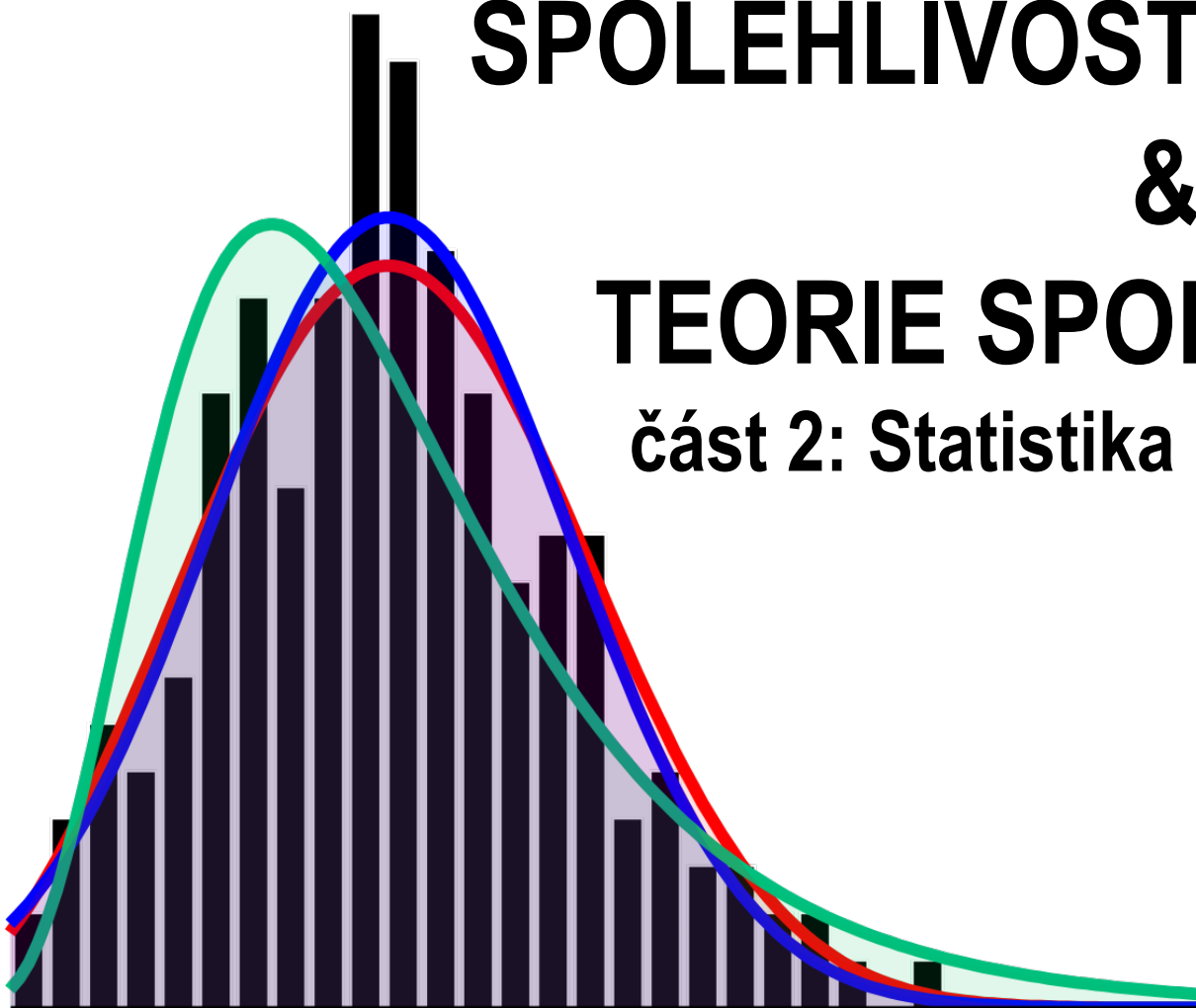




SPOLEHLIVOST KONSTRUKCÍ & TEORIE SPOLEHLIVOSTI

část 2: Statistika a pravděpodobnost

Drahomír Novák
Jan Eliáš
David Lehký





část 2 Statistika a pravděpodobnost



Náhodná veličina

Definice: Veličina X se nazývá **náhodná veličina** pokud její možné realizace x jsou výsledkem náhodného procesu.

Náhodné veličiny dělíme na:

- **Diskrétní** – nabývá pouze **izolovaných** hodnot (hod kostkou, počet osob).
- **Spojitě** – nabývá **libovolné** hodnoty z určitého interval (pevnost, teplota).

Základní statistický soubor (populace) – souhrn všech možných realizací x veličiny X (diskrétní vs. spojitě).

Statistický výběr – pouze některé realizace (např. výsledky měření).

Hustota p-sti, rozdělovací funkce (pdf)

– označení

$$f_X(x)$$

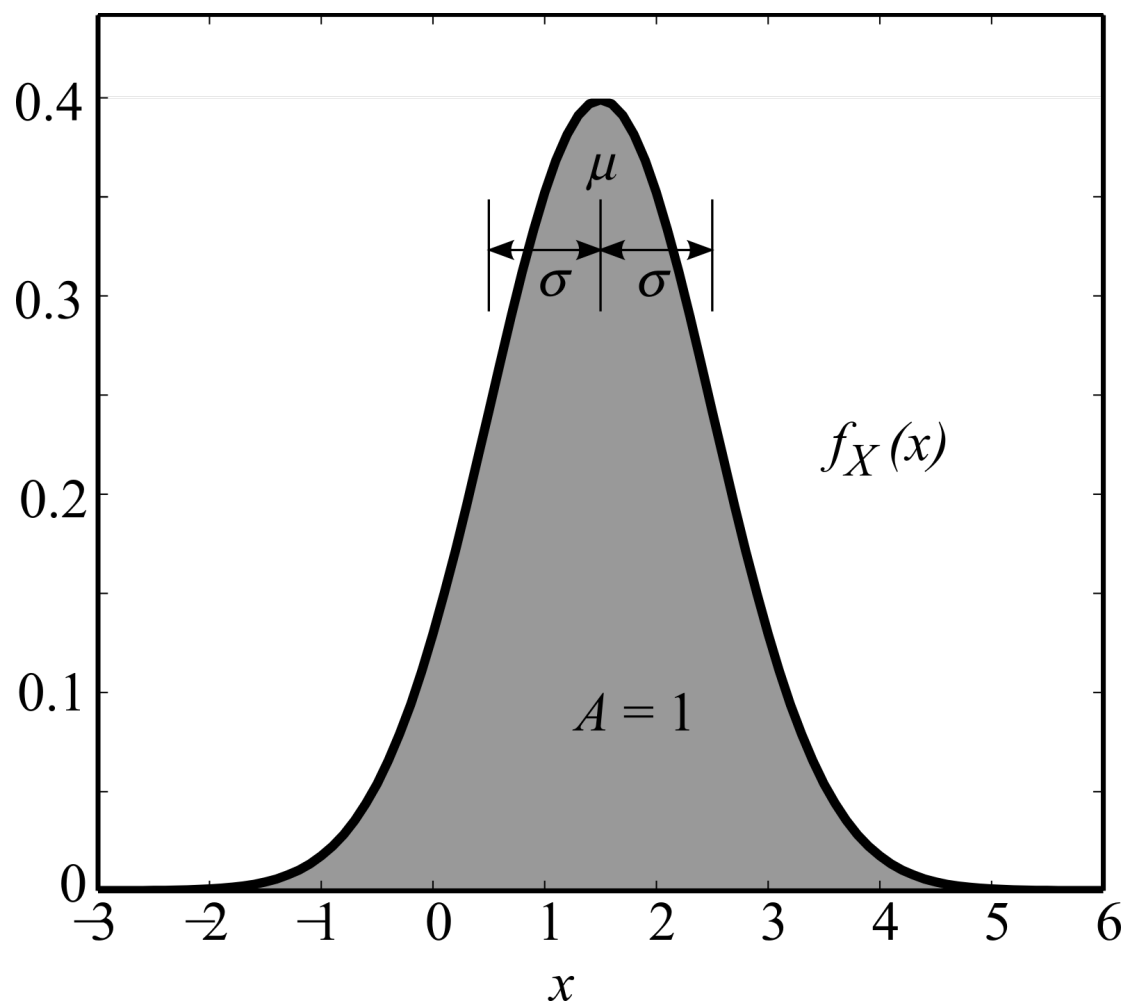
– vlastnosti

nezáporná

$$f_X(x) \geq 0$$

jednotková plocha

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$



Obrázek 1. pdf

Distribuční funkce, kumulativní fce (cdf)

–označení

$$F_X(x)$$

–Definition

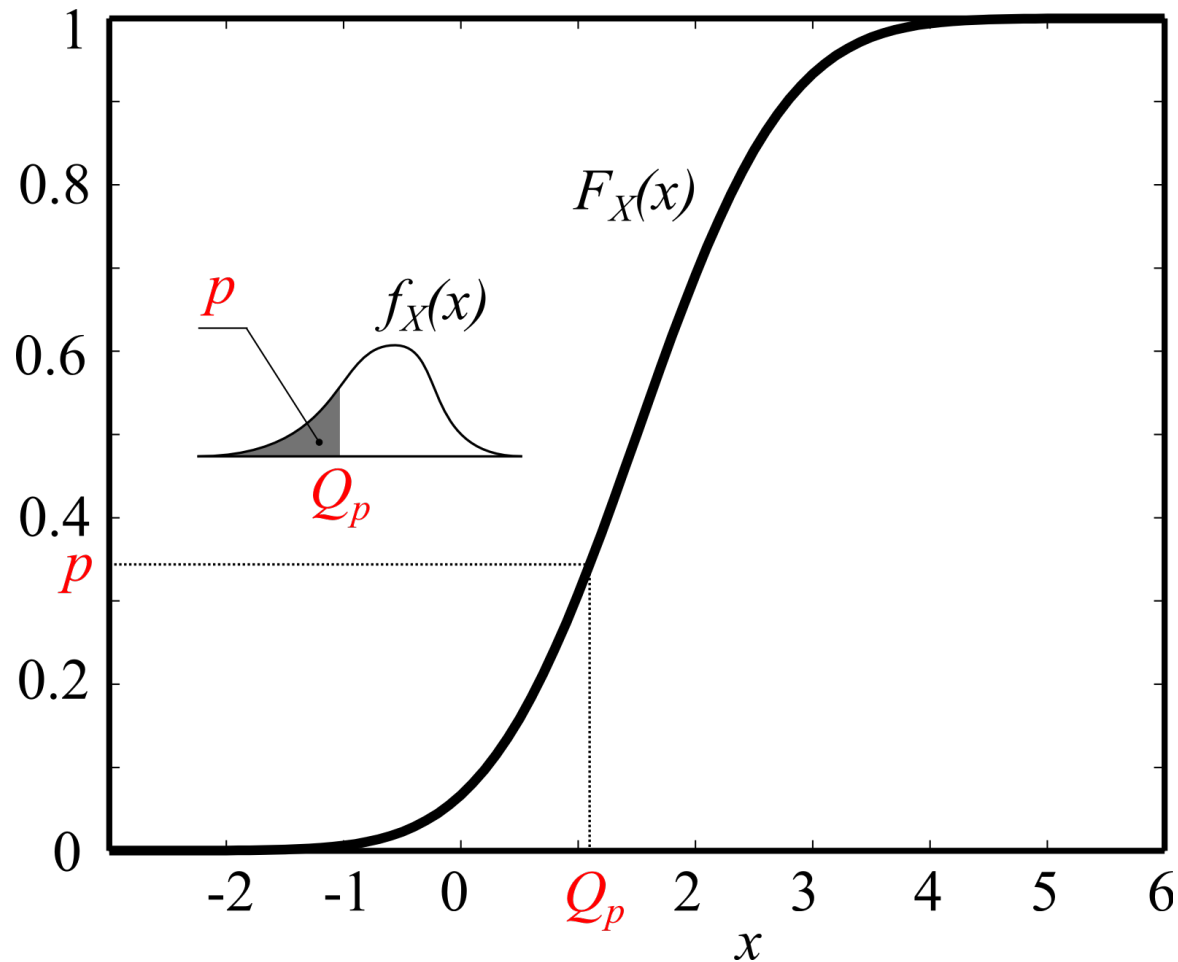
$$\begin{aligned} F_X(x) &= P(X \leq x) = \\ &= \int_{-\infty}^x f_X(t) dt \end{aligned}$$

–vlastnosti

$$\lim_{x \rightarrow \infty} F_X(x) = 1$$

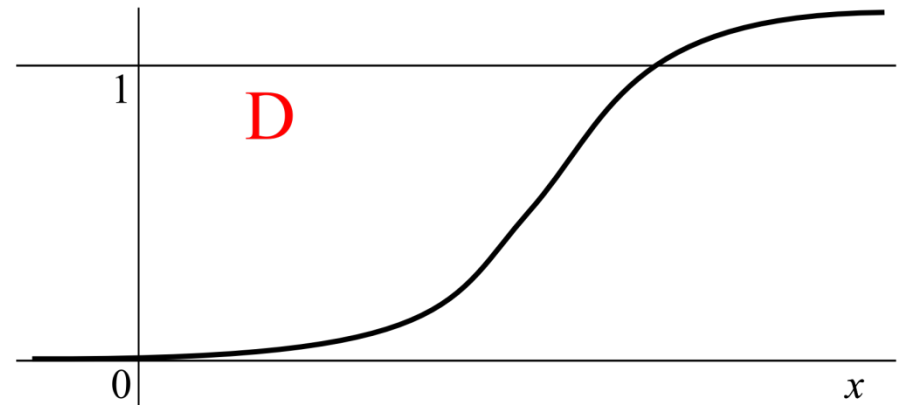
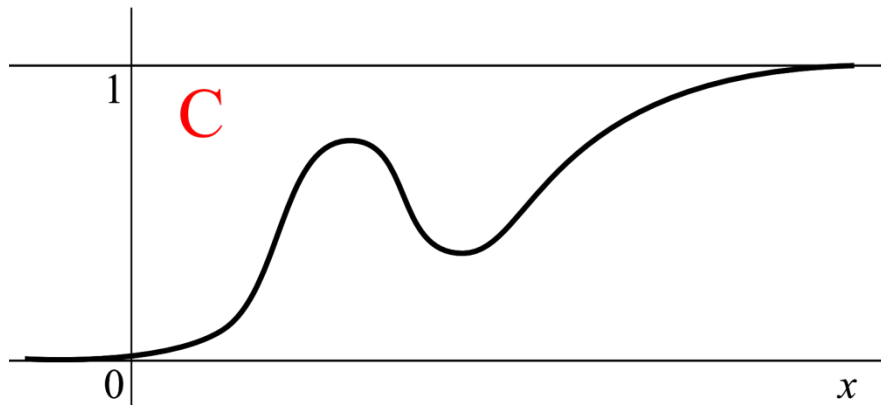
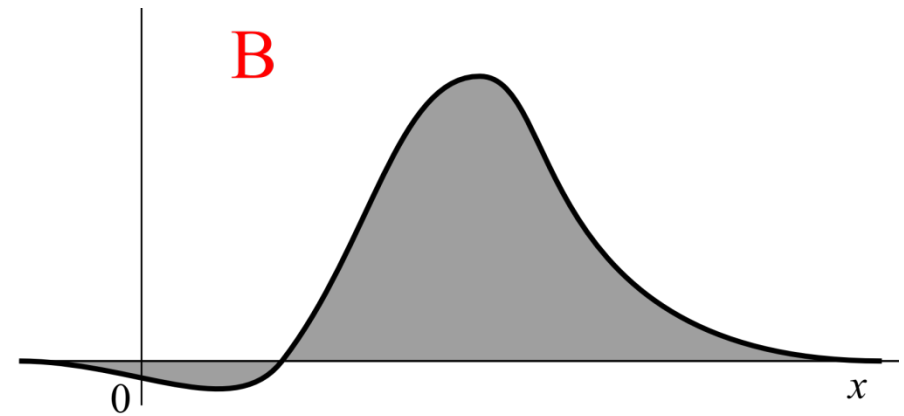
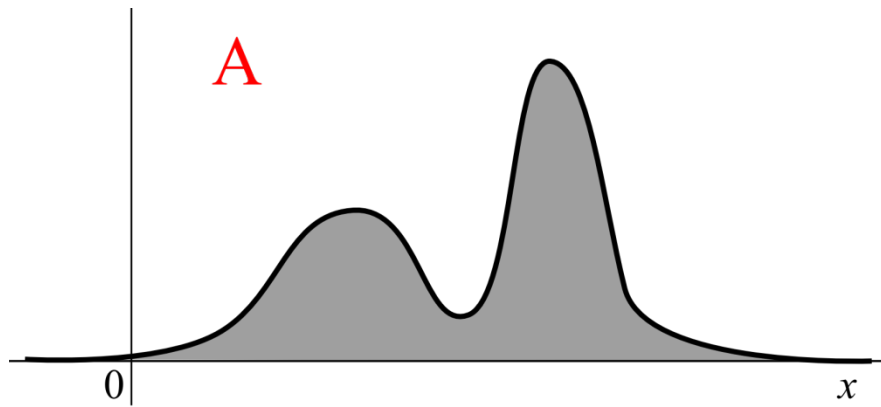
$$\lim_{x \rightarrow -\infty} F_X(x) = 0$$

$$\frac{\partial F_X(x)}{\partial x} \geq 0 \quad \text{neklesající}$$



Obrázek 3. cdf

Co je a co není platná cdf nebo pdf



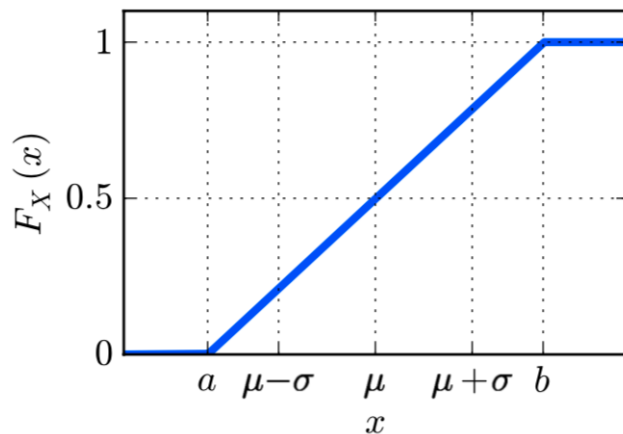
Obrázek 4. Testování kritérií cdf a pdf

Rovnoměrné (obdelníkové) rozdělení

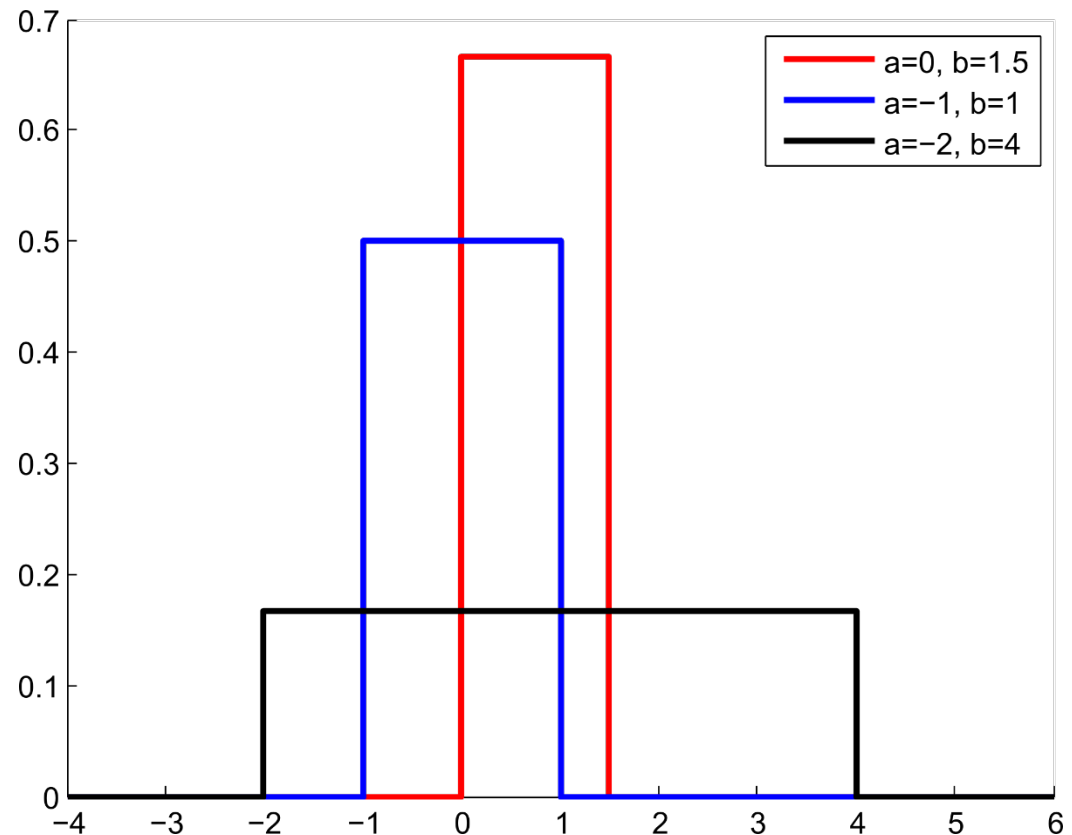
–dva parametry

- a – spodní mez
- b – horní mez

– lze ho jednoduše
simulovat na počítači



$$f(x) = \frac{1}{b-a} \quad x \in \langle a, b \rangle$$



Obrázek 9. Rovnoměrné rozdělení

Normální (Gaussovo) rozdělení

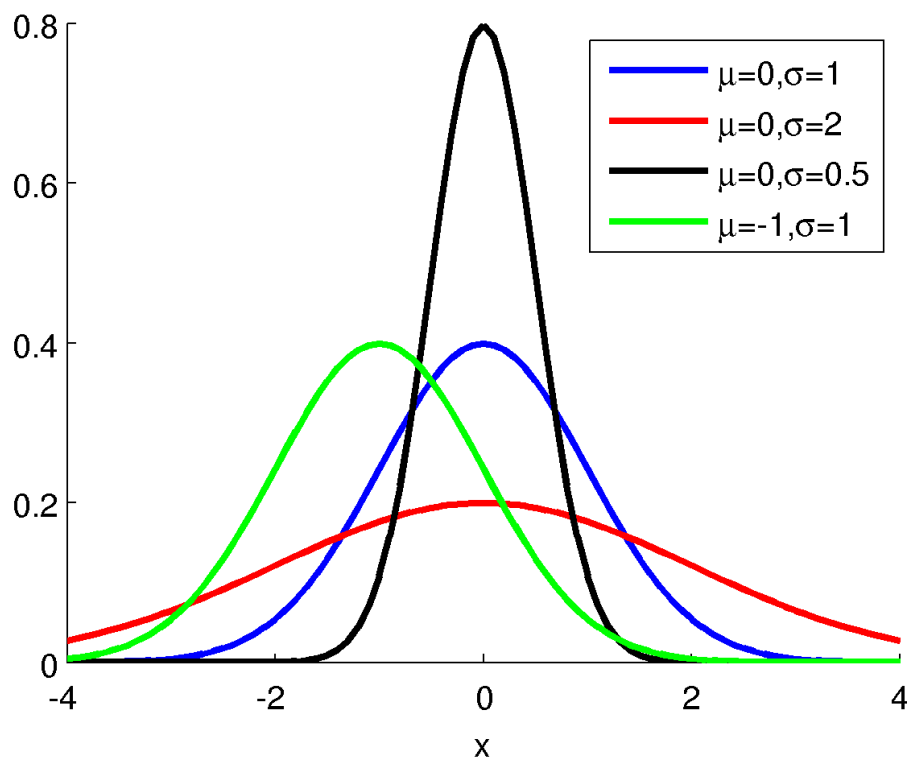
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad x \in (-\infty, \infty)$$

–dva parametry

- μ – střední hodnota
- σ – směrodatná odchylka

–výsledek centrálního limitního teorému (součet mnoha veličin vždy vede na Normální rozdělení)

–například geometrické veličiny, IQ, ...

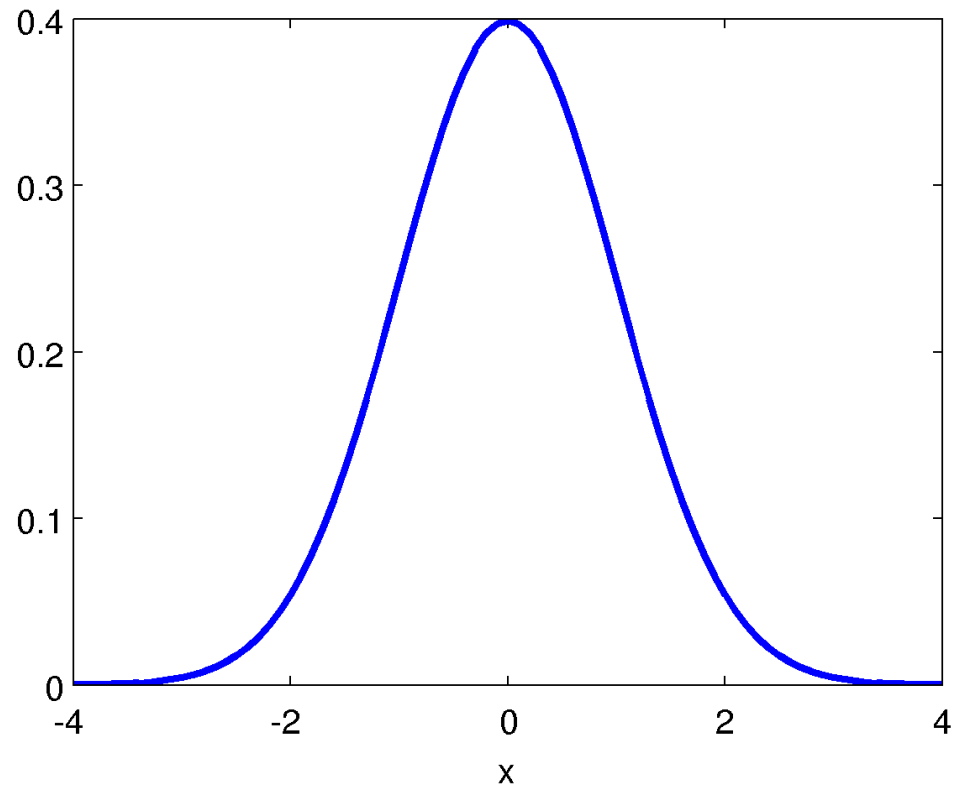


Obrázek 10. Normalní rozdělení

Standardizované normální rozdělení

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad x \in (-\infty, \infty)$$

- střední hodnota 0
- směrodatná odchylka 1
- symbol ϕ



Obrázek 11. Standardizované normální rozdělení

Lognormální rozdělení

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\log(x) - \mu)^2}{2\sigma^2}} \quad x \in \langle 0, \infty \rangle$$

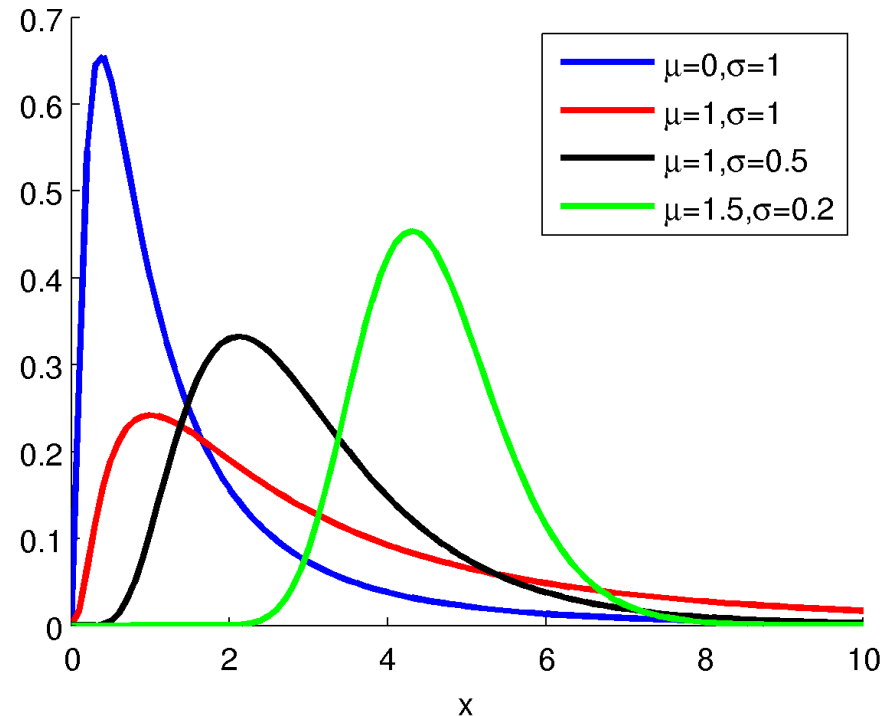
–dva parametry

- μ – střední hodnota v log-prostoru
- σ – směrodatná odchylka v log-prostoru

–omezeno zdola nulou

–např. pevnost, hustota, nejistoty, Youngův modul,

...



Obrázek 12. Lognormální rozdělení

Weibullovo rozdělení (2-parametrické)

$$f(x) = \frac{m}{\lambda} \left(\frac{x}{\lambda}\right)^{m-1} e^{-\left(\frac{x}{\lambda}\right)^m} \quad x \in \langle 0, \infty \rangle$$

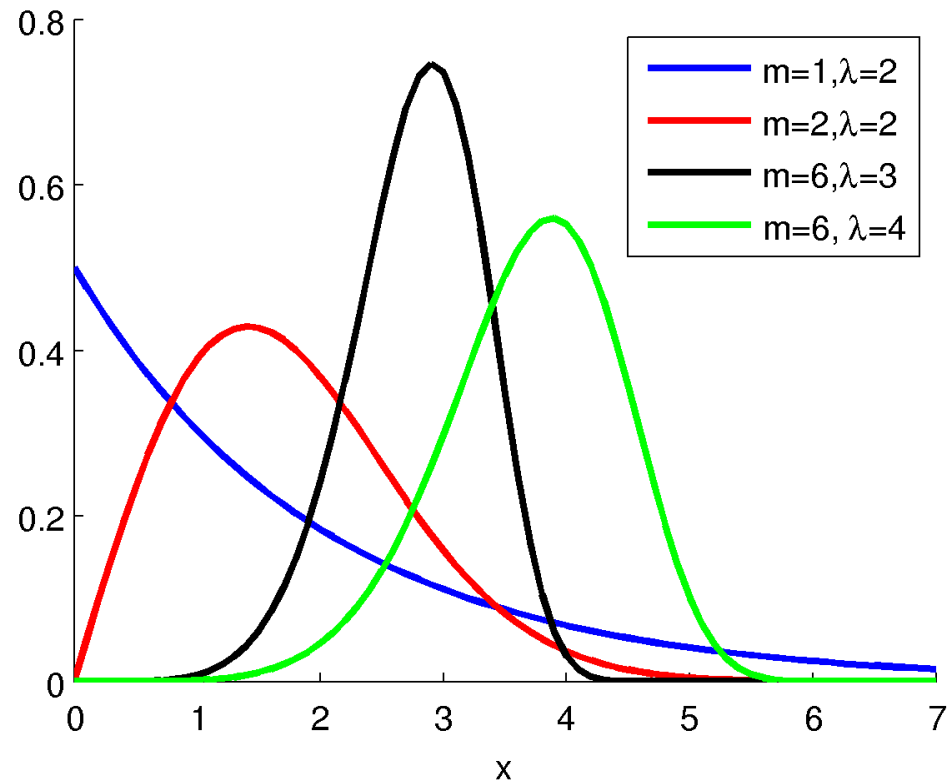
– dva parametry

- m – parametr tvaru
- λ – parametr měřítka

– výsledek teorie extrémních hodnot (minimum z mnoha veličin může vést na Weibullovo rozdělení)

– např. pevnost, ...

– omezeno zdola nulou



Obrázek 13. Weibullovo rozdělení

Beta rozdělení

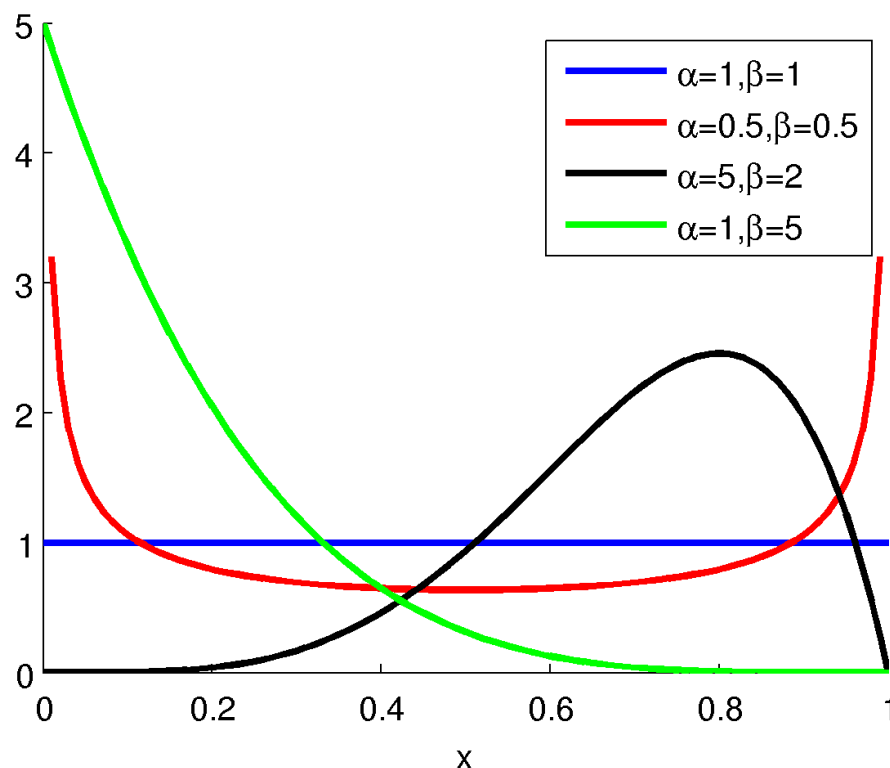
$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \quad x \in \langle 0, 1 \rangle$$

–dva parametry

- α, β – parametry tvaru

–oboustranně omezené

–meze mohou být
jednoduše upraveny
z intervalu $[0, 1]$ na
libovolný reálný interval
=> další dva parametry



Obrázek 14. Beta rozdělení

Charakteristiky rozdělení

Statistické momenty:

–střední hodnota (1. centrální moment = **těžiště** hustoty)

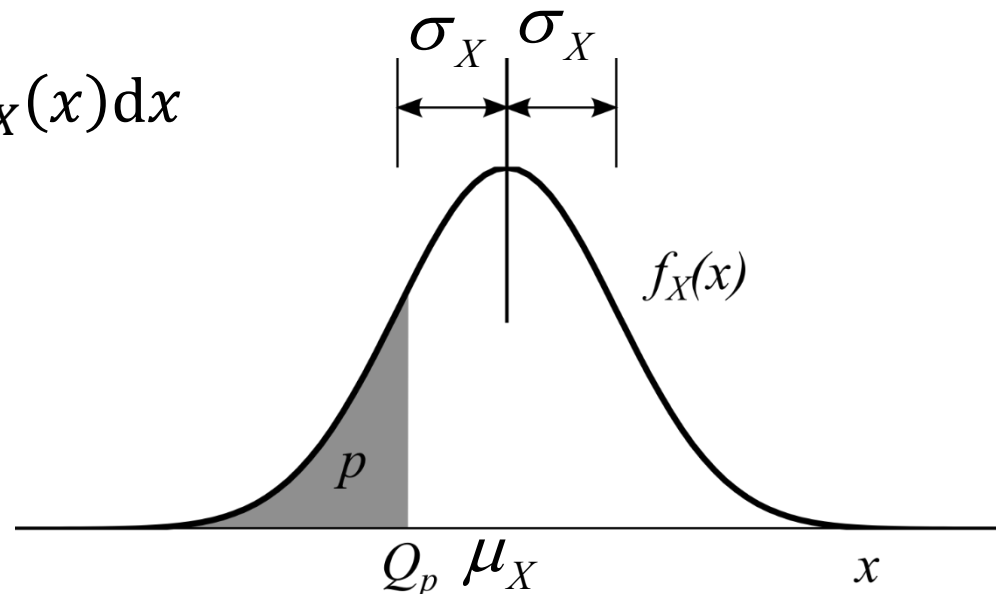
$$\mu_X = \int_{-\infty}^{\infty} x f_X(x) dx$$

–rozptyl (2. CM, absolutní míra rozptýlení kolem stř. hodnoty)

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx$$

–směrodatná odchylka

$$\sigma_X = \sqrt{\sigma_X^2}$$



Charakteristiky rozdělení

– variační koeficient (míra relativního rozptýlení)

$$CoV_X = V_X = \frac{\sigma_X}{\mu_X}$$

– šikmost (míra nesymetrie základního souboru)

$$\alpha_X = \frac{1}{\sigma_X^3} \int_{-\infty}^{\infty} (x - \mu_X)^3 f_X(x) dx$$

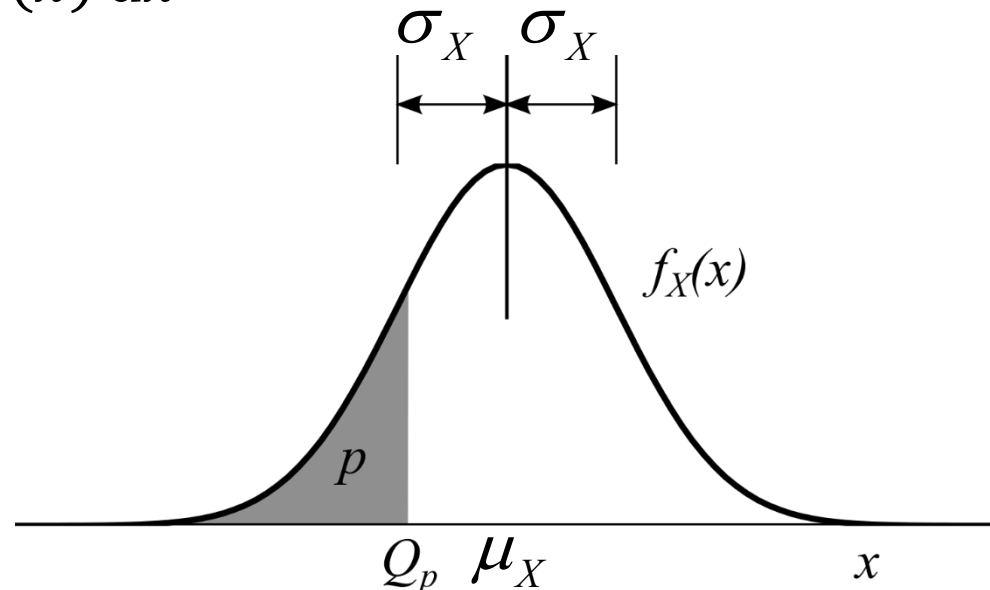
Kvantily

– 100p% kvantil

$$Q_p = F_X^{-1}(p)$$

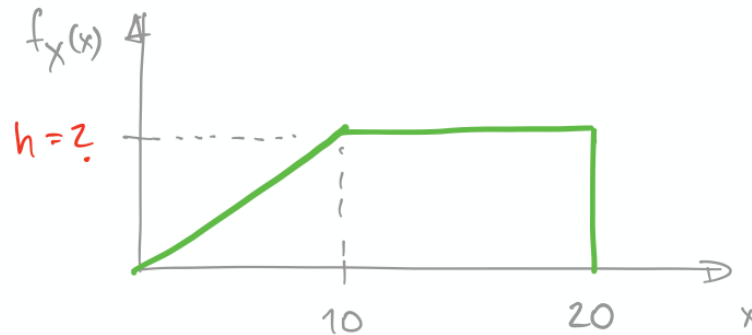
– medián

$$Q_{0.5} = F_X^{-1}(0.5)$$



Charakteristiky rozdělení – příklad

Př.: Spočítejte střední hodnotu a rozptyl náh. veličiny X .



$$A = \int_{-\infty}^{\infty} f_X(x) dx = 1$$

$$A = \frac{1}{2} 10 \cdot h + 10 \cdot h = 15h = 1$$

↓

$$\underline{h = \frac{1}{15}}$$

Hustota p-stí:

$$f_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{15}x & 0 \leq x \leq 10 \\ \frac{1}{15} & 10 \leq x \leq 20 \\ 0 & x > 20 \end{cases}$$

Charakteristiky rozdělení – příklad



Hustota p-stí :

$$f_X(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{150}x & 0 \leq x \leq 10 \\ \frac{1}{15} & 10 \leq x \leq 20 \\ 0 & x > 20 \end{cases}$$

Střední hodnota :

$$\begin{aligned} E(X) = \mu_X &= \int_{-\infty}^{\infty} x \cdot f_X(x) dx = \int_0^{10} \frac{1}{150} x^2 dx + \int_{10}^{20} \frac{1}{15} x dx = \frac{1}{150} \left[\frac{x^3}{3} \right]_0^{10} + \frac{1}{15} \left[\frac{x^2}{2} \right]_{10}^{20} = \\ &= \frac{1}{150} \cdot \frac{1000}{3} + \frac{1}{15} \left(\frac{400}{2} - \frac{100}{2} \right) = \frac{20}{9} + 10 = \frac{110}{9} = \underline{\underline{12,2}} \end{aligned}$$

Pomocí těžiště :

$$E(X) = \mu_X = x_t = \frac{S_{x10}}{A} = \left(\frac{1}{2} \cdot 10 \cdot \frac{1}{15} \cdot \frac{2}{3} \cdot 10 + 10 \cdot \frac{1}{15} \cdot 10 \right) / 1 = \frac{20}{9} + 10 = \underline{\underline{12,2}}$$

Charakteristiky rozdělení – příklad

Rozptyl:

$$D(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \quad \text{nebo} \quad \underline{D(X) = E(X^2) - [E(X)]^2}$$

$$\begin{aligned} E(X^2) &= \int_0^{10} \frac{1}{150} X^3 dx + \int_{10}^{20} \frac{1}{15} X^2 dx = \frac{1}{150} \left[\frac{X^4}{4} \right]_0^{10} + \frac{1}{15} \left[\frac{X^3}{3} \right]_{10}^{20} = \\ &= \frac{1}{150} \left(\frac{10000}{4} \right) + \frac{1}{15} \left(\frac{8000}{3} - \frac{1000}{3} \right) = \frac{50}{3} + 155\overline{15} = \underline{172\overline{2}} \end{aligned}$$

$$D(X) = E(X^2) - [E(X)]^2 = 172\overline{2} - 12\overline{12}^2 = \underline{\underline{22,8395}}$$

Nebo moment setrvačnosti $f_X(x)$:

$$\begin{aligned} D(X) = I_{4,1} &= \frac{1}{36} b_1 h_1^3 + A_1 \cdot x_1^2 + \frac{1}{12} b_2 h_2^3 + A_2 \cdot x_2^2 = \frac{1}{36} \cdot \frac{1}{15} \cdot 10^3 + \frac{1}{2} \cdot \frac{1}{15} \cdot 10 \cdot \left(12\overline{2} - \frac{2}{3} \cdot 10 \right)^2 + \\ &+ \frac{1}{12} \cdot \frac{1}{15} \cdot 10^3 + \frac{1}{15} \cdot 10 \cdot \left(15 - 12\overline{12} \right)^2 = 1,851851 + 10,28807 + 5,5 + 5,14403 = \underline{\underline{22,8395}} \end{aligned}$$

Jak určit charakteristiky a rozdělení?

- Přímé hodnocení **experimentálních dat** pomocí nástrojů matematické statistiky (laboratorní zkoušky, měření in situ).
- **Doporučení v normách** (Eurocode, JCSS probabilistic model code atd.).
- **Inženýrský odhad**.

Jak určit charakteristiky a rozdělení?

Experimenty:

Často máme k dispozici **omezený počet realizací** x veličiny X , tj. pracujeme se **statistickým výběrem**. Cílem je pak najít tzv. **nejlepší odhady** statistických momentů z daného výběru.

- **Aritmetický průměr** – průměr hodnot výběru, nejlepší nestranný odhad střední hodnoty

$$m_X = \frac{1}{n} \sum_{i=1}^n x_i$$

Jak určit charakteristiky a rozdělení?

- Výběrová rozptyl – nestranný odhad rozptylu

$$s_X^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m_X)^2$$

- Výběrová směrodatná odchylka

$$s_X = \sqrt{s_X^2}$$

- Výběrová šikmost – nestranný odhad šikmosti

$$a_X = \frac{1}{(n-1)(n-2)s_X^3} \sum_{i=1}^n (x_i - m_X)^3$$

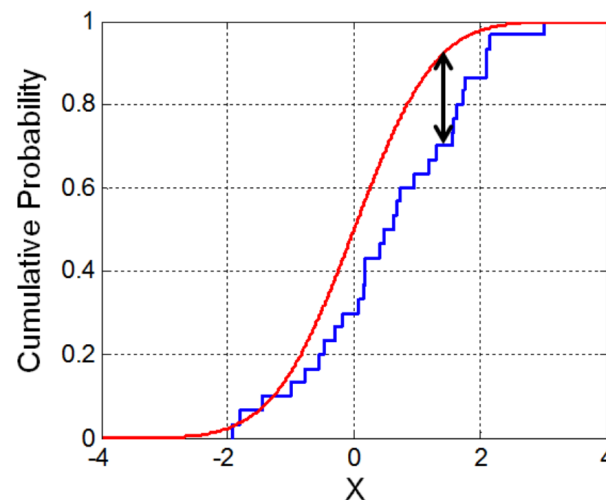


Jak určit charakteristiky a rozdělení?

Vhodné rozdělení se hledá pomocí tzv. **testů dobré shody** (angl. goodness of fit test) – jedná se o **testování statistické hypotézy o shodě** vybraného rozdělení a daného vzorku dat z populace.

Statistické testy:

- Kolmogorov-Smirnov
- Chi-kvadrát (Chi-square) test

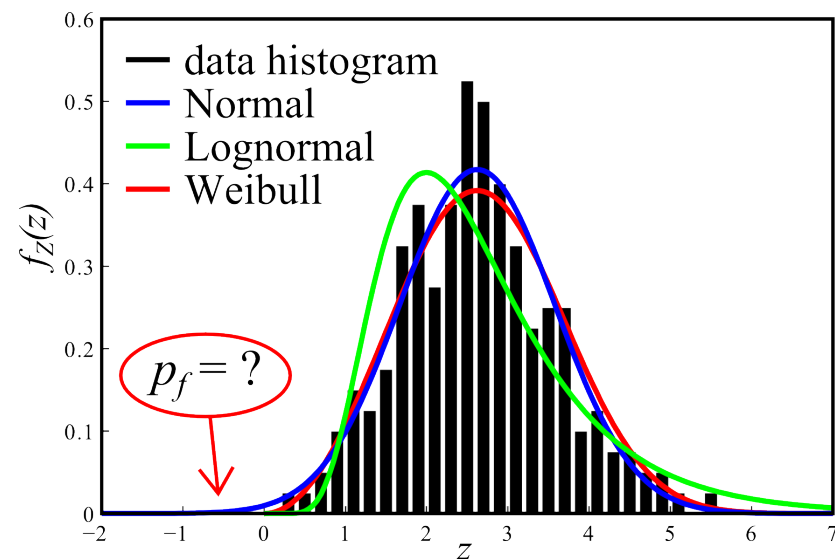
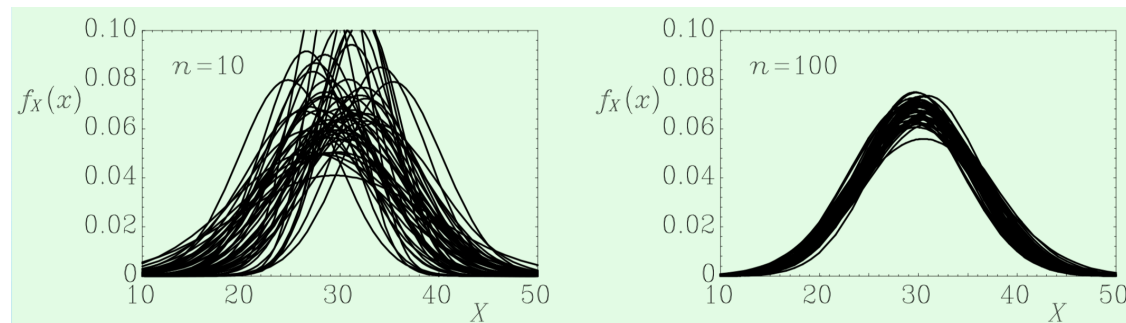


Kolmogorovův–Smirnovův test kvantifikuje vzdálenost mezi empirickou distribuční funkcí vzorků a kumulativní distribuční funkcí referenčního rozdělení:

$$D_n = \sup_x |F_n(x) - F_X(x)|$$

Charakteristiky a rozdělení – poznámky

- **Malý statistický výběr** (<30 vzorků) – omezená informace, reflektovat apriorní informaci o veličině)
- Střední část rozdělení může být určena z histogramů, ale pro **spolehlivostní výpočty** jsou nejdůležitější **okrajové části !!!**
- Nutnost určení rozdělení teoreticky (z fyzikálních principů)

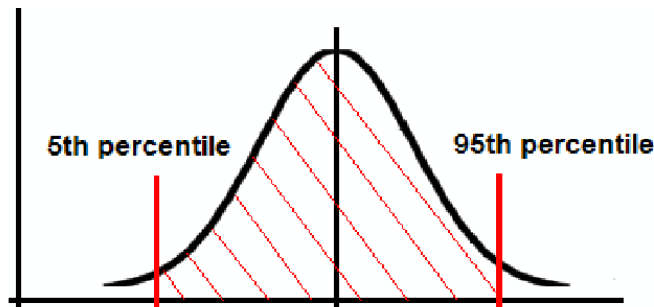


Jak určit charakteristiky a rozdělení?

Normové předpisy:

Pokud nejsou k dispozici experimentální data, lze získat statistické charakteristiky s pomocí normových předpisů a doporučení. Např. norma uvádí pevnostní třídy betonu a jim odpovídající charakteristické pevnosti (5% kvantily).

EN1992-1-1:



Strength classes for concrete														
f_{ck} (MPa)	12	16	20	25	30	35	40	45	50	55	60	70	80	90
$f_{ck,cube}$ (MPa)	15	20	25	30	37	45	50	55	60	67	75	85	95	105
f_{cm} (MPa)	20	24	28	33	38	43	48	53	58	63	68	78	88	98
f_{ctm} (MPa)	1,6	1,9	2,2	2,6	2,9	3,2	3,5	3,8	4,1	4,2	4,4	4,6	4,8	5,0
$f_{ctk, 0,05}$ (MPa)	1,1	1,3	1,5	1,8	2,0	2,2	2,5	2,7	2,9	3,0	3,1	3,2	3,4	3,5
$f_{ctk, 0,95}$ (MPa)	2,0	2,5	2,9	3,3	3,8	4,2	4,6	4,9	5,3	5,5	5,7	6,0	6,3	6,6
E_{cm} (GPa)	27	29	30	31	33	34	35	36	37	38	39	41	42	44
ε_{c1} (‰)	1,8	1,9	2,0	2,1	2,2	2,25	2,3	2,4	2,45	2,5	2,6	2,7	2,8	2,8



Jak určit charakteristiky a rozdělení?

Charakteristické hodnoty pevnostních veličin jsou typicky 5% kvantily. Pak **střední hodnotu** lze spočítat jako:

- *Normální rozdělení:*

$$X_m = \frac{X_k}{(1 - 1.645V_X)}$$

- *Lognormální rozdělení:*

$$X_m = \frac{X_k}{e^{-1.645V_X}}$$

$$X_m = \frac{X_k}{(1 - 2V_x)}$$

(simplified formula in JCSS)

$$X_m = X_k + 8 \text{ (MPa)}$$

(even more simplified formula in EN 1992-1-1)

Jak určit charakteristiky a rozdělení?

Pro **jiná rozdělení** lze použít např. software FReET (viz 4. přednáška):

Distribution details

Status: O.K.

Distribution: Lognormal (2 par)

Moments:

Mean: 61.304

Std: 7.3565

COV: 0.12

Skewness: 0.36173

Kurtosis: 0.23353

Parameters:

lambda: 4.1087

zeta: 0.11957

Moments & params:

Mean: 61.304

Std: 7.3565

Other:

Median: 60.867

Mode: 60.003

Random variable transformation:

☒ No transformation

☐ Minimum of 0 IID

☐ Maximum of 0 IID

Estimated moments:

Mean: 61.304

Std: 7.3565

COV: 0.12

Skewness: 0.36173

Kurtosis: 0.23353

Calculator:

x: 0

PDF(x): 0

CDF(x): 0

INV(p): 0.05

INV(p): 50

Transformed variable:

PDF(x): 0

CDF(x): 0

INV(p): 50

Apply



Jak určit charakteristiky a rozdělení?

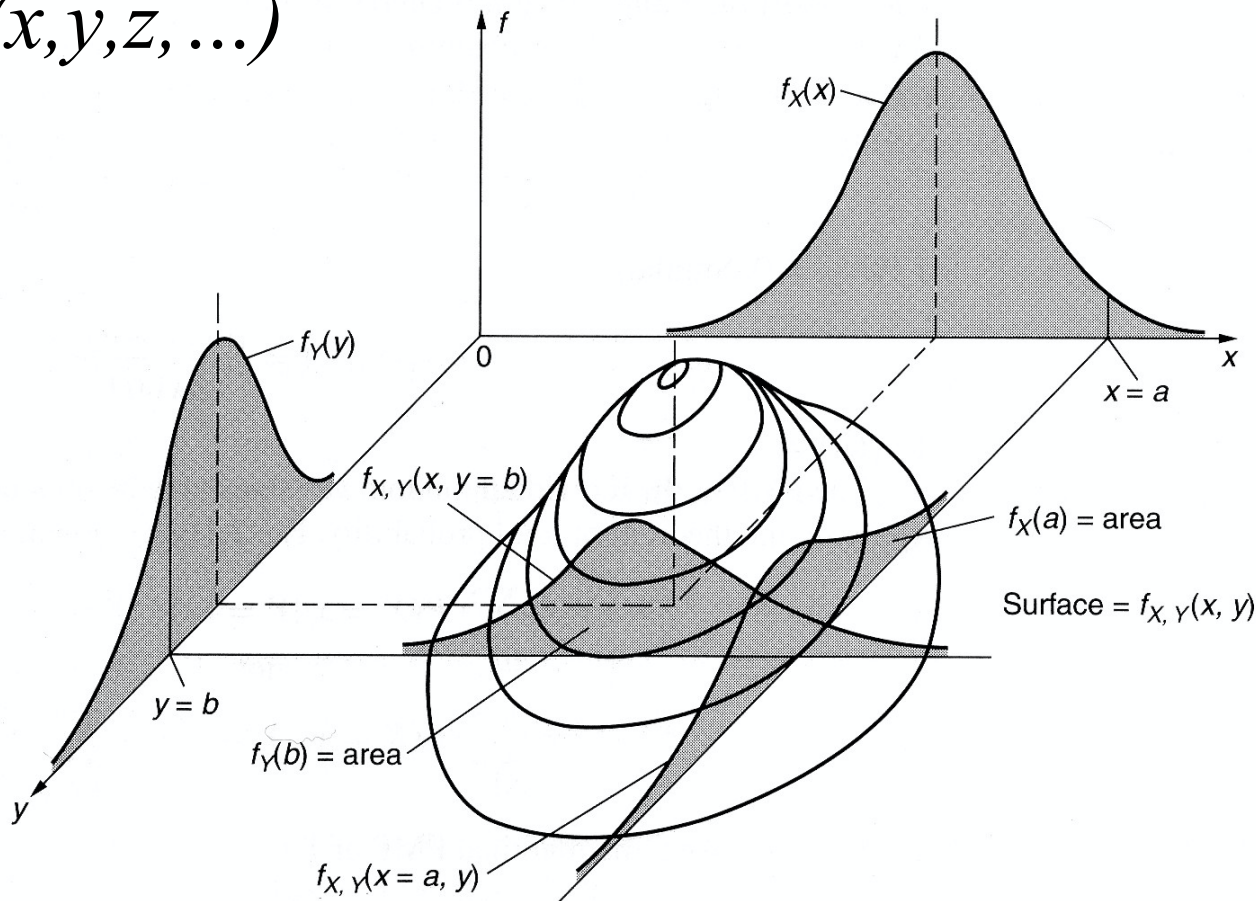
Vhodné **typy rozdělení** a **variační koeficienty** vybraných veličin lze nalézt např. v **JCSS probabilistic model code** apod.:

Material	Variable	Unit	PDF	Mean	CoV
Steel	Weight density	γ_s (kN/m ³)	Normal	77	< 0.01
Steel	Yield stress	f_y (MPa)	Lognormal (2 par)	$\frac{f_{yk}}{1 - 2CoV}$	0.05–0.08
Concrete	Weight density	γ_c (kN/m ³)	Normal	23–25	0.03–0.04
Concrete	Compressive strength	f_c [MPa]	Lognormal (2 par)	$\frac{f_{ck}}{1 - 2CoV}$	0.06–0.15
Concrete	Tensile strength	f_t (MPa)	Lognormal (2 par)	f_{tm}	0.3
Concrete	Modulus of elasticity	E (MPa)	Lognormal (2 par)	E_m	0.15
Concrete	Cover to reinforcement	c (m)	Beta	c_{nom}	0.15–0.20

Sdružená hustota pravděpodobnosti

–Rozdělení pravděpodobnost dvou či více veličin

$$f_{X,Y,Z,\dots}(x,y,z,\dots)$$

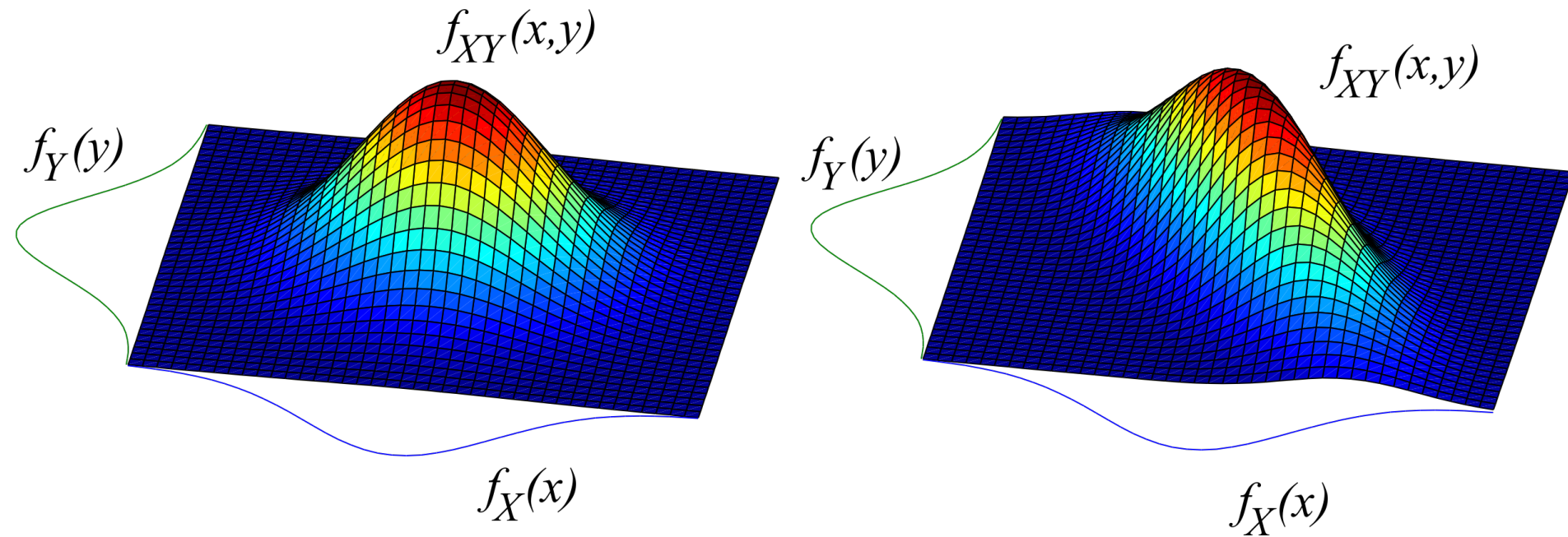


Obrázek 5. Sdružená hustota pravděpodobnosti

Statistická závislost

–veličiny X a Y jsou nezávislé právě když

$$f_{XY}(x, y) = f_X(x)f_Y(y)$$



Obrázek 6. Nezávislé a závislé veličiny

Pearsonův korelační koeficient

– míra **lineární** závislosti

– $\rho = 0$ **mohou** či **nemusí** být závislé

– $\rho \neq 0$ závislé

– $\rho > 0$ pozitivní korelace, $\rho < 0$ negativní korelace

– limity $\rho \in \langle -1, 1 \rangle$

$$\rho_{i,j} = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{var}(X_i) \cdot \text{var}(X_j)}}, \quad \langle -1, 1 \rangle$$

$$\rho(X, Y) = \frac{\int_{\Omega_X} \int_{\Omega_Y} xy f_{XY}(x, y) dx dy - \mu_X \mu_Y}{\sigma_X \sigma_Y}$$

Korelační matice

Diagonální

Symetrická

Pozitivně definitní

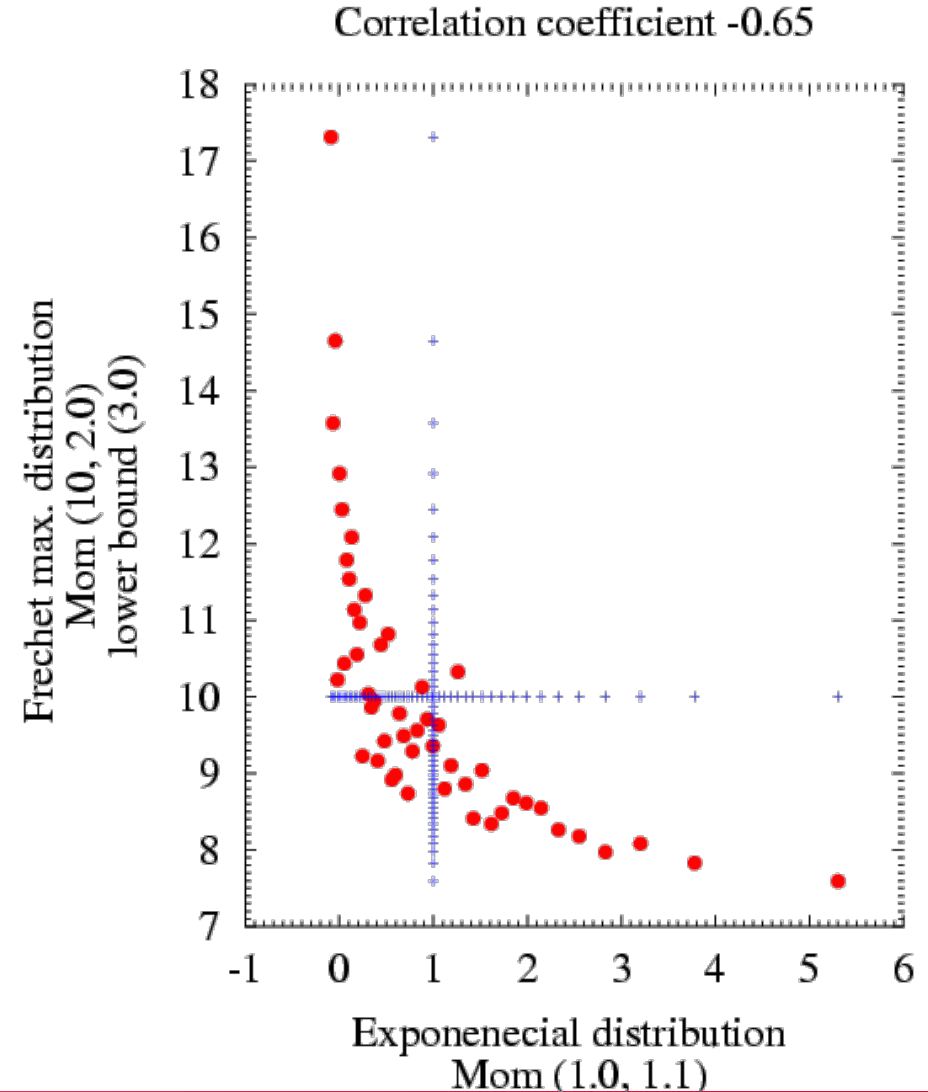
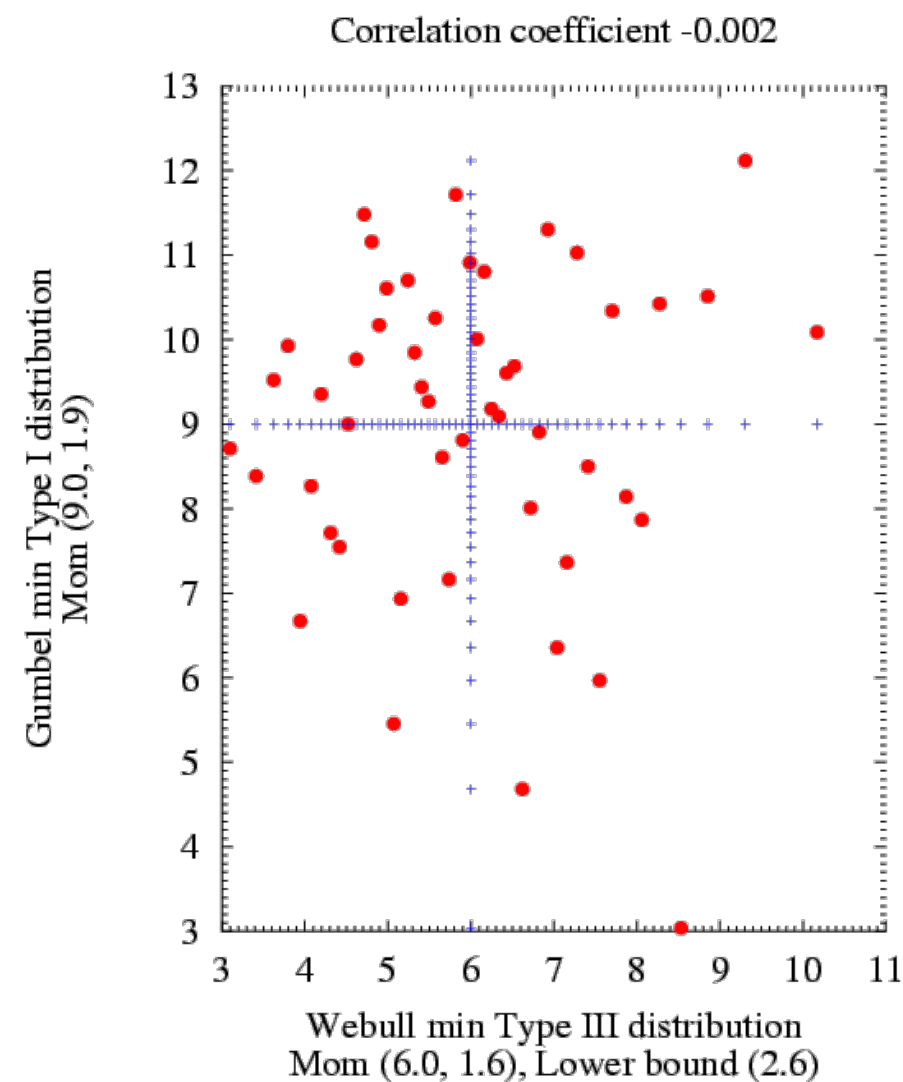
náhodná veličina

1	$A_{1,2}$	...		$A_{1,NVar}$
	1	...		$A_{2,NVar}$
		1		...
			1	...
				1

Symetrie
korelační matice

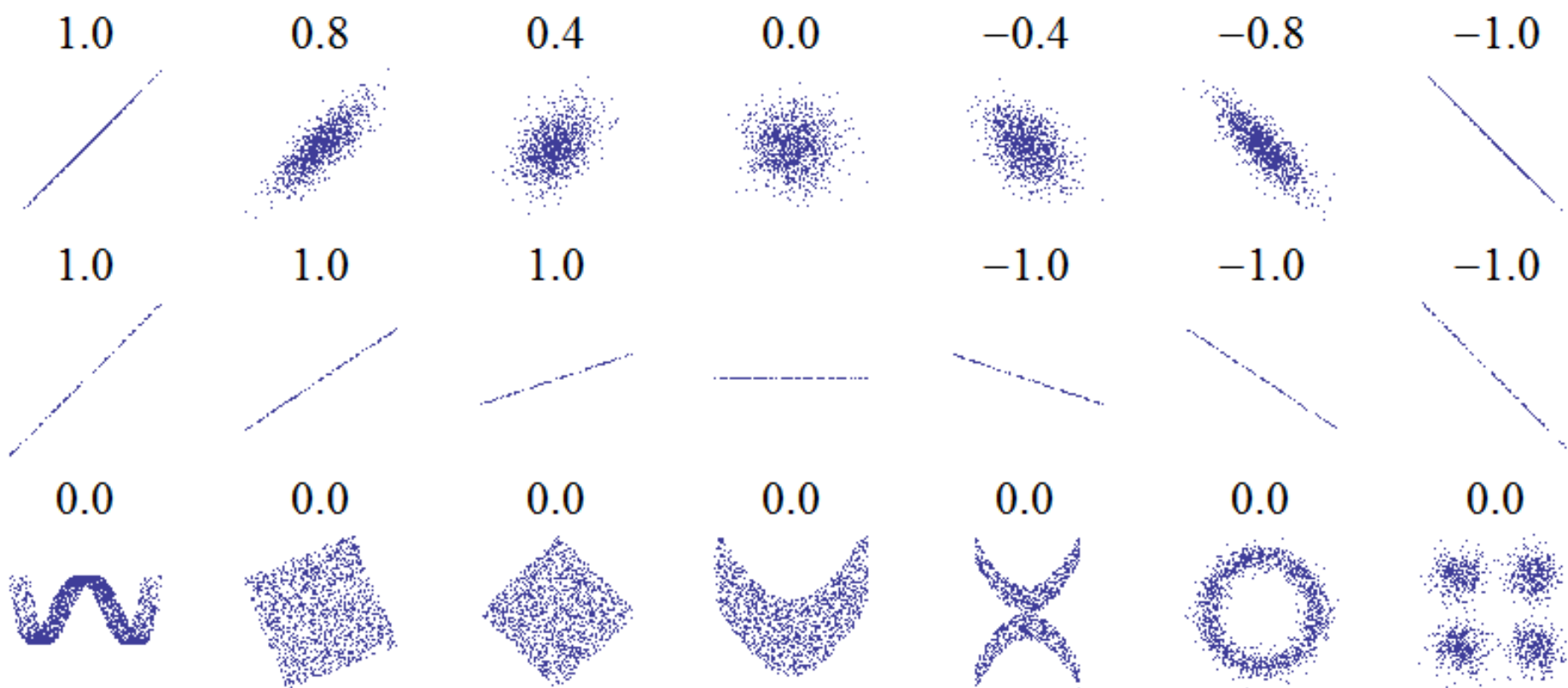
náhodná veličina

2D reprezentace závislosti



Příklady ne/závislých veličin

–poslední řádek = **nulová** Pearsonova korelace, jinak ale **závislé** náhodné veličiny



Obrázek 7. Příklady Personovy korelační míry

Neparametrické pořadové míry korelace

Spearman, nezávislá na typu rozdělení, robustní, **místo** samotných **hodnot** používá jejich **pořadí**

$$A_{i,j} = 1 - \frac{6 \sum_k (R_{ki} - R_{kj})^2}{N_{Sim}^3 - N_{Sim}}, \quad \langle -1, 1 \rangle$$

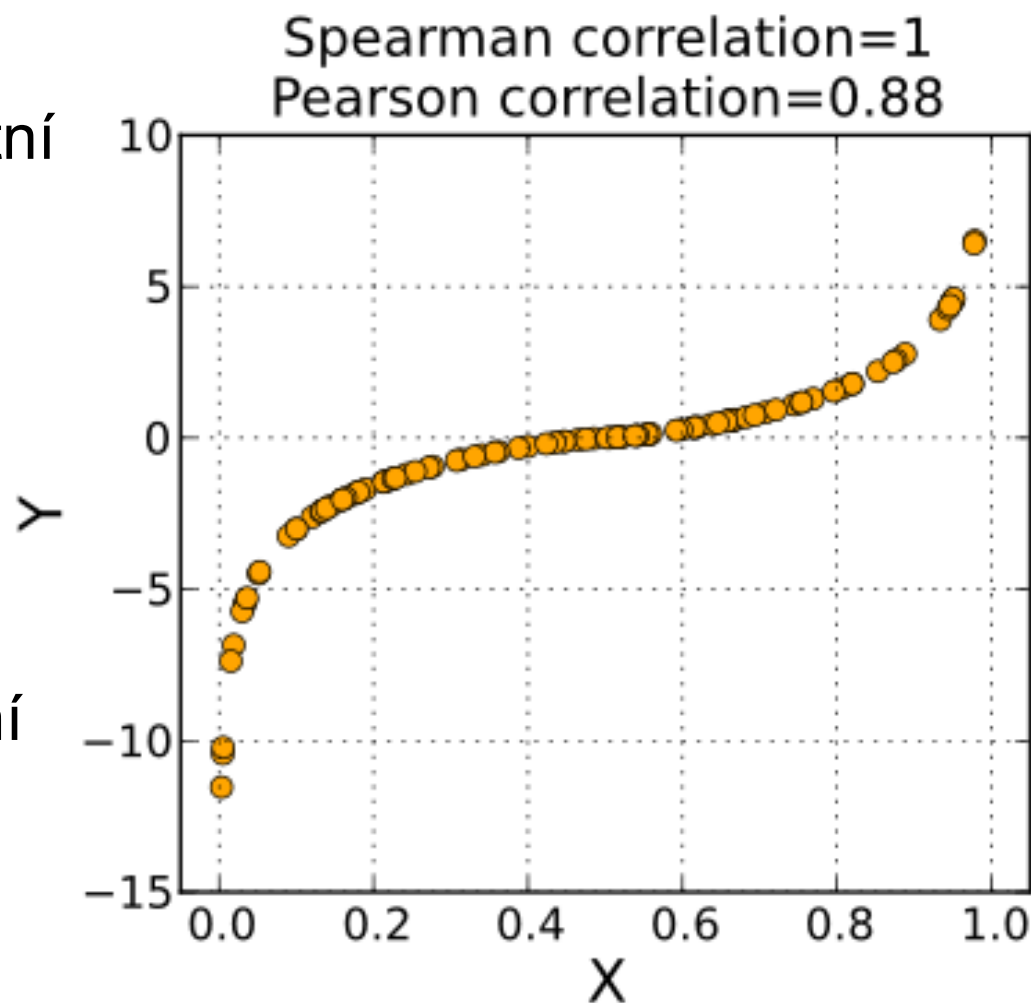
Kendallovo tau

$$\tau_i = \tau(q_{ji}, p_j), \quad j = 1, 2, \dots, N$$

Odolné k defektům v datech, používají jen pořadí

Srovnání korelačních měř

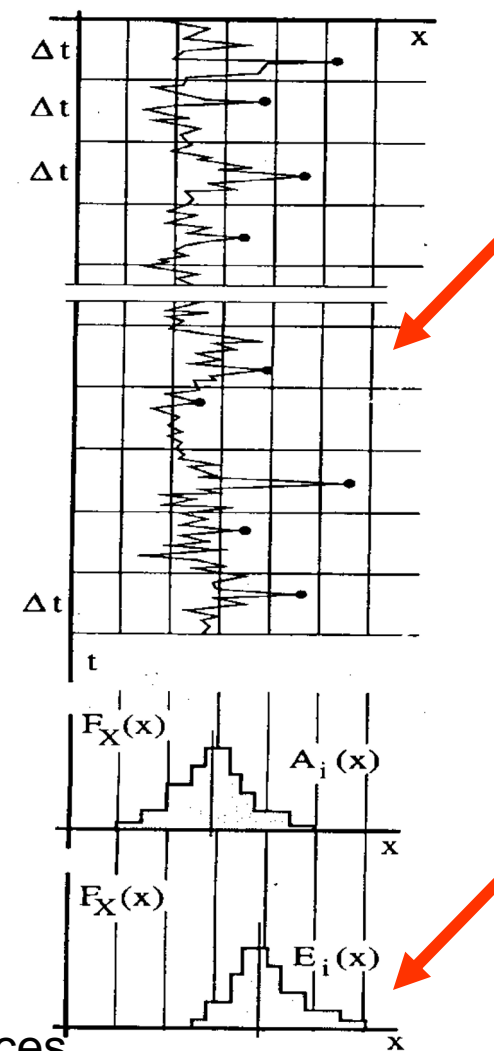
- pro praktické spolehlivostní výpočty je rozdíl nepodstatný, protože je beztak často minimální znalost skutečné závislosti vstupních veličin
- Pearson – klasická
- Neparametrická - robustní



Obrázek 8. Spearmanova versus Pearsonova korelace

Stochastický proces v čase

- ne pouze jednotlivé hodnoty ale také jejich **časová posloupnost**
- např. meteorol. data, výška hladiny, rychlost větru
- stacionární proces** – charakteristiky procesu (střední hodnota, rozptyl, vyšší statistické momenty + korelace) se s časem nemění
- A – klasický histogram
- E – histogram **extrémních hodnot**
- auto-korelační funkce, spektrální hustota
- stochastická metoda konečných prvků
- nestacionární proces**

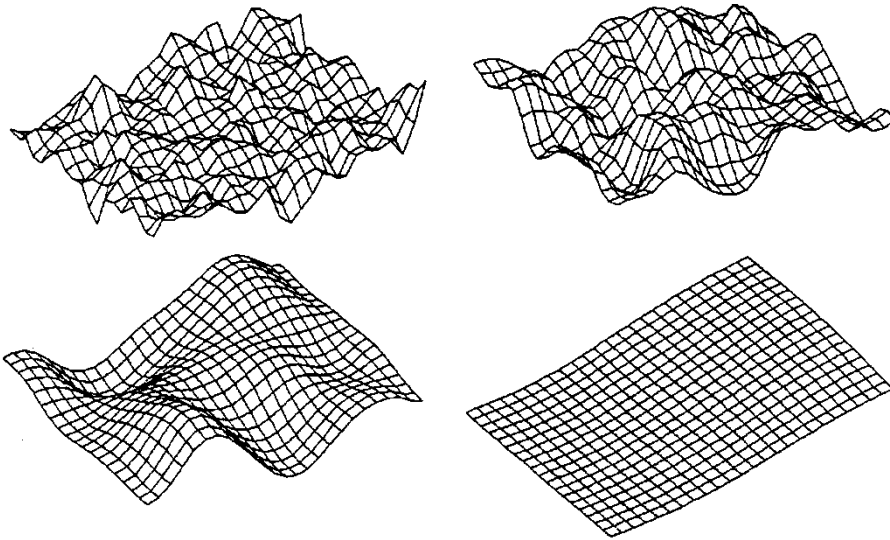


Obrázek 16. Stochastický proces

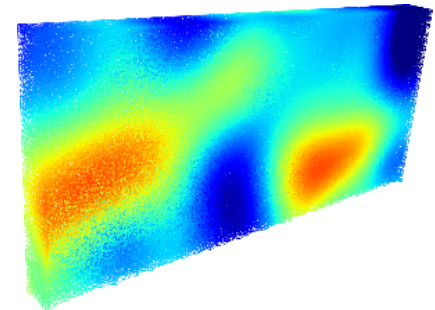
Náhodné pole – prostorová závislost

- Pevnost materiálu, Youngův modul, ... veličiny se mění v prostoru spíše **spojitě**
- spojitost lze zajistit požadováním kladné korelační závislosti jejíž velikost je dána vzdáleností

$$\rho_{ij} = \exp \left[- \left(\frac{\| \mathbf{x}_i - \mathbf{x}_j \|}{d} \right)^2 \right]$$



Obrázek 17. Náhodné pole ve 2D



Obrázek 18. Náhodné pole ve 3D