

# PST

## 10. Způsoby popisu rozdělení náhodných veličin a vektorů. Odhady parametrů rozdělení. Základní statistické testy. Markovské řetězce a jejich asymptotické vlastnosti.

---

[http://cmp.felk.cvut.cz/~navara/stat/PMS\\_ebook.pdf](http://cmp.felk.cvut.cz/~navara/stat/PMS_ebook.pdf)

<http://cmp.felk.cvut.cz/~navara/stat/>

---

[https://cs.wikipedia.org/wiki/Rozd%C4%9Blen%C3%AD\\_pravd%C4%9Bpodobnosti](https://cs.wikipedia.org/wiki/Rozd%C4%9Blen%C3%AD_pravd%C4%9Bpodobnosti) - dobrý shrnutí popisu náhodných veličin a vektorů

Testování hypotéz - příklad

[https://youtu.be/VMRZO\\_yYWRA](https://youtu.be/VMRZO_yYWRA)

## Základní pojmy pravděpodobnosti

### Laplaceova (klasická) pravděpodobnost

- **Náhodný pokus** má  $n \in \mathbb{N}$  různých, vzájemně se vylučujících výsledků, které jsou stejně možné.
- **Elementární jevy** = výsledky náhodného pokusu
- **Množina všech elementárních jevů**:  $\Omega$
- **Jev** je podmnožina všech elementárních jevů ( $A \subseteq \Omega$ )
- **Pravděpodobnost jevu**  $A$ :

$$P(A) = \frac{|A|}{|\Omega|}$$

(Prostě počet správných výsledků děleno počet všech možných)

- **Jevové pole**: všechny jevy pozorovatelné v náhodném pokusu, zde  $\exp \Omega$  (=množina všech podmnožin množiny  $\Omega$ )

### Kolmogorovova pravděpodobnost

- Elementárních jevů (=prvků množiny  $\Omega$ ) může být nekonečně mnoho, nemusí být stejně pravděpodobné
- **Jevy** jsou podmnožiny množiny  $\Omega$ , ale ne nutně všechny. Tvoří podmnožinu  $A \subseteq \exp \Omega$ , která splňuje podmínky  $\sigma$ -algebry (viz. níže).
- **Pravděpodobnost** není určena strukturou jevů jako u Laplaceova modelu, je to funkce  $P: A \rightarrow \langle 0, 1 \rangle$ , splňující podmínky:
- $(P1) P(\Omega) = 1$ ,
- $(P2) P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} P(A_n)$ , pokud jsou množiny (=jevy)  $A_n$ ,  $n \in \mathbb{N}$ , po dvou neslučitelné
- **Pravděpodobnostní prostor** je trojice  $(\Omega, A, P)$ , kde  $\Omega$  je neprázdná množina,  $A$  je  $\sigma$ -algebra podmnožin množiny  $\Omega$  a  $P$  je pravděpodobnost.

### $\sigma$ -algebra

$\sigma$ -algebra je teoretický koncept výběru jistých podmnožin dané množiny, který splňuje pevně definované podmínky. Koncept  $\sigma$ -algebry umožňuje například zavést míru, čehož se dále využívá zejména v matematické analýze k budování pojmu integrál a právě v teorii pravděpodobnosti [wikipedia]. Systém podmnožin  $A$  nějaké množiny  $\Omega$  musí splňovat podmínky:

1.  $\emptyset \in A$
2.  $A \in A \Rightarrow A^c \in A$  (uzavřenost vůči doplňku)
3.  $(\forall n \in \mathbb{N} : A_n \in A) \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in A$  (uzavřenost vůči sjednocení)

Nejmenší  $\sigma$ -algebra podmnožin  $R$ , která obsahuje všechny intervaly, se nazývá **Borelova  $\sigma$ -algebra**. Obsahuje všechny intervaly otevřené, uzavřené i polouzavřené, i jejich spočetná sjednocení, a některé další množiny, ale je menší než  $\exp R$ . Její prvky nazýváme borelovské množiny.

## Nezávislost náhodných jevů

Pokud jsou jevy  $A$  a  $B$  nezávislé pak platí:

$$P(A \cap B) = P(A) \cdot P(B).$$

$$P(A|B) = P(A).$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

# Podmíněná pravděpodobnost

$$P(A) = P(B) * P(A|B) + P(\neg B) * P(A|\neg B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

## Bayesova věta

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

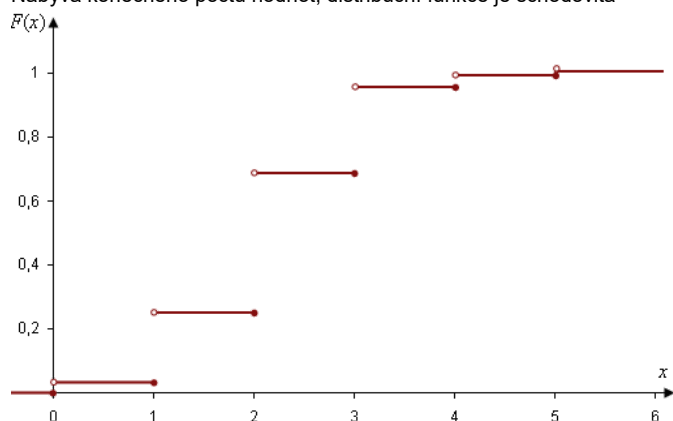
$$P(A \cap B) = P(A|B) * P(B) = P(B|A) * P(A)$$

## Náhodná veličina (NV)

Měřitelná f-ce, její chování popisuje neklesající distribuční  $F_x(t) = P(X \leq t)$ . Přiděluje hodnoty od 0 do 1. Pravděpodobnostní fce  $p_x(t) = P(X = t)$ .

### Diskrétní

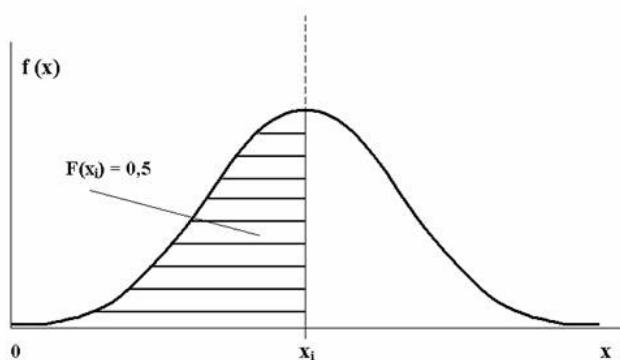
Nabývá konečného počtu hodnot, distribuční funkce je schodovitá



### Spojité

Nekonečný počet hodnot,  $P(X = t) = 0$

$F_x(t) = \int_{-\infty}^t f(t)dt$ , kde  $f(t)$  je funkce udávající hustotu.



## Smíšená NV

Směs diskrétní a spojitě

$$Mix_c(D, S), c \in (0, 1)$$

Platí:

$$F_x(t) = cF_D(t) + (1 - c)F_S(t)$$

$$p_x(t) = cp_D(t) + (1 - c)p_S(t)$$

## Střední hodnota

Průměrná hodnota NV, nemusí existovat

Diskrétní -  $\sum t * P(t)$

Spojité -  $\int_{-\infty}^{\infty} t * f_x(t)dt$

- Je to číslo

- $E(x + y) = E(x) + E(y)$  //platí i pro minus
- $E(rX) = r * E(X)$
- $E(r + X) = r + E(X)$

## Rozptyl

Jak moc jsou hodnoty odchýlené od průměru  $E(x)$

$$D(x) = \sigma_x^2 \text{ (směrodatná odchylka)}$$

$$D(x) = E((x - E(x))^2) = E(x^2) - E^2(x)$$

$$E(x^2) = E^2(x) + D(x)$$

$$\text{Diskrétní - } E(x^2) = \sum k^2 P(X = k)$$

$$\text{Spojité - } E(x^2) = \int_{-\infty}^{\infty} t^2 * f_x(t) dt$$

- $D(x) \geq 0$
- $D(x + r) = D(x)$
- $D(r) = 0$
- $D(xr) = D(x) * r^2$

## Směrodatná odchylka

Průměrná odchylka od průměru

- $\sigma_x \geq 0$
- $\sigma_r = 0$
- $\sigma_x * r = |r| \sigma_x$
- $\sigma_x + r = \sigma_x$

## Normovaná NV

$$\text{norm } X = \frac{X - E(X)}{\sigma_x}$$

## Momenty náhodných veličin

Pro náhodnou veličinu N je k-tý moment definován jako  $E(X^k)$

1. moment = střední hodnota
2. moment = rozptyl + střední hodnota nadruhou

## Základní typy diskrétních rozdělení

### Rovnoměrné

m možných výsledků, kde každý má stejnou pravděpodobnost, hod kostkou

$$E(x) = (A + B)/2, \text{ kde A je nejmenší hodnota a B největší, } D(x) = (b - a)^2/12$$

### Alternativní

2 možné hodnoty (pst q a (1-q)), hod mincí

$$E(x) = P(1) = q, D(x) = q(1 - q)$$

### Binomické

Vyjadřuje počet úspěchů v sérii m pokusů, každý pokus má pst úspěchu q, počet orlů při hodu 20 mincemi

$$E(x) = m * q, D(x) = mq(1 - q)$$

### Poissonovo

počet událostí během intervalu délky t,  $\lambda$  průměrný počet za jednotku času (třeba 5 hovorů za hodinu)

$$E(x) = \lambda = D(x)$$

### Geometrické

Počet neúspěchů než nastane úspěch, pokaždé je stejná pst úspěchu "p", hod mincí dokud nespadne orel

$$E(x) = (1 - p)/p, D(x) = (1 - p)/p^2$$

## Hypergeometrické

Například "M" losů z nichž "J" vyhrává, tak udává počet výherních losů, z vytažených "S" losů

$$E(x) = \frac{S \cdot J}{M}, D(x) = \frac{S \cdot J \cdot (M - J) \cdot (M - S)}{M^2 \cdot (M - 1)}$$

## Základní typy spojitých rozdělení

### Rovnoměrné

Stejně dlouhé intervaly mají stejnou pst

$$E(x) = (A + B)/2, \text{ kde } A \text{ je nejmenší hodnota a } B \text{ největší, } D(x) = (b - a)^2/12$$

### Normální

$$E(x) = \mu, D(x) = \sigma^2$$

### Exponenciální

doba čekání mezi 2 následnými výskyty událostí

### Logaritmicko normální

existuje 👍

## Náhodné vektory

n-rozměrná náhodná veličina, měřitelná funkce  $X: \Omega \rightarrow R^n$

Používáme ho v případech, kdy je k popisu výsledku náhodného pokusu nutné použít více čísel.

**Příklad:** Chceme popsat vztah např. mezi výškou a váhou osob. K tomu potřebujeme více informací než jen popis jednotlivých náhodných veličin.

#### Kovariance

- určuje míru statistické závislosti mezi náhodnými veličinami
- $cov(X, Y) = E(XY) - E(X)E(Y)$
- $cov(X, X) = D(X)$
- $cov(Y, X) = cov(X, Y)$
- $cov(aX + b, cY + d) = a * c * cov(X, Y)$
- $cov(X, Y) = 0$  pro nezávislé veličiny

#### Korelace

- je to kovariance pro normované náhodné veličiny
- $\rho(X, Y) = cov(normX, normY)$
- Korelace nabývá hodnot  $\in (-1, 1)$
- Pro  $\rho = 0$  říkáme, že jsou veličiny **nekorelované**

pro vektory platí  $D(X + Y) = DX + DY + 2cov(X, Y)$

## Čebyševova nerovnost

Čebyševovou nerovností I. typu označujeme tvrzení, že pro libovolnou nezápornou náhodnou veličinu  $X$  se střední hodnotou  $E(X)$  je pravděpodobnost, že veličina  $X$  nabude alespoň hodnoty  $\varepsilon$  dána podmínkou

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

pro všechna  $\varepsilon > 0$ . (Tato nerovnost se někdy v literatuře označuje jako Markovova.)

Pro libovolnou náhodnou veličinu  $X$  se střední hodnotou  $E(X)$  a rozptylem  $D(X)$  je pravděpodobnost, že absolutní hodnota  $|X - E(X)|$  nabude hodnoty menší než libovolné  $\varepsilon > 0$  omezena Čebyševovou nerovností II. typu

$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}$$

nebo také

$$P(|X - E(X)| < \varepsilon \cdot \sigma) \geq 1 - \frac{1}{\varepsilon^2}$$

$$\text{kde } \sigma = \sqrt{D(X)}$$

# Centrální limitní věta

Pokud platí předpoklady centrální limitní věty, tak výběrový průměr má jakožto náhodná veličina *asymptoticky normální rozdělení*. Jinak řečeno - rozdělení součtu  $n$  nezávislých NV se stejným rozdělením se blíží k normálnímu rozdělení.

**Předpoklady:**

- Stejné rozdělení
- Nezávislé NV
- $D(x) \neq 0$

$$P(X \leq a) = P(\text{norm}(X) \leq \frac{a - EX}{\sqrt{DX}})$$

pak lze vyčíst hodnotu normálního rozdělení pro  $y \leq \frac{a - EX}{\sqrt{DX}}$  z tabulek a máme výsledek.

## Základ statistiky

Výběrový soubor rozsahu  $n$  označujeme jako náhodný výběr  $X = (X_1, \dots, X_n)$ .

**Statistika** je každá měřitelná funkce, definovaná na náhodném výběru libovolného rozsahu. Následující statistiky se používají pro odhady číselných parametrů.

**Empirické rozdělení**

**Výběrový průměr**

z náhodného výběru  $X = (X_1, \dots, X_n)$  je nestranný konzistentní odhad střední hodnoty:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

**Výběrový rozptyl**

náhodného výběru  $X = (X_1, \dots, X_n)$  je nestranný konzistentní odhad střední rozptylu:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

## Obecné vlastnosti odhadů parametrů. Odhady střední hodnoty, rozptylu, směrodatné odchylky, momentů.

Snaha o odhad neznámých veličin a jejich parametrů v závislosti na pozorovaných datech.

Značení:

$\vartheta$  ... jakákoli hodnota parametru (reálné číslo),

$\vartheta^*$  ... skutečná (správná) hodnota parametru (reálné číslo),

$\hat{\Theta}, \hat{\Theta}_n$  ... odhad parametru založený na náhodném výběru rozsahu  $n$  (náhodná veličina)

$\hat{\vartheta}, \hat{\vartheta}_n$  ... realizace odhadu (reálné číslo)

Žádoucí vlastnosti odhadů:

- $E\hat{\Theta}_n = \vartheta^*$ , tj.  $E(\hat{\Theta}_n - \vartheta^*) = 0$  **nestranný** (opak: **vychýlený**)
- $\lim_{n \rightarrow \infty} E\hat{\Theta}_n = \vartheta^*$ , tj.  $\lim_{n \rightarrow \infty} E(\hat{\Theta}_n - \vartheta^*) = 0$  **asymptoticky nestranný**
- **eficientní** = s malým rozptylem, což posuzujeme podle  $E((\hat{\Theta}_n - \vartheta^*)^2) = D\hat{\Theta}_n + (E\hat{\Theta}_n - \vartheta^*)^2$ , pro nestranný odhad se redukuje na  $D\hat{\Theta}_n$
- **nejlepší nestranný** odhad je ze všech nestranných ten, který je nejvíce eficientní (mohou však existovat více eficientní vychýlené odhady)
- $\lim_{n \rightarrow \infty} E(\hat{\Theta}_n - \vartheta^*) = 0$ ,  $\lim_{n \rightarrow \infty} \sigma_{\hat{\Theta}_n} = 0$  **konzistentní**
- **robustní**, tj. odolný vůči šumu („i při zašuměných datech dostáváme dobrý výsledek“) – přesné kritérium chybí, ale je to velmi praktická vlastnost

## Odhady parametrů metodou momentů

Metoda momentů je založena na rovnosti výběrových momentů a momentů rozdělení. Nelze jednoznačně rozhodnout, která z metod dává lepší výsledky. Rozhodování provádíme podle konkrétní situace, nejčastěji rozhoduje jednoduchost získaných vzorců. Metoda momentů zohledňuje všechna data z výběru a volíme ji v případech, kdy je soustava věrohodnostních rovnic obtížně řešitelná.

Pro základní rozdělení dávají obě metody shodné výsledky a v případě složitějších rozdělení můžeme jako další kritérium uvažovat, které vzorce jsou méně citlivé na zavlečené chyby do hodnot výběru.

Porovnáme teoretické  $k$ -té momenty  $E(x^k)$  s jejich odhady

Z náhodného výběru lze spočítat **výběrový  $k$ -tý obecný moment**:

$$m_{X^k} = \frac{1}{n} \sum_{j=1}^n x_j^k.$$

Pro  $k = 1$  to je  $EX$

Pro  $k = 2$  to je  $DX + E(x)^2$

## Odhady parametrů metodou maximální věrohodnosti.

### Definice:

Nechť náhodný výběr  $X_1, \dots, X_n$  má rozdělení určené sdruženou hustotu

$$f(\mathbf{x}; \theta) = \prod_{i=1}^n f_{\theta}(x_i) \quad \text{pro spojité rozdělení, resp.}$$

$$p(\mathbf{x}; \theta) = \prod_{i=1}^n \mathbb{P}_{\theta}(X_i = x_i) \quad \text{pro diskrétní rozdělení.}$$

Při pevné hodnotě  $\mathbf{x} = (x_1, \dots, x_n)$  se funkce  $f(\mathbf{x}; \theta)$  resp.  $p(\mathbf{x}; \theta)$  jakožto funkce  $\theta$  nazývá **věrohodnostní funkce** a značí  $L(\theta; \mathbf{x})$ .

Obvykle je vhodné nehledat maximum věrohodnostní funkce přímo, ale zlogaritmovat ji. Součiny se nám změň na součty ale na výsledek to nebude mít vliv.

Myšlenkou této metody je, že hledáme takové hodnoty parametrů, které by nejlépe vysvětlovali realizaci náhodného výběru, tj. při kterých by pozorované výsledky byly "nejméně nepravděpodobné".

Pravděpodobnost je určena pro předpovídání výsledků budoucího pokusu při známém pravděpodobnostním modelu a tudíž má definiční obor jevy z nějaké  $\sigma$ -algebry. Naopak věrohodnost vyhodnocuje sérii pokusů již realizovaných a dovoluje jim přizpůsobit neznámé parametry pravděpodobnostního modelu, a proto je definována na prostoru  $\Pi$  všech možných hodnot parametrů rozdělení.

**Metoda maximální věrohodnosti** považuje za správný odhad parametrů takové hodnoty, které maximalizují věrohodnost.

**Řešení:** Alternativního rozdělení je popsáno pravděpodobnostmi

$$f_{Alt}(X = 1) = p$$

a

$$f_{Alt}(X = 0) = 1 - p,$$

kde  $p$  je pravděpodobnost úspěchu (jevu 1).

Věrohodnostní funkce pak je součin pravděpodobností jednotlivých realizací za dané hodnoty parametru  $p$ :

$$\begin{aligned} L(\{0, 0, 1\}, p) &= \prod_{i=1}^3 f_{Alt}(x_i | p) \\ &= (1 - p) \cdot (1 - p) \cdot p \\ &= (1 - p)^2 p. \end{aligned}$$

Tuto funkci bychom mohli již přímo maximalizovat (to by vedlo ke kvadratické rovnici), nebo můžeme podobně jako dříve tuto funkci logaritmovat a úlohu si zjednodušit.

Logaritmická věrohodnostní funkce pak je

$$l(\{0, 0, 1\}, p) = 2\ln(1 - p) + \ln(p)$$

a její derivace podle  $p$  položíme rovnu 0:

$$\frac{l(\{0, 0, 1\}, p)}{\partial p} = 2 \frac{-1}{1 - p} + \frac{1}{p} = 0$$

a po úpravě (vynásobení  $p(1 - p)$ )

$$\begin{aligned} 2\hat{p} &= (1 - \hat{p}) \\ 3\hat{p} &= 1 \\ \hat{p} &= \frac{1}{3} \end{aligned}$$

Odhadem parametru  $p$  metodou max. věrohodnosti je tedy  $\hat{p} = 1/3$ .

## Intervalové odhady.

Dosud jsme skutečnou hodnotu parametru  $\vartheta^*$  nahrazovali **bodovým odhadem**  $\hat{\Theta}$  (což je náhodná veličina). Nyní místo toho hledáme **intervalový odhad**, tzv. **interval spolehlivosti**.

**vosti**  $I$ , což je minimální interval takový, že

$$P(\vartheta^* \in I) \geq 1 - \alpha,$$

kde  $\alpha \in (0, 1)$  je pravděpodobnost, že meze intervalu  $I$  budou překročeny;  $1 - \alpha$  je **koefficient spolehlivosti**. Obvykle hledáme **horní**, resp. **dolní jednostranný** odhad,

$$I = (-\infty, q_{\hat{\Theta}}(1 - \alpha)), \text{ resp. } I = (q_{\hat{\Theta}}(\alpha), \infty),$$

nebo (**symetrický**) **oboustranný** odhad,

$$I = (q_{\hat{\Theta}}(\frac{\alpha}{2}), q_{\hat{\Theta}}(1 - \frac{\alpha}{2})).$$

Jde vlastně o odhad kvantilů.

K tomu potřebujeme znát rozdělení odhadu  $\hat{\Theta}$ .

Symetrický odhad ve smyslu, že se stejnou pst ho překročíme nahoru nebo dolů (nemusí být vždy stejná vzdálenost od střední hodnoty).

## Princip statistického testování hypotéz.

Na základě testu chceme klasifikovat objekty (nebo skupiny objektů) na zdravé/nemocné, střízlivé/opilé, nevinné/vinné apod.

- **Nulová hypotéza:** objekt je „normální“, „negativní“; výsledkem testu je náhodná veličina  $T$ .
- **Alternativní hypotéza:** objekt je „anomální“, „pozitivní“; výsledkem (též) testu je náhodná veličina  $T^*$ .

**Příklad:** Máme považovat test nemoci za pozitivní?

**Nulová hypotéza  $H_0$ :** Člověk není nakažený.

**Alternativní hypotéza  $H_1$ :** Člověk je nakažený.

Vždy se testuje na nějaké hladině významnosti  $\alpha$ .

## Testy střední hodnoty a rozptylu, porovnání dvou rozdělení

$\bar{x} = EX$

Test střední hodnoty -  $t = \frac{\bar{x}-c}{\sigma} * \sqrt{n}$ , kde  $c$  je odhad  $EX$  podle hypotézy,  $n$  stupně volnosti, pokud neznáme odchylku tak ji nahradíme odhadem směrodatné odchylky.

Test rozptylu -  $t = \frac{(n-1)*S_x^2}{c}$ , kde  $c$  je odhad  $\sigma^2$  podle hypotézy,  $n$  stupně volnosti

## $\chi^2$ -test

Slouží k testování hypotézy, že náhodná veličina má předpokládané rozdělení. Protože umíme hypotézy jen zamítnat, nikdy nepotvrdíme, že takové rozdělení opravdu má.

Testujeme **diskrétní rozdělení** (mohlo vzniknout diskretizací spojitého).

$H_0$  : Náhodná veličina má diskrétní rozdělení do  $k$  tříd s nenulovými pravděpodobnostmi  $p_1, \dots, p_k$ .

Testujeme pomocí realizace náhodného výběru rozsahu  $n$ . Není důležité pořadí výsledků, pouze jejich **četnosti**  $N_i$ , resp. **realizace četností**  $n_i$  (**empirické četnosti**) nebo **realizace relativních četností**  $\frac{n_i}{n}$  ( $i = 1, \dots, k$ ). Porovnáváme je s **teoretickými četnostmi**  $n p_i$ .

## $\chi^2$ -test dobré shody

$H_0$  : Dva náhodné výběry pocházejí ze stejného diskrétního rozdělení.

Rozsahy výběrů jsou  $m, n$ , četnosti výsledků  $M_i, N_i$ , jejich realizace  $m_i, n_i$  ( $i = 1, \dots, k$ ).

Předpokládáme rozdělení s neznámými teoretickými pravděpodobnostmi  $p_i$  ( $i = 1, \dots, k$ ).

## Test nezávislosti v kontingenční tabulce.

Je to  $\chi^2$ -test nezávislosti dvou rozdělení

$H_0$  : Dvě diskrétní náhodné veličiny (jejichž rozdělení neznáme) jsou nezávislé.

$X$  nabývá  $k$  hodnot s pravděpodobnostmi  $p_1, \dots, p_k$ ,

$Y$  nabývá  $m$  hodnot s pravděpodobnostmi  $q_1, \dots, q_m$ .

Realizace dvojrozměrného náhodného výběru  $((x_1, y_1), \dots, (x_n, y_n))$  obsahuje dvojice realizací náhodných veličin  $X, Y$ ; potřebujeme pouze četnosti  $N_{ij}$ , resp. jejich realizace  $n_{ij}$  ( $i = 1, \dots, k$ ;  $j = 1, \dots, m$ ). Ty bývají uspořádány do tzv. **kontingenční tabulky**. Počet tříd je  $k m$ .

## Markovské řetězce a jejich asymptotické vlastnosti

[https://cs.wikipedia.org/wiki/Markov%C5%AFv\\_%C5%99et%C4%9Bzec](https://cs.wikipedia.org/wiki/Markov%C5%AFv_%C5%99et%C4%9Bzec)

$P \rightarrow$  Matice přechodu - je stochastická - součty řádků je 1

Vektor je sloupcový  $\rightarrow$  Wernerovo náboženství.

## Klasifikace stavů

- Nedosažitelné - do takového stavu se nedostaneme nikdy
- Absorpční - z takového stavu se nelze dostat
- Dosažitelné - lze do něj dojít s nenulovou pst
- Trvalý - pokud z něj odejdeme je pst 1, že se sem v budoucnu vrátíme, opak je stav přechodný
- Komunikují - lze z 1. se dostat do 2. a zároveň naopak (silně souvislý graf)



- Periodické

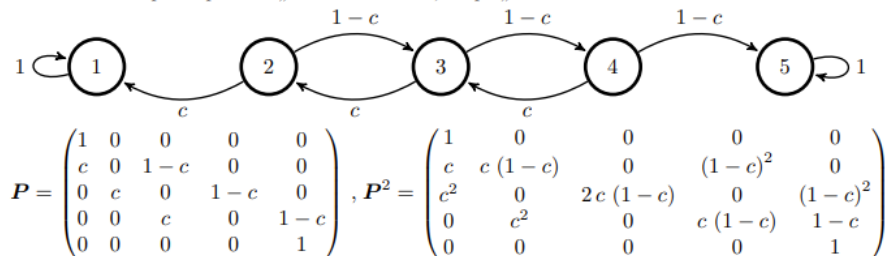
**Perioda** stavu  $i$  je největší společný dělitel všech čísel  $t$ , pro která  $p_{ii}(t) > 0$ , tj. největší číslo  $t$  takové, že  $p_{ii}(u) = 0$  pro všechna  $u$ , která nejsou násobky  $t$ .

- ergodický - trvalý a neperiodický

## Absorpční řetězce

Absorpční řetězce jsou takové řetězce, které obsahují mimo stavy přechodný i stavy absorpční. Tzn., že pravděpodobnost setrvání v takovém stavu je rovna 1.

**Řešení.** Jde opět o příklad „shoda v tenisu“, resp. „basketbal“.



## Asymptotické chování Markovových řetězců

Do jakého stavu se konverguje po nekonečně mnoha krocích  $t$ .

Přechodové stavy - pst konverguje k 0

Vždy skončí v některém trvalém stavu jejich součet musí být 1.

$$P = \left( \begin{array}{c|c} I_i & 0 \\ \hline R & Q \end{array} \right) \begin{array}{l} R \dots \text{z přechod. do} \\ \text{trvalých} \\ Q \dots \text{mezi přechod.} \end{array}$$

$$F = (I - Q)^{-1} \cdot R$$

## Rozložitelné a nerozložitelné řetězce

Asi je tím myšleno, že rozložitelné lze rozdělit na více silných komponent souvislosti, které se dají řešit separátně, kdežto nerozložitelné mají pouze jednu komponentu silné souvislosti + absorpční stavy.