Pravděpodobnost a statistika

Hynek Lavička[1]

[1]Katedra fyziky
Fakulta jaderná a fyzikálně inženýrská
České vysoké učení technicé v Praze

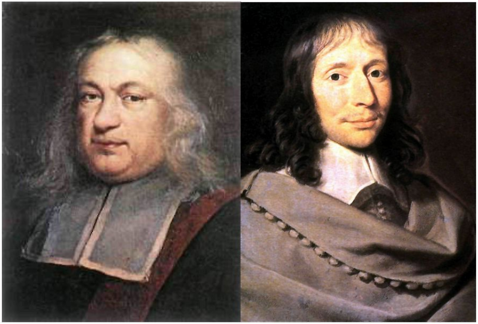
26.3.2007

Obsah

Outline

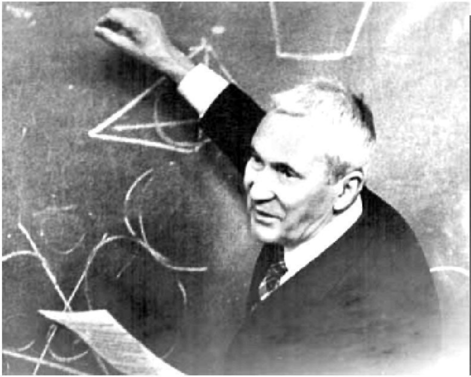
Historický úvod - kořeny

Teorie pravděpodobnosti má své kořeny v práci Pierre de Fermata a Blaise Pascala ze 17. století, kdy se snažili analyzovat možnosti výhry v hazardních hrách.



Historický úvod - moderní teorie

Moderní teorie pravděpodobnosti je založena na práci Andreje Nikolajeviče Kolmogorova a Richarda von Misese.



Úvod

Ústředním termínem, který se zde vyskytuje je náhodná proměnná a stochastický proces, které rozšiřují pochopení chování systémů, které nemáme plně pod kontrolou, které mohou být nezávislé na čase nebo se v čase vyvíjet. Teorie pravděpodobnosti je základem pro většinu lidských aktivit, které vyžadují kvalitativní analýzu velkých souborů dat - statistiku. Teorie pravděpodobnosti se stala v 20. století základem pro

- Statistickou fyziku
- Kvantovou mechaniku a její rozšíření

Zákon velkých čísel

Jestliže je událost ω s pravděpodobností $\mathbb{P}(\omega)$ pozorována opakovaně a pozorování jsou nazávislá, potom poměr pozorovananých událostí a celkového počtu pozorování konverguje k $\mathbb{P}(\omega)$, když počet opakování roste nad všechny meze.

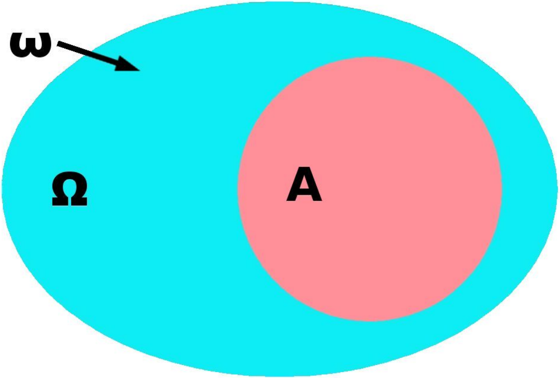
$$\mathbb{P}(\omega) = \lim_{N \to +\infty} \frac{N_\omega}{N} \tag{1}$$

Outline

Outline

Událost

Elementární událost ω je jednoprvková podmnožina množiny výsledků nějakého jevu Ω. Událost A je podmnožina množiny výsledků nějakého jevu Ω.



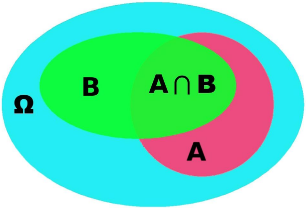
Událost

Máme 2 nezávislé události A, B potom pravděpodobnost konjunkce je

$$\mathbb{P}(A \bigcap B) = \mathbb{P}(A)\mathbb{P}(B) \tag{2}$$

Máme 2 události A, B potom podmíněná pravděpodobnost je

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \bigcap B)}{\mathbb{P}(B)} \tag{3}$$



Outline

Náhodná proměnná

Proměnná, která nabývá náhodné (nepredikovatelné) hodnoty (událost) z množiny možných hodnot Ω

$$X : \Omega \rightarrow \mathbb{R} \tag{4}$$

Typy náhodných proměnných podle vlastností Ω

- Diskrétní
- Spojitá

Outline

Stochastický proces

Stochastický proces je množina náhodných proměnných. Pokud množina má strukturu 1-D prostoru, pak se nazývá časová řada v ostatních případech náhodné pole.

$$F : T \rightarrow X \tag{5}$$

Typy stochastických procesů (časových řad) podle vlastností Ω

- Diskrétní
- Spojitá

Outline

Charakteristiky náhodných proměnných

- Kumulativní distributivní funkce (distributivní funkce)

$$F(x) = \mathbb{P}(X(\omega) \leq x) \tag{6}$$

- Pravěpodobnostní hustota

$$\rho(x) = \frac{dF}{dx} \tag{7}$$

- Střední hodnota

$$\mu = \int_{-\infty}^{+\infty} x\rho(x)\, dx \tag{8}$$

$$\mu = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(X(\omega)) \tag{9}$$

Charakteristiky náhodných proměnných

- Variance

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 \rho(x)\; dx \tag{10}$$

$$\sigma^2 = \sum_{\omega \in \Omega} (X(\omega)-\mu)^2 \mathbb{P}(X(\omega)) \tag{11}$$

Charakteristiky náhodných 2 proměnných - korelace

Nechť máme náhodné proměnné X a Y

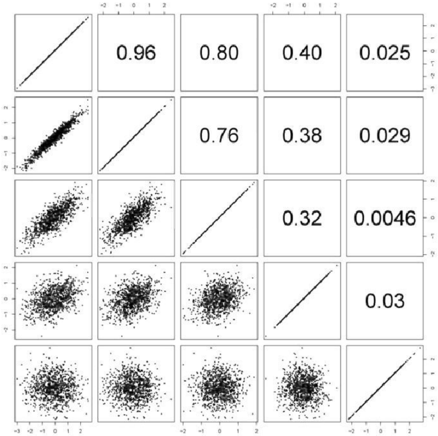
- Korelační funkce

Měří míru závislosti 2 náhodných proměnných.

$$\rho_{X,Y} = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_x)(y-\mu_y)\rho(x,y)\ dx\ dy}{\sigma_x\sigma_y} \tag{12}$$

$$\rho_{X,Y} = \frac{\sum_{\omega_X\in\Omega_X}\sum_{\omega_Y\in\Omega_Y}(X(\omega_X)-\mu_x)(Y(\omega_Y)-\mu_y)\mathbb{P}(X(\omega_X),Y(\omega_Y))}{\sigma_x\sigma_y} \tag{13}$$

Charakteristiky náhodných 2 proměnných - korelace



Součet náhodných proměnných

Nechť máme nezávislé náhodné proměnné X a Y

$$Z = X + Y \tag{14}$$

$$\rho_Z(z) = \int_{-\infty}^{+\infty} \rho_X(x)\rho_Y(z-x)\ dx \tag{15}$$

$$\mathbb{P}(z) = \sum_{\omega_X \in \Omega_X} \sum_{\omega_Y \in \Omega_Y} \mathbb{P}(\omega_X)\mathbb{P}(\omega_Y)\delta_{X(\omega_X)+Y(\omega_Y),z} \tag{16}$$

Outline

Outline

Binomické rozdělení

Pravděpodobnost k úspěchů při ano/ne experimentech v sérii N pokusů, kde každý získaný úspěch má pravděpodobnost p.

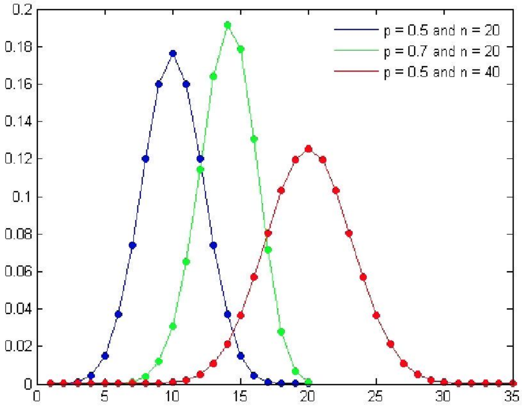
$$\Omega = \{0, 1, 2, \ldots, N\} \tag{17}$$

$$\mathbb{P}(k) = \binom{N}{k} p^k (1-p)^{N-k} \tag{18}$$

$$\mu = Np \tag{19}$$

$$\sigma^2 = Np(1-p) \tag{20}$$

Binomické rozdělení



Hypergeometrické rozdělení

Máme dodávku N objektů, ale D je defektních. Hypergeometrické rozdělení popisuje pravděpodobnost, že ve vzorku n objektů je k defektních objektů.

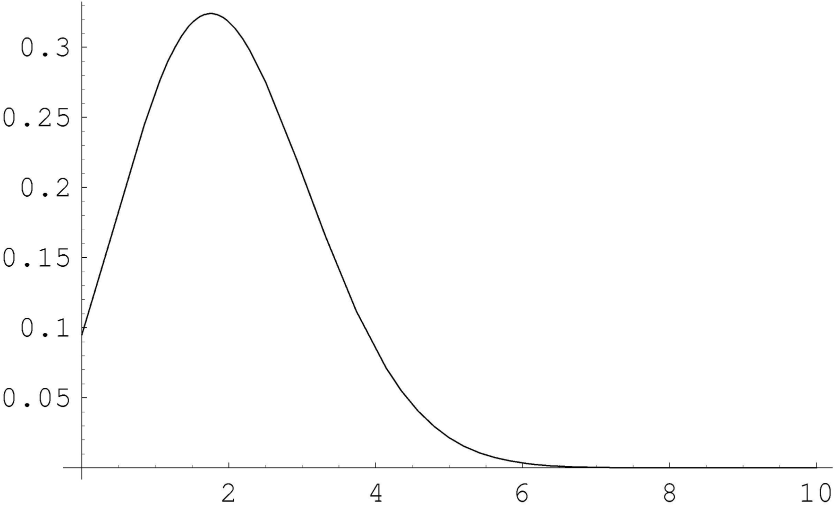
$$\Omega = \{0, 1, 2, \ldots, n\} \tag{21}$$

$$\mathbb{P}(k) = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}} \tag{22}$$

$$\mu = \frac{nD}{N} \tag{23}$$

$$\sigma^2 = \frac{n\frac{D}{N}(1-\frac{D}{N})(N-n)}{N-1} \tag{24}$$

Hypergeometrické rozdělení



Poissonovo rozdělení

Rozdělení vyjadřuje pravděpodobnost počtu úspěchů, které se stanou v pevném časovém úspěchu, když uspěchy se objeví se známým průměrným tempem.

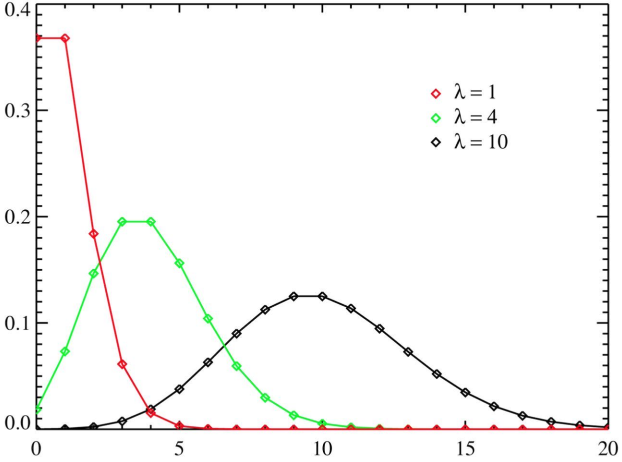
$$\Omega = \mathbb{N}_0 \tag{25}$$

$$\mathbb{P}(k) = \frac{\exp(-\lambda)\lambda^k}{k!} \tag{26}$$

$$\mu = \lambda \tag{27}$$

$$\sigma^2 = \lambda \tag{28}$$

Poissonovo rozdělení



Outline

Gaussovo (normální) rozdělení

Popisuje fluktuaci měřené veličiny.

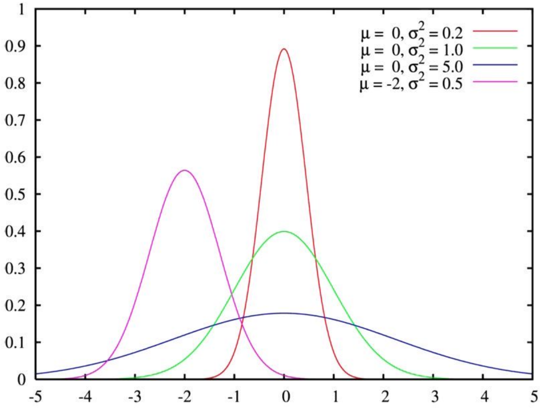
$$\Omega = \mathbb{R} \tag{29}$$

$$\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) \tag{30}$$

$$\mu = \mu \tag{31}$$

$$\sigma^2 = \sigma^2 \tag{32}$$

Gaussovo (normální) rozdělení



Gibbs-Boltzmanovo rozdělení (exponenciální rozdělení)

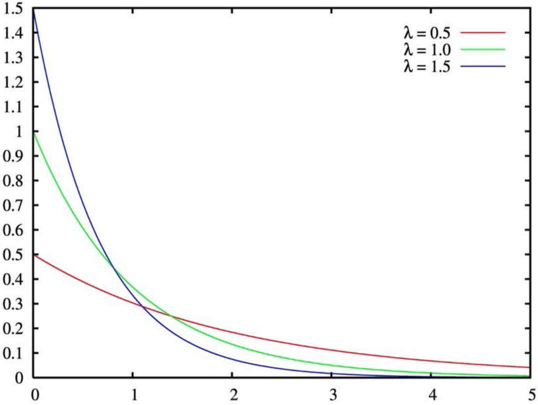
$$\Omega = \mathbb{R}_0^+ \tag{33}$$

$$\rho(x) = \lambda \exp(-\lambda x) \tag{34}$$

$$\mu = \frac{1}{\lambda} \tag{35}$$

$$\sigma^2 = \frac{1}{\lambda^2} \tag{36}$$

Gibbs-Boltzmanovo rozdělení (exponenciální rozdělení)



Léviho stabilní rozdělení

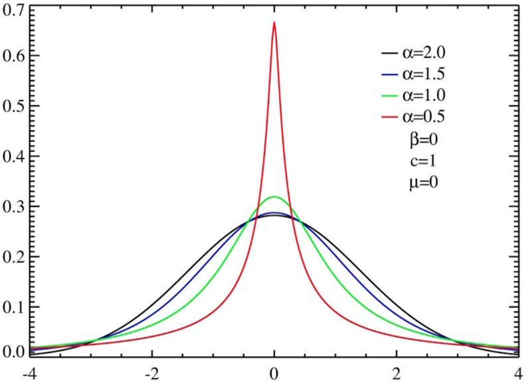
Popsáno μ a α

$$\Omega = \mathbb{R} \tag{37}$$

$$\mu = \begin{cases} \alpha > 1 & \mu \\ \alpha \leq 1 & +\infty \end{cases}$$

$$\sigma^2 = \begin{cases} \alpha = 2 & 2 \\ \alpha \neq 2 & +\infty \end{cases}$$

Léviho stabilní rozdělení



Outline

Outline

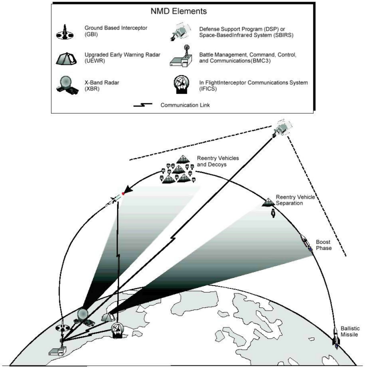
Příklad stochastického procesu

Hra házení kuliček na cíl, přičemž hráč postupuje od jednoho hřiště ke druhému.

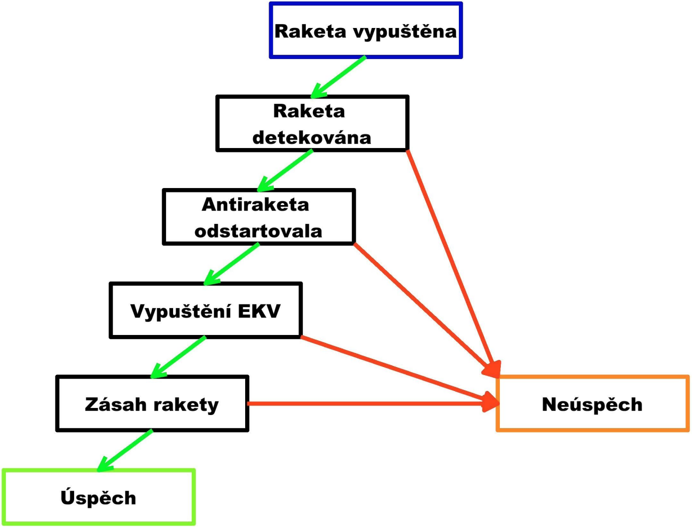
GMD

GMD (Groundbased mid-course defence) je systém proti balistickým raketám.

The NMD Concept of Operations



GMD - schéma simulace



Outline

Statistika nudná věda?

Systém	Plánovaná efektivita	Uskutečněná efektivita	Cena
Patriot	$> 99\%$	$< 1\%$	$1G\$$
Raketoplány	$1 - \frac{1}{10^6}$	1-$\frac{2}{350}$	$10G\$$
GMD	$> 90\%$	$< 10\%$	$100G\$(200G\$)$

Outline

Závěr

Co si odnést?

- Definice elementárního jevu, jevu, náhodné veličiny a stochastického procesu
- Příklady diskrétních a spojitých distribucí
- Pravěpodobnost a statistika jsou důležitými obory pro posuzování efektivity