Definice Entropie stochastického procesu
$\mathbb{X} = (X_1, X_2, \dots)$
je definována jako

$$H(\mathbb{X}) = \lim_{n\to\infty} \frac{H(X_1, X_2, \dots, X_n)}{n}$$

pokud tato limita existuje

Příklady: a) Psací stroj generuje náhodně znak z množiny m znaků se stejnou pravděpodobností! nezávisle na sobě.

$$H(X_1, \dots, X_n) = \log m^n = n \cdot \log m$$

$$H(\mathbb{X}) = \lim_{n\to\infty} \frac{n \cdot \log m}{n} = \log m$$

---

b) zobecnění i.i.d

$$H(\mathbb{X}) = \frac{1}{n} \lim_{n\to\infty} H(X_1, \dots, X_n) =$$

$$= \dots \frac{1}{n} n \cdot H(X_1) = H(X_1).$$

Definice Podmíněná entropie $H'(\mathbb{X})$ je definována jako

$$H'(\mathbb{X}) = \lim_{n \to \infty} H(X_n \mid X_{n-1}, \dots, X_1).$$

Věta Pro stacionární stochastický proces $\mathbb{X}$ existují $H(\mathbb{X})$ a $H'(\mathbb{X})$ a platí, že

$$H(\mathbb{X}) = H'(\mathbb{X})$$

Důkaz:

$$H(X_{n+1} \mid X_1, \dots, X_n) \le H(X_{n+1} \mid X_n, \dots, X_2)$$
$$= H(X_n \mid X_{n-1}, \dots, X_1)$$

Jedná se tedy o nerostoucí omezenou posloup.
Která má vlastní limitu.
Tedy $H'(\mathbb{X})$ existuje.

Tvrzení: Cesarovské sumy

Jestliže $a_n \to a$ a $b_n = \frac{1}{n}\sum_{i=1}^{n} a_i$ pak $b_n \to a$.

Dk: $\forall \varepsilon > 0\ \exists n_0\ \forall n > n_0\ |a_n - a| \le \varepsilon$

Pak $|b_n - a| = \left|\frac{1}{n}\sum_{i=1}^{n}(a_i - a)\right| \le$

$$\le \frac{1}{n}\sum_{i=1}^{n}|a_i - a| \le \frac{1}{n}\sum_{i=1}^{n_0}|a_i - a| + \frac{n-n_0}{n}\varepsilon$$

$$\le \frac{1}{n}\sum_{i=1}^{n_0}|a_i - a| + \varepsilon \le 2\varepsilon.$$

pro n dostatečně velké je suma $\le \varepsilon$

$\Rightarrow |a - b_n| \le 2\varepsilon$

dokončení důkazu

řetězové pravidlo

$$\frac{1}{n}H(X_1, X_2, \ldots X_n) = \frac{1}{n}\big(H(X_1) + H(X_2|X_1) + H(X_3|X_2, X_1) + \ldots + H(X_n|X_{n-1}, \ldots X_1)\big)$$

Cesarovské součty

$a_k = H(X_k | X_{k-1}, \ldots, X_1)$.

$\Rightarrow H(\mathbb{X}) = H'(\mathbb{X})$.

Jednodušeji vyjde pro Markovův řetězec:

<u>Věta</u>: Je-li $\mathbb{X}$ stacionární Markovův proces pak

$$H(\mathbb{X}) = H(X_2 | X_1)$$

<u>Dk</u>: $H(\mathbb{X}) = H'(\mathbb{X}) = \lim_{n\to\infty} H(X_n | X_{n-1}, \dots, X_1)$

$$= \lim_{n\to\infty} H(X_n | X_{n-1}) = H(X_2 | X_1)$$

<u>Příklad</u>: Stacionární Markovův proces s přechodovou maticí

$$\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \qquad 0 \le \alpha, \beta \le 1$$

má entropii

$$H(\mathbb{X}) = H(X_2 / X_1) = \frac{\beta}{\alpha+\beta} H(\alpha) + \frac{\alpha}{\alpha+\beta} H(\beta)$$

BOBO

obecně: $\vec{p} = (p_1, \dots, p_n)$ – stac. rozdělení Markovova řetězce

(P_{ij}) – matice přechodu

$$H(\mathbb{X}) = H(X_2 / X_1) =$$
$$= \sum_{i,j=1}^{n} p_i P_{ij} \log P_{ij}$$

Příklad: Stac. Markovův řetězec má matici přechodu

$$P = \begin{pmatrix} 1-p & p \\ 1 & 0 \end{pmatrix}$$

Kdy je rychlost entropie maximální?

$$H(\mathbb{X}) = \frac{H(p)}{p+1} = \frac{p \log p + (1-p)\log(1-p)}{p+1}$$

Dcv: kalkulus.

$$\Rightarrow p = \frac{3-\sqrt{5}}{2} = 0{,}382$$

a $H(p) = 0{,}694$.

10. Informační zdroje nad abecedou $\mathfrak{X} = \{1, \ldots, 5\}$ jsou popsány dvěma pravděpodobnostními funkcemi p a q. Uvažujme binární kódy C_1 a C_2 pro tyto zdroje:

x	$p(x)$	$q(x)$	$C_1(x)$	$C_2(x)$
1	1/2	1/2	0	0
2	1/4	1/8	10	100
3	1/8	1/8	110	101
4	1/16	1/8	1110	110
5	1/16	1/8	1111	111

Ověřte, že kód C_1 je optimální pro zdroj s pravděpodobnostní funkcí p a kód C_2 je optimální pro zdroj s pravděpodobnostní funkcí q. Pokud užijeme kód C_1 pro zdroj popsaný pomocí q, jaké chyby se dopustíme?

Řešení:

Určíme entropie p a q:

$$H(p) = 1.875,$$
$$H(q) = 2.$$

Avšak střední délka kódu C_1 je 1.875, střední délka kódu C_2 je 2: oba kódy jsou optimální. Střední délka $L_q(C_1)$ kódu C_1 použitého na zdroj s pravděpodobnostní funkcí q je

$$\frac{1}{2} + (2+3) \cdot \frac{1}{8} + 2 \cdot 4 \cdot \frac{1}{8} = 2.125.$$

Kód C_1 tedy není pro q optimální: za nevhodné kódování zaplatíme chybou o velikosti rozdílu

$$L_q(C_1) - H(q) = 0.125,$$

což je právě hodnota informační divergence $D(q||p)$.

11. Markovský řetězec nad abecedou $\Lambda = \{a, b, c\}$ je popsán těmito pravděpodobnostmi přechodu: pravděpodobnosti přechodů ze stavu a i ze stavu b jsou rovnoměrné, přechod ze stavu c do c nikdy nemůže nastat, ostatní podmíněné pravděpodobnosti jsou shodné. Určete rychlost entropie odpovídajícího markovského zdroje informace.

Řešení:

Ze zadání dostaneme matici přechodu:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

Rychlost entropie markovského zdroje splňuje $H((X_n)_{n\in\mathbb{N}}) = H(X_2|X_1)$. Stačí tedy určit

$$H(X_2|X_1) = \sum_{x_1\in\{a,b,c\}} p(x_1)\cdot H(X_2|x_1),$$

kde $p = (p(a), p(b), p(c))$ je stacionární rozdělení řetězce a $H(X_2|x_1)$ jsou podmíněné entropie jednotlivých řádku matice $\mathbf{P}$. Platí

$$H(X_2|a) = H(X_2|b) = \log 3, \quad H(X_2|c) = \log 2 = 1.$$

Stacionární rozdělení nalezneme řešením soustavy

$$p\cdot\mathbf{P} = p, \quad p(a) + p(b) + p(c) = 1.$$

Dostaneme $p = (0.375, 0.375, 0.25)$. Proto

$$H(X_2|X_1) = 2\cdot 0.375\log 3 + 0.25\cdot 1 = 0.75\log 3 + 0.25 \doteq 1.435.$$

12. Určete rychlost entropie markovského zdroje informace s abecedou $\mathfrak{X} = \{a, b, c\}$, jehož matice přechodu je

$$\mathbf{P} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Jaká je maximální rychlost entropie libovolného markovského zdroje s abecedou $\mathfrak{X}$? Odpověď zdůvodněte.

Řešení:

Nejprve nalezneme stacionární rozdělení $\mathbf{p} = (p_1, p_2, p_3)$ Markovova řetězce zadaného maticí $\mathbf{P}$. Řešíme soustavu lineárních rovnic

$$\mathbf{p} = \mathbf{p}\mathbf{P},$$
$$p_1 + p_2 + p_3 = 1.$$

Jejím řešením je vektor $\mathbf{p} = \left(\frac{2}{5}, \frac{2}{5}, \frac{1}{5}\right)$. Informační zdroj $X_1, X_2, \ldots$ je markovský, pokud pro jeho počáteční rozdělení $\mathbf{p}(0)$ platí $\mathbf{p}(0) = \mathbf{p}$. Potom je rychlost entropie tohoto zdroje rovna střední podmíněné entropii $H(X_2|X_1)$. Tedy

$$H(X_2|X_1) = \sum_{i=0}^{2} p_i(0)\cdot H(X_2|X_1 = i) = \frac{2}{5}\cdot H\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{2}{5}.$$

Maximální rychlost entropie markovského zdroje pro tento zdroj je maximum entropie na tříprvkové množině, neboť $H(X_2|X_1)$ je konvexní kombinací podmíněných entropií na tříprvkových množinách. V našem případě tedy $\log 3 \doteq 1.585$.

Zpět k blokovému kódování!

$\mathbb{X} = (X_1, X_2, \ldots)$ – časová řada

blokujeme po n znacích $(X_1, \ldots, X_n)$

"supersymboly" $p(x_1, \ldots, x_n)$

z $\mathcal{X}^n$

střední délka na jeden znak

$$L_n = \frac{1}{n} \sum_{(x_1, \ldots, x_n) \in \mathcal{X}^n} \ell(C(x_1, \ldots, x_n)) \cdot p(x_1, \ldots, x_n)$$

$$\frac{H(X_1, \ldots X_n)}{n} \le L_n \le \frac{H(X_1, \ldots, X_n)}{n} + \frac{1}{n}$$

Má-li $\mathbb{X}$ entropii $H(\mathbb{X})$, pak

$$\lim_{n \to \infty} L_n = H(\mathbb{X})$$

! Neboli, pro jakýkoliv blokový instantní kód jde L_n k $H(\mathbb{X})$. !

13. Nalezněte střední délku Huffmanova kódu pro informační zdroj generující řetězec

aaababbcabc

Určete, jaké úspory dosáhneme použitím blokového Huffmanova kódování s délkou bloku 2.

Řešení:

Při použití Huffmanova kódu o délce bloku 1 máme tuto tabulku:

	a	b	c
$p(x)$	$\frac{5}{11}$	$\frac{4}{11}$	$\frac{2}{11}$
C_H	0	11	10

→ odhad z četnosti

Při použití bloku délky 2 dostaneme tento Huffmanův kód:

	aa	ab	ba	ac	ca	bb	bc	cb	cc
$p(x,y)$	$\frac{25}{121}$	$\frac{20}{121}$	$\frac{20}{121}$	$\frac{10}{121}$	$\frac{10}{121}$	$\frac{16}{121}$	$\frac{8}{121}$	$\frac{8}{121}$	$\frac{4}{121}$
C_H^2	01	110	111	000	1011	100	0011	1010	0010

Střední délky kódů vztažené na 1 znak zdrojové abecedy:

$$L(C_H) = \frac{17}{11} \doteq 1.545, \quad \frac{L(C_H^2)}{2} \doteq \frac{3.041}{2} = 1.521.$$

Tedy dosáhneme komprese lepší o

$$\left(1 - \frac{1.521}{1.545}\right) \times 100\,\% = 1.55\,\%.$$

14. Při vysílání dvouprvkové abecedy $\{\cdot, -\}$ se zkreslí 8% teček a 2% čárek. Zpráva, ve které se tečky a čárky vyskytují rovnočetně, obsahuje 154 bitů informace. Kolik z nich uvedený informační kanál v průměru přenese správně?

Řešení:

Rozdělení informačního zdroje X s abecedou $\{\cdot, -\}$ je na vstupu popsáno rovnoměrnou pravděpodobnostní funkcí p_X. Podmíněné pravděpodobnosti $p_{Y|X}(y|x)$, $x, y \in \{\cdot, -\}$, vyjadřující možnou záměnu symbolu jsou dány maticí

$$\begin{pmatrix} 0.92 & 0.08 \\ 0.02 & 0.98 \end{pmatrix}.$$

Univerzální kódování

1. Metoda typů

Csiszár, Körner 1981 aplikace: univerzální kód.
testování hypotéz

Značení: • $\mathcal{X} = \{a_1, a_2, \dots, a_{|\mathcal{X}|}\}$ abeceda
- symboly

• $\vec{x} = (x_1, x_2, \dots x_n) \in \mathcal{X}^n$

• $P_{\vec{x}}(a) = N(a|\vec{x})/n$ ... relativní četnost
$N(a|\vec{x})$ ... kolikrát je symbol $a \in \mathcal{X}$ v posloupnosti $\vec{x}$

• Typ $P_{\vec{x}}$ (empirické rozdělení) je rozdělení $(P_{\vec{x}}(a) : a \in \mathcal{X})$ na $\mathcal{X}$.

• $\mathcal{P}_n$ ... množina všech typů (se jmenovatelem n)

Příklad: $\mathcal{X} = \{0, 1\}$

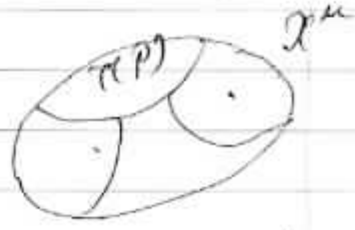
$$\mathcal{P}_n = \{(P(0), P(1)) : \left(\frac{0}{n}, \frac{n}{n}\right), \left(\frac{1}{n}, \frac{n-1}{n}\right) \dots \left(\frac{n}{n}, 0\right)\}$$

$|\mathcal{P}_n| = n+1$

BOBO

<u>Definice</u>: Ať $P \in \mathcal{P}_n$. Množina všech posloupností délky n a typu P se nazývá <u>třída typu P</u>. Značíme $T(P)$

$$T(P) = \{\vec{x} \in \mathcal{X}^n : P_{\vec{x}} = P\}$$



(posloupnosti s předepsaným počtem symbolů)

<u>Příklad</u>: $\mathcal{X} = \{1, 2, 3\}$

$\vec{x} = 11321$

$P_{\vec{x}} = \dots \quad P_{\vec{x}}(1) = \frac{3}{5}, \quad P_{\vec{x}}(2) = \frac{1}{5}, \quad P_{\vec{x}}(3) = \frac{1}{5}$

$T(P_{\vec{x}}) = \{11123, 11132, 11213, \dots, 32111\}$

$$|T(P_x)| = \binom{5}{3,1,1} = \frac{5!}{3!\,1!\,1!} = 20$$

Kombinatorická úloha: Jak velká je množina $T(P_{\vec{x}})$:

$$\binom{n}{x_1, x_2, \dots x_n} = \frac{n!}{x_1!\,x_2! \dots x_n!}$$

$$\binom{n}{N(a_1|\vec{x}), N(a_2|\vec{x}), \dots N(a_{|\mathcal{X}|}|\vec{x})} =$$

$$= \frac{n!}{N(a_1|\vec{x})!\, N(a_2|\vec{x})! \cdots N(a_{|X|}|\vec{x})!}$$

Pro $X = \{0,1\}$, $\vec{x} \in X^n$ a $P_{\vec{x}} = \left(\frac{k}{n}, \frac{n-k}{n}\right)$

je $|T(P_{\vec{x}})| = \binom{n}{k}$

Věta

$$|\mathcal{P}_n| \le (n+1)^{|X|}$$

__Důkaz__: Čitatelé zlomků v $P_{\vec{x}}$ se jmenovatelem n se mohou pohybovat od 0 do n.
Tedy $|\mathcal{P}_n| \le (n+1)^{|X|}$

Je to hrubý odhad, ale důležitá je skutečnost, že počet prvků v $\mathcal{P}_n$ roste polynomiálně s n!

Zdrojová distribuce na $\mathcal{X}$.

$Q(x)$; $x \in \mathcal{X}$, $Q(x) > 0$; $\forall x \in \mathcal{X}$

Rozšíříme na $\mathcal{X}^n$

$$Q^n(x_1, x_2 \ldots x_n) = Q(x_1) \cdot Q(x_2) \cdots Q(x_n)$$

i.i.d. - model.

<u>Věta</u> Pro $\vec{x} \in \mathcal{X}^n$ $Q(\vec{x})$ závisí pouze na typu $P_{\vec{x}}$ a platí

$$\boxed{Q^n(\vec{x}) = 2^{-n\left(H(P_{\vec{x}}) + D(P_{\vec{x}} \| Q)\right)}}$$

<u>Dk</u>: $$Q^n(\vec{x}) = \prod_{i=1}^{n} Q(x_i) = \prod_{a \in \mathcal{X}} Q(a)^{N(a|\vec{x})}$$

$$= \prod_{a \in \mathcal{X}} Q(a)^{nP_{\vec{x}}(a)} = \prod_{a \in \mathcal{X}} 2^{nP_{\vec{x}}(a) \log Q(a)}$$

$$= \prod_{a \in \mathcal{X}} 2^{n(P_{\vec{x}}(a) \log Q(a) - P_{\vec{x}}(a) \log P_{\vec{x}}(a) + P_{\vec{x}}(a) \log P_{\vec{x}}(a))}$$

$$= 2^{n \sum_{a \in \mathcal{X}} \left(-P_{\vec{x}}(a) \log \frac{P_{\vec{x}}(a)}{Q(a)} + P_{\vec{x}}(a) \log P_{\vec{x}}(a)\right)}$$

$$= 2^{n(-D(P_{\vec{x}} \| Q) - H(P_{\vec{x}}))}$$

Důsledek: Je-li $\vec{x}$ třídy Q (tj. $Q = P_{\vec{x}}$) pak

$$\boxed{Q^n(\vec{x}) = 2^{-nH(Q)} = 2^{-nH(P_{\vec{x}})}}$$

Příklad: Férová kostka: Posloupnost délky n bude mít přímé $n/6$ výskytů jedné strany a pravděpodobnosti

$$2^{-nH\left(\frac{1}{6}, \frac{1}{6}, \dots, \frac{1}{6}\right)} = 6^{-n}$$

To není překvapivé.

Zajímavé: $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{12}, \frac{1}{12}, 0\right)$

Pravděpodobnost pozorování posloupnosti s těmito frekvencemi je

$$2^{-nH\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{6}, \frac{1}{12}, \frac{1}{12}, 0\right)}$$

pro n násobek 12.

Věta (Velikost $T(P)$). Pro každé $P \in \mathcal{P}_n$ máme:

$$\frac{2^{nH(P)}}{(n+1)^{|\mathcal{X}|}} \leq |T(P)| \leq 2^{nH(P)}$$

Důkaz:

Dá se spočítat přímá hodnota ale s tou se špatně manipuluje.

rozšíření P na $\mathcal{X}^n$ dle i.i.d

$$1 \geq P^n(T(P)) = \sum_{\vec{x} \in T(P)} P^n(\vec{x}) =$$

$$= \sum_{\vec{x} \in T(P)} 2^{-nH(P)} = |T(P)| \, 2^{-nH(P)}$$

$\Rightarrow$ horní odhad: $|T(P)| \leq 2^{nH(P)}$

Nyní dolní odhad:

nejdříve ukážeme, že $T(P)$ - třída má největší pravděpodobnost pro P^n tj.

$$P^n(T(P)) \geq P^n(T(\hat{P})) \quad \forall \hat{P} \in \mathcal{P}_n.$$

poměr pravděpodobností:

$$\frac{P^n(T(P))}{P^n(T(\hat{P}))} = \frac{|T(P)| \cdot \prod_{a \in \mathcal{X}} P(a)^{nP(a)}}{|T(\hat{P})| \cdot \prod_{a \in \mathcal{X}} P(a)^{n\hat{P}(a)}}$$

$$= \frac{\binom{n}{nP(a_1), nP(a_2), \dots, nP(a_{|\mathcal{X}|})} \prod_{a \in \mathcal{X}} P(a)^{nP(a)}}{\binom{n}{n\hat{P}(a_1), n\hat{P}(a_2), \dots, n\hat{P}(a_{|\mathcal{X}|})} \prod_{a \in \mathcal{X}} P(a)^{n\hat{P}(a)}}$$

$$= \prod_{a\in\mathcal{X}} \frac{(n\hat{P}(a))!}{(nP(a))!} \, P(a)^{n(P(a)-\hat{P}(a))}$$

Dů: $\frac{m!}{n!} \geq n^{m-n}$; $m \geq n$ a $m < n$

užitím tohoto

$$\frac{P^n(T(P))}{P^n(T(\hat{P}))} \geq \prod_{a\in\mathcal{X}} (nP(a))^{n\hat{P}(a)-nP(a)} \, P(a)^{n(P(a)-\hat{P}(a))}$$

$$= n^{n(\sum_{a\in\mathcal{X}} \hat{P}(a) - \sum_{a\in\mathcal{X}} P(a))}$$

$$= n^{n(1-1)} = 1.$$

Nyní

$$1 = \sum_{\hat{P}\in\mathcal{P}_n} P^n(T(\hat{P})) \leq \sum_{\hat{P}\in\mathcal{P}_n} \max P^n(T(\hat{P}))$$

$$\leq \sum_{\hat{P}\in\mathcal{P}^n} P^n(T(P)) \leq (n+1)^{|\mathcal{X}|} P^n(T(P))$$

$$= (n+1)^{|\mathcal{X}|} \sum_{\vec{x}\in T(P)} P^n(\vec{x}) = (n+1)^{|\mathcal{X}|} \sum_{\vec{x}\in T(P)} 2^{-nH(P)}$$

$$= (n+1)^{|\mathcal{X}|} |T(P)| \cdot 2^{-nH(P)}$$

BOBO

Pro binární abecedu to znamená:

$$\frac{2^{nH(\frac{k}{n})}}{(n+1)^2} \le \binom{n}{k} \le 2^{nH(k/n)}$$

Dá se vylepšit

$$\boxed{\text{Tvrzení: } \frac{2^{nH(k/n)}}{n+1} \le \binom{n}{k} \le 2^{nH(\frac{k}{n})}}$$

Důkaz: skripta) T-C s. 353

<u>Věta</u> (Pravděp. typů tříd)

Pro libovolné $P \in \mathcal{P}_n$ a libovolné rozdělení Q na $\mathcal{X}$ platí

$$\boxed{\frac{2^{-nD(P\|Q)}}{(n+1)^{|\mathcal{X}|}} \le Q^n(T(P)) \le 2^{-nD(P\|Q)}}$$

<u>Důkaz</u>: Máme

$$Q^n(T(P)) = \sum_{\vec{x} \in T(P)} Q^n(\vec{x})$$

$$= \sum_{\vec{x} \in T(P)} 2^{-n(D(P\|Q)+H(P))}$$

$$= |T(P)| \cdot 2^{-n(D(P\|Q)+H(P))}$$

užitím odhadů pro $|T(P)|$ vychází

$$\frac{2^{-nD(P\|Q)}}{(n+1)^{|X|}} \le Q^n(T(P)) \le 2^{-nD(P\|Q)}$$

$\boxed{\text{důsledek: při velké divergenci } P \text{ a } Q \text{ a } T(P) \text{ málo pravděpodobná}}$

souhrn:

$$|\mathcal{P}_n| \le (n+1)^{|X|}$$

$$Q^n(\vec{x}) = 2^{-n(D(P_{\vec{x}}\|Q)+H(P_{\vec{x}}))}$$

$$|T(P)| \doteq 2^{nH(P)}$$

$$Q^n(T(P)) \doteq 2^{-nD(P\|Q)}$$

aplikace: univerzální kódy
Sanov's theorem
Testování hypotéz
Limitní věty

Zákon velkých čísel (diskrétní varianta)

$\varepsilon > 0$

Typická množina posloupností:

$$T_Q^{\varepsilon} = \{ \vec{x} \in X^n \mid D(P_{\vec{x}} \| Q) \leq \varepsilon \}$$

má velkou pravděpodobnost v Q^n pro velké n

$$1 - Q^n(T_Q^{\varepsilon}) =$$

$$= \sum_{\{P : D(P\|Q) > \varepsilon\}} Q^n(T(P))$$

$$\leq \sum_{\{P \mid D(P\|Q) > \varepsilon\}} 2^{-n\varepsilon}$$

$$\leq (n+1)^{|X|} 2^{-n\varepsilon}$$

$$= 2^{-n\left(\varepsilon - |X| \frac{\log(n+1)}{n}\right)} \xrightarrow[\text{pro } n \to \infty]{} 0$$

vede k

Věta:

$$\text{Pravděp}\left(D(P_{\vec{x}} \| Q) > \varepsilon\right\} \leq 2^{-n\left(\varepsilon - |X| \frac{\log(n+1)}{n}\right)}$$

$\downarrow n \to \infty$

fr ve pozorujeme skutečnost 0