right] = \pi_0 * \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu} + \underbrace{\sum_{n=3}^{\infty} \left[\frac{\lambda^3}{6\mu^3} * \left(\frac{\lambda}{4\mu} \right)^{n-3} \right]}_S \right) = 1$$

$$S = \sum_{n=3}^{\infty} \left[\frac{\lambda^3}{6\mu^3} * \left(\frac{\lambda}{4\mu} \right)^{n-3} \right] [\lambda = \mu = 20] = \sum_{n=3}^{\infty} \frac{20^3}{6 * 20^3} * \left(\frac{20}{4 * 20} \right)^{n-3} = \frac{1}{6} \sum_{n=3}^{\infty} \left(\frac{1}{4} \right)^{n-3} = \dots$$

Součet nekonečné geometrické řady spočteme jako $S_g = \frac{a_0}{1-q}$ pro $|q| < 1$, tedy:

$$\dots = \frac{1}{6} * \frac{1}{1 - \frac{1}{4}} = \dots = \boxed{\frac{2}{9}}$$

$$\pi_0 * \left(1 + 1 + \frac{1}{2} + \frac{2}{9} \right) = 1 \Rightarrow \pi_0 = \frac{18}{\underline{\underline{49}}}$$

Výsledné stacionární rozdělení je

$$\pi = \left(\pi_0, \pi_0, \frac{1}{2}\pi_0, \frac{1}{6}\pi_0, \frac{1}{24}\pi_0, \dots \right) = \boxed{\left(\frac{18}{49}, \frac{18}{49}, \frac{9}{49}, \frac{3}{49}, \frac{3}{196}, \dots \right)}$$

C)

Fronta je prázdná tehdy, pokud se v systému nachází 0, 1, 2 nebo 3 požadavky.

$$P(X(t) \leq 3) = \pi_0 + \pi_1 + \pi_2 + \pi_3 = \frac{18}{49} + \frac{18}{49} + \frac{9}{49} + \frac{3}{49} = \frac{48}{\underline{\underline{49}}}$$

Příklad 12

Uvažujte markovský řetězec s konečnou množinou stavů $S = \{s_1, s_2, \dots, s_n\}$.

- (A) Definujte Markovskou podmínku.
- (B) Definujte stacionární rozdělení pro tento markovský řetězec.
- (C) Pro stav $s_i \in S$ definujte čas návratu τ_i .
- (D) Jak souvisí střední hodnota času návratu se stacionárním rozložením?

A) Markovova podmínka

$$P(X_n = s_j | X_0 = s_0, X_1 = s_1, \dots, X_{n-1} = s_i) = P(X_n = s_j | X_{n-1} = s_i) = P_{ij}$$

B) Stacionární rozdělení

π je stacionární rozdělení, jestliže platí

$$\pi * P = \pi \text{ (přesněji zapsáno: } \sum_{\text{all } i} \pi_i P_{ij} = \pi_j)$$

(P je přechodová matice)

Pro stacionární rozdělení platí

$$\pi_i \geq 0 \text{ a } \sum_{i=1}^n \pi_i = 1$$

C) Čas návratu

Pro Markovův řetězec

$$(X_n)_{n \geq 0}$$

začínající ve stavu $s_i \in S$, definujeme **náhodný čas návratu**

$$\tau_i = \inf \{n \geq 1 : X_n = s_i\}.$$

Náhodný čas návratu je nejmenší čas n , ve kterém se vrátíme do stavu s_i , přičemž počáteční stav se jako návrat nepočítá (proto $n \geq 1$).

Příklad 13

Náhodný výběr 16 hodnot z normálního rozdělení. Je zadán 98% konfidenční interval (7,398; 12,602). Otestujte hypotézu $H_0 : \mu = 7,5$ vs. $H_A : \mu \neq 7,5$ s chybou 1. stupně $\alpha = 5\%$.
(Pozn. jsou dostupné tabulky studentova a chi-kvadrát rozdělení.)

Máme zadáný 98% CI, my však potřebujeme 95% interval.
 σ^2 neznáme, použijeme tedy \mathcal{T} -rozdělení:

$$\mu \in \left(\bar{X}_n - \mathcal{T}_{\frac{\alpha}{2}, n-1} \frac{s_n}{\sqrt{n}}; \bar{X}_n + \mathcal{T}_{\frac{\alpha}{2}, n-1} \frac{s_n}{\sqrt{n}} \right)$$

$$\bar{X}_n = \frac{7,398 + 12,602}{2} = \frac{20}{2} = 10$$

$$\mathcal{T}_{\frac{\alpha}{2}, n-1} = \mathcal{T}_{0,01; 16-1} = \boxed{2,602}$$

(Máme 98% interval /jen na jedné straně/ $\Rightarrow \alpha = 2\% = \frac{0,02}{2} = 0,01$)
(Hodnotu 2,602 jsme získali z tabulky Studentova rozdělení / \mathcal{T} -rozdělení/: řádek 15 a sloupec 0,99.)

Nyní nám chybí směrodatná odchylka s , tu získáme ze zadaného 98% intervalu:

$$CI_L = \bar{X}_n - \mathcal{T}_{\frac{\alpha}{2}, n-1} * \frac{s_n}{\sqrt{n}}$$

$$7,398 = 10 - 2,602 * \frac{s_n}{4}$$

$$s_n = \boxed{4}$$

V tuto chvíli máme všechny potřebné údaje na sestavení 95% CI.

$$CI_{95} = \left(10 - \mathcal{T}_{\frac{0,05}{2}, 15}; 10 + \mathcal{T}_{\frac{0,05}{2}, 15} \right) = (10 - 2,131; 10 + 2,131) = \boxed{(7,869; 12,131)}$$

(Opět hledáme v tabulce Studentova rozdělení, tentokrát na řádku 15 a sloupci 0,975 – chceme 95%, na každé straně 2,5%)

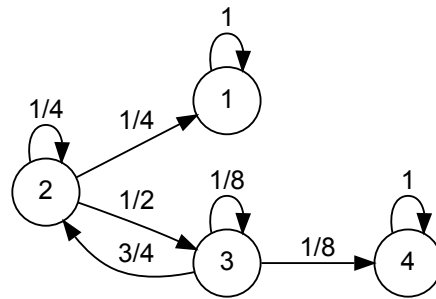
Ptáme se

$$H_0 \stackrel{?}{\in} CI_{95}$$

$$7,5 \notin CI_{95} \Rightarrow \text{Hypotézu } H_0 \text{ **zamítáme**.}$$

Příklad 14

Mějme Markovův řetězec



se dvěma absorbujícími stavy (s_1, s_2) a dvěma neabsorbujícími stavy (s_2, s_3).

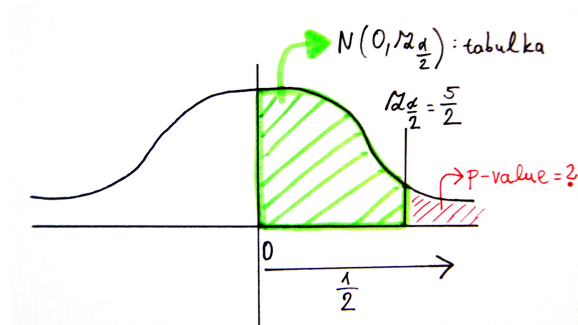
- (A) Definujte fundamentální matici.
- (B) Jaký bude střední počet kroků do pohlcení, pokud budeme začínat ve stavu s_3 ?
- (C) Jaká je pravděpodobnost, že budeme pohlceni stavem s_4 , pokud začínáme ve stavu s_1 ?

Příklad 15

Mějme N nezávislých jevů $X_1 = \text{Exp}(\lambda_1), \dots, X_n = \text{Exp}(\lambda_n)$ a jev $Y = \min(X_1, \dots, X_n)$. Najděte hustotu a distribuční funkci Y .

Příklad 16

Mějme 25 naměřených hodnot. Jejich součet je 250, $\sigma^2 = 16$. Otestujte hypotézu $H_0 : \mu = 12$ proti $H_A : \mu \neq 12$. Při jaké nejmenší chybě můžeme H_0 ještě zamítnout?



Hledáme p -value, tedy pravděpodobnost, že H_A nastane.

$$\bar{X}_n = \frac{\sum}{N} = \frac{250}{25} = 10$$

Pracujeme jen s jednou stranou intervalu, tedy

$$\begin{aligned}\mu_0 &= \bar{X}_n + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{N}} \\ z_{\frac{\alpha}{2}} &= \frac{\bar{X}_n - \mu_0}{\frac{\sigma}{\sqrt{N}}} \\ z_{\frac{\alpha}{2}} &= \frac{10 - 12}{\frac{\sqrt{16}}{\sqrt{25}}} \\ z_{\frac{\alpha}{2}} &= -\frac{5}{2}\end{aligned}$$

V tabulce normálního rozdělení nalezneme příslušnou hodnotu

$$N\left(0, z_{\frac{\alpha}{2}}\right) = N\left(0, \frac{5}{2}\right) = 0,4938$$

$$\frac{\alpha}{2} = \frac{1}{2} - 0,4938 = 0,0062$$

(Chceme zjistit hodnotu pod křivkou jen na jedné straně intervalu)

$$p = 2 * \frac{p}{2} = 2 * 0,0062 = \underline{\underline{0,0124}}$$

(Výsledná pravděpodobnost je však pod křivkou na obou stranách grafu)

p -value je tedy $0,0124 = 1,24\%$. V tomto bodě se přesně láme zamítnout/nezamítnout hypotézu. Pro zamítnutí potřebujeme hodnotu větší, tedy výsledek

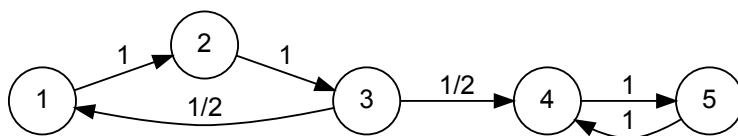
$$p\text{-value} = \underline{\underline{0,0124 + \epsilon}}$$

Příklad 17

Mějme dvě veličiny $X_1 \sim \text{Geom}(p_1)$ a $X_2 \sim \text{Geom}(p_2)$. Určete rozdělení $Y = \min(X_1, X_2)$.

Příklad 18

Mějme Markovský řetězec v diskrétním čase



- (A) Definujte transientní stav.
- (B) Spočtěte stacionární rozdělení.
- (C) Nalezněte transientní stavy. Zdůvodněte.
