PŘEHLED VZORCŮ A HLAVNÍCH POJMŮ
na přednáškách KMA/PSA
7. až 12. týdne ZS 2015/16

(následující bude mj. promítnuto během těchto přednášek PSA - volná místa na stránkách jsou úmyslně : po vytištění lze používat k doplnění vlastních poznámek)

Vysvětlení, použití, grafy, příklady, etc. budou na přednášce.

ÚVOD DO MATEMATICKÉ STATISTIKY

1. POPISNÁ STATISTIKA

(statistický) **soubor**: $x_1, x_2, \ldots, x_n$

[$\approx$ hodnoty (diskr. nebo spoj.) náh. veličiny]

$x_i \in R \ldots$ prvek souboru, hodnota v souboru

$n \in N \ldots$ rozsah souboru

četnost, resp. relativní četnost prvku

uspořádaný soubor: $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$

CHARAKTERISTIKY POLOHY

• **(aritmetický) průměr**:

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

Je-li n_i četnost i-té hodnoty x_i, $(i = 1, 2, \ldots, k)$ souboru, pak $n = n_1 + n_2 + \ldots + n_k$, tj. průměr lze napsat jako $\overline{x} = \frac{x_1 n_1 + x_2 n_2 + \ldots + x_k n_k}{n} = \sum_{i=1}^{k} x_i w_i\,,$ kde $w_i = \frac{n_i}{n}$ $(i = 1, 2, \ldots, k)$.

• **medián** $\tilde{x}$ (uspořádaného) souboru je jeho:
prostřední hodnota (je-li n liché), resp.
aritm. průměr dvou prostředních hodnot (je-li n sudé)

• **modus** $\hat{x}$ je hodnota(-y) s nejvyšší četností

CHARAKTERISTIKY VARIABILITY

• **rozpětí** je rozdíl mezi největší a nejmenší hodnotou

• **(výběrový) rozptyl**:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

• **(výběrová) směrodatná odchylka**:

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

2. ODHADY PARAMETRŮ

náhodný výběr rozsahu n z rozdělení náhodné veličiny X je posloupnost *nezávislých* náh. veličin $X_1, X_2, \ldots, X_n$, které mají stejné rozdělení jako X

$x_1, x_2, \ldots, x_n$ je tzv. **realizace** náh. výběru (= statistický soubor)

neznámou hodnotu parametru v rozdělení X nahrazujeme vhodným reálným číslem, spočteným z realizace náh. výběru - tzv. **bodový odhad** parametru

• stř. hodnotu $E(X)$ odhadujeme aritm. průměrem:

$$E(X) \approx \bar{x}$$

odtud: $p \approx \bar{x}$ pro $X \sim A(p)$

$\lambda \approx \bar{x}$ pro $X \sim Po\,(\lambda)$

$\delta \approx \bar{x}$ pro $X \sim Exp\,(\delta)$

$\mu \approx \bar{x}$ pro $X \sim N(\mu, \sigma^2)$

• rozptyl $D(X)$ odhadujeme výběrovým rozptylem:

$$D(X) \approx s^2$$

odtud: $\sigma^2 \approx s^2$ pro $X \sim N(\mu, \sigma^2)$

intervalové odhady (neznámého) parametru θ rozdělení náh. veličiny X:

pro $\alpha \in (0, 1)$ [často bývá $\alpha = 0.01,\ 0.05$ nebo 0.10]

interval (a, b) nazveme **100 (1−α) %-ní** (oboustranný) **interval spolehlivosti**, je-li

$$\boxed{P(a < \theta < b) \,=\, 1 - \alpha}$$

[a, resp. b je tzv. dolní, resp. horní mez spolehlivosti]

jednostranné intervaly spolehlivosti: $P(\theta < b) = 1 - \alpha$, resp. $P(a < \theta) = 1 - \alpha$

- je-li $x_1, x_2, \ldots, x_n$ realizace náhodného výběru z rozdělení $X \sim A(p)$, pak $100\,(1-\alpha)$%-ní **interval spolehlivosti pro parametr** p je:

$$\left(\hat{p} - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \; , \; \hat{p} + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

  kde $\hat{p} \approx \bar{x}$, přitom n je tak velké, aby $\boxed{n\hat{p}(1-\hat{p}) \geq 9}$,
  $u_{1-\frac{\alpha}{2}}$ jsou kvantily rozdělení $N(0,1)$

- je-li $x_1, x_2, \ldots, x_n$ realizace náhodného výběru z rozdělení $X \sim N(\mu, \sigma^2)$, kde $\boxed{n > 30}$,
  pak $100\,(1-\alpha)$%-ní **interval spolehlivosti**
  pro parametr μ je:

$$\left(\bar{x} - u_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \, , \; \bar{x} + u_{1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right)$$

  kde $u_{1-\frac{\alpha}{2}}$ jsou kvantily rozdělení $N(0,1)$

Pozn.: pro $\boxed{n \leq 30}$ tento interval spolehlivosti je:

$$\left(\bar{x} - t_{\nu,\,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \, , \; \bar{x} + t_{\nu,\,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{n}} \right)$$

$t_{\nu,\,1-\frac{\alpha}{2}} \ldots$ kvantily t-rozdělení s ν stupni volnosti, kde $\nu = n - 1$

t-rozdělení, též **Studentovo rozdělení**, $t(\nu)$, je typem spojitého rozdělení ppsti. Křivka, která je grafem hustoty t-rozdělení, má "obdobný"tvar a symetrii jako normální rozdělení (kde $\mu = 0$), ale odráží větší "tvárnost"prostřednictvím *stupně volnosti*, označovaném ν.

Tabulky obsahují jen kvantily t-rozdělení pro běžné volby α.

Pro $n > 30$ se používá aproximace $t(\nu) \approx N(0, 1)$.

- je-li $x_1, x_2, \ldots, x_n$ realizace náhodného výběru z rozdělení $X \sim N(\mu, \sigma^2)$,
  pak $100\,(1-\alpha)$%-ní **interval spolehlivosti pro parametr** σ^2 je:

$$\left(\frac{(n-1)\,s^2}{\chi^2_{\nu,\,1-\frac{\alpha}{2}}}\,,\ \frac{(n-1)\,s^2}{\chi^2_{\nu,\,\frac{\alpha}{2}}} \right)$$

  kde $\chi^2_{\nu,\,1-\frac{\alpha}{2}}$, $\chi^2_{\nu,\,\frac{\alpha}{2}}$ jsou kvantily χ^2-rozdělení

  s ν stupni volnosti, kde $\nu = n-1$

chí-kvadrát rozdělení, $\chi^2(\nu)$, je typem spojitého rozdělení. Křivka, která je grafem hustoty χ^2-rozdělení, je identicky nulová pro všechny nekladné argumenty a nemá symetrický tvar. Tvar hustoty χ^2-rozdělení závisí na čísle zvaném *stupeň volnosti*, označovaném ν.

Tabulky obsahují jen kvantily χ^2-rozdělení pro běžné volby α.

3. TESTOVÁNÍ HYPOTÉZ

Použitím naměřených hodnot testujme určitou (statistickou) hypotézu, označme ji H_0.

Náš závěr bude vždy jeden z následujících dvou:
1. Zamítáme H_0.
2. Nezamítáme H_0.

Můžeme dopustit chyb dvou druhů:

chyba 1. druhu: H_0 je pravdivá, ale test vede k zamítnutí H_0;

chyba 2. druhu: H_0 je nepravdivá, ale test vede k nezamítnutí H_0.

Ppst chyby 1. druhu se značí α, požaduje se malá a volí se před testem (běžně: 0.1, 0.05, 0.01)

$\alpha \ldots$ tzv. **hladina významnosti**

Ppst chyby 2. druhu se označuje β (závisí na volbě α).

Rozhodnutí o případném zamítnutí H_0 provádíme pomocí testu zvaného **testové kritérium**.

Množina všech hodnot test. kritéria vedoucí k zamítnutí H_0, se nazývá **kritický obor**, označuje se W. Hodnota, která odděluje W od hodnot, které vedou k nezamítnutí H_0, je tzv. **kritická hodnota**

Obecný postup:

1) Stanovíme H_0 - tzv. **nulová hypotéza**

2) Stanovíme H_1 - tzv. **alternativní hypotéza**

3) Zvolíme **hladinu významnosti** α.

4) Vybereme vhodné test. kritérium, stanovíme jeho rozdělení ppsti za předpokladu platnosti H_0. Toto rozdělení, hladina význ. a formulace H_1 určují **kritický obor** W (kritickými hodnotami jsou pak příslušné kvantily rozdělení test.kritéria).
Načrtneme graf rozdělení test. kritéria, vyznačíme W.

5) Hodnota test. kritéria $\in W \Rightarrow$ **zamítáme** H_0.
Hodnota test. kritéria $\notin W \Rightarrow$ **nezamítáme** H_0.

Podle formulace H_1 dostáváme testy:

- **levostranné:**

  např. $H_1 : \mu < \mu_0$

- **pravostranné:**

  např. $H_1 : \mu > \mu_0$

- **oboustranné:**

  např. $H_1 : \mu \neq \mu_0$

Předpokládejme, že známe typ rozdělení, z něhož pocházejí hodnoty v souboru, a testujeme jen neznámé hodnoty parametrů. Testujeme tzv. **PARAMETRICKÉ HYPOTÉZY**
- viz následující testy:

Test parametru p rozdělení $A(p)$

Používáme testové kritérium

$$u = \frac{\hat{p} - p}{\sqrt{p\,(1-p)}}\,\sqrt{n}\,,$$

kde $\hat{p} = \bar{x}$ je odhad parametru p,

p je testovaný parametr (daný v H_0),

n je rozsah souboru (test požaduje n tak velké, aby $n\hat{p}(1-\hat{p}) \geq 9$).

Je-li H_0 pravdivá, pak $u \approx N(0,1)$.

Test parametru μ rozdělení $X \sim N(\mu, \sigma^2)$, tzv. **t-test**

Testové kritérium:

$$t = \frac{\bar{x} - \mu}{s} \sqrt{n}$$

Je-li H_0 pravdivá , pak t má t-rozdělení ppsti s $n-1$ stupni volnosti, tj.

$$t \sim t(\nu), \quad \text{kde } \nu = n - 1.$$

Pozn.: Pro velký rozsah n souboru (řekněme $n > 30$), platí $t(\nu) \approx N(0,1)$. Kritickými hodnotami jsou pak příslušné kvantily u_p rozdělení $N(0,1)$.

Test parametru σ^2 rozdělení $X \sim N(\mu, \sigma^2)$

Testové kritérium:

$$\chi^2 = \frac{(n-1)\,s^2}{\sigma^2}\ ,$$

kde n je rozsah souboru,
s^2 je výběrový rozptyl,
σ^2 je testovaný rozptyl (daný v H_0).

Je-li H_0 pravdivá, pak χ^2 má χ^2-rozdělení ppsti s $n-1$ stupni volnosti, tj.

$$\chi^2 \sim \chi^2(\nu), \quad \text{kde } \nu = n-1.$$

Jiný druh hypotéz jsou tzv. **NEPARAMETRICKÉ HYPOTÉZY** - např. testy, zda naměřené hodnoty "shodují"s určitým typem rozdělení.

Často používaný je **chí-kvadrát test dobré shody**:

$n \ldots$ **rozsah** souboru - tj. počet všech (ne nutně různých) naměřených hodnot experimentu

$k \ldots$ **všechny možné výsledky** experimentu

pro $i = 1, 2, \ldots, k$ označme: $n_i \ldots$ **pozorovaná četnost** i-tého výsledku,
$n_i^O \ldots$ **očekávaná četnost** i-tého výsledku.

testové kritérium:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(n_i - n_i^O\right)^2}{n_i^O}$$

Je-li n tak velké, $n_i^O \geq 5 \ \ \forall\, i = 1, \ldots, k$, pak

$$\boxed{\chi^2 \approx \chi^2(\nu)\,, \text{ kde } \nu = k - 1}$$

K zamítnutí H_0 : *"Není (významný) rozdíl mezi pozorovanými a očekávanými četnostmi* vede "velká"hodnota χ^2, tj. test je **pravostranný**.

χ^2-test dobré shody lze použít i pro hodnoty získané ze *spojitých rozdělení*, jestliže jak pozorované tak očekávané četnosti náleží k nějakým disjunktním třídám (intervalům):

4. KONTINGENČNÍ TABULKY
(χ^2-test nezávislosti)

Testujme např. zda technologie má vliv na podíl vadných výrobků. Pozorované četnosti n_{ij} $(i = 1, 2, 3;\ j = 1, 2)$ zapišme ve formě tzv. **kontingenční tabulky**

n_{ij}	Vyhovující výrobky	Vadné výrobky
Technologie A	226	24
Technologie B	136	14
Technologie C	95	5

kontingence … "statistická"závislost jedné náhodné veličiny (např. technologie), ozn. ji X na jiné náhodné veličině (např. vyhovující/vadný výrobek), ozn. ji Y

Testujeme hypotézu H_0: **"X, Y jsou nezávislé.**

.

Testové kritérium:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(n_{ij} - n_{ij}^O\right)^2}{n_{ij}^O} \,,$$

kde r je počet možných hodnot veličiny X (**řádky**),

s je počet možných hodnot veličiny Y (**sloupce**),

n_{ij} … **pozorované četnosti**

n_{ij}^O … **očekávané četnosti** příslušné k n_{ij}

Za předpokladu platnosti H_0 je

$$n_{ij}^O = \frac{n_{i\cdot}\ n_{\cdot j}}{n} \,,$$

kde $n_{i\cdot}$ je součet četností v i-tém řádku tabulky,

$n_{\cdot j}$ je součet četností v j-tém sloupci tabulky,

n je celkový počet pozorování v tabulce.

Jsou-li **všechny očekávané četnosti aspoň 5**, pak $\chi^2 \approx \chi^2(\nu)$, kde $\nu = (r-1)(s-1)$.

5. TESTY VÝBĚRŮ Z ROZDĚLENÍ $N_2(\vec{\mu}, \Sigma)$

TEST NEZÁVISLOSTI

Víme (viz Náh. vektor): Pro $(X, Y) \sim N_2(\vec{\mu}, \Sigma)$ platí:
X, Y jsou nezávislé $\Leftrightarrow$ korelační koeficient $\varrho = 0$.

Jsou-li (x_i, y_i), $i = 1, 2, \ldots, n$ realizace náh. výběru, odhadujeme $\varrho \approx r$, kde

$$\boxed{r = \frac{\frac{1}{n}\sum x_i\, y_i - \bar{x}\;\bar{y}}{s_x \cdot s_y}}$$

$\bar{x}$ je aritm. průměr hodnot $x_1, \ldots, x_n$,
$\bar{y}$ - " - $y_1, \ldots, y_n$,
s_x je výběr. směrodatná odchylka hodnot $x_1, \ldots, x_n$,
s_y - " - $y_1, \ldots, y_n$.

$r \ldots$ **výběrový korelační koeficient**

Platí: $-1 \leq r \leq +1$,
kde $r \approx 0$ naznačuje *lineární nezávislost* x_i a y_i (resp. nezávislost - pro rozdělení N_2),
$r \to \pm 1$ naznačuje silnou *lineární závislost* (resp. závislost pro N_2).

K testu nezávislosti (při výběrech z N_2) používáme testové kritérium

$$t = \frac{r}{\sqrt{1-r^2}}\sqrt{n-2}$$

Platí-li $H_0 : \varrho = 0$, pak $t \sim t(\nu)$, kde $\nu = n - 2$.

(Alternativa $H_1 : \varrho \neq 0 \Rightarrow$ test je oboustranný.)

Zamítneme-li H_0 na hladině $\alpha = 0.05$, resp. $\alpha \ = \ 0.01$, říkáme, že r je **významný**, resp. **vysoce významný**, a hodnoty x_i a y_i považujeme za závislé.

t-TESTY SHODY DVOU STŘEDNÍCH HODNOT

Nechť $[x_i, y_i]$ je náhodný výběr rozsahu n z rozdělení $(X, Y) \sim N_2(\vec{\mu}, \Sigma)$, kde $\vec{\mu} = (\mu_1, \mu_2)$, $\mu_1 = E(X)$, $\mu_2 = E(Y)$.

Testujeme

$$\boxed{H_0 : \mu_1 = \mu_2}$$

• ***t*-test pro závislé výběry** - tzv. **párový *t*-test**

Nechť $z_i = y_i - x_i$ $(i = 1, \ldots, n)$ jsou realizace náhodné veličiny $Z = Y - X$ se střední hodnotou $\mu = \mu_2 - \mu_1$.

Pak $H_0 : \mu = 0$ a testové kritérium je

$$t = \frac{\overline{z}}{s}\sqrt{n}$$

kde $\overline{z}$ je aritm. průměr hodnot $z_1, \ldots, z_n$,
s je jejich výběr. směrodatná odchylka.

Při platnosti H_0 je $t \sim t(\nu)$, kde $\nu = n - 1$.

• **t-test pro nezávislé výběry**

Výběr rozsahu n_1 z rozdělení $N(\mu_1, \sigma_1^2)$:
$x_1, x_2, \ldots, x_{n_1}$, arim. průměr $\bar{x}$, výběr.rozptyl s_1^2

Výběr rozsahu n_2 z rozdělení $N(\mu_2, \sigma_2^2)$:
$y_1, y_2, \ldots, y_{n_2}$, aritm. průměr $\bar{y}$, výběr.rozptyl s_2^2

Testové kritérium :

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{1}{n_1}s_1^2 + \frac{1}{n_2}s_2^2}}$$

Platí-li $H_0 : \mu_1 = \mu_2$, je $t \sim t(\nu)$, kde $\nu = min(n_1, n_2) - 1$.

F-TEST SHODY DVOU ROZPTYLŮ

Označme dva nezávislé náhodné výběry z normálního rozdělení tak, aby $s_1^2 \geq s_2^2$.

Testové kritérium:

$$f = \frac{s_1^2}{s_2^2}$$

Platí-li $H_0 : \sigma_1^2 = \sigma_2^2$, pak $f \sim F(\nu_1, \nu_2)$, kde $\nu_1 = n_1 - 1,\ \nu_2 = n_2 - 1$.

$F(\nu_1, \nu_2)$ je tzv. **F-rozdělení** s ν_1 a ν_2 stupni volnosti

Test je pravostranný a kvantily (=kritické hodnoty testu) F-rozdělení jsou v tabulkách.

6. REGRESE, METODA NEJMENŠICH ČTVERCŮ

dáno: $[x_i, y_i], i = 1, 2 \ldots, n$

$x_i \ldots$ dány "přesně"či se zanedbatelnou chybou

$y_i \ldots$ určeny "nepřesně"($\approx$ realizace náh. veličiny)

$y_i \ldots$ naměřené hodnoty **neznámé** funkce $y = f(x)$ v bodech x_i, tj. $y_i \approx f(x_i)$

cíl: určit funkci $f(x)$, resp. její odhad $\hat{f}(x)$, které tyto (s "chybou") naměřené hodnoty "nejlépe"odpovídají

Regresní přímka

Předpokládáme, že f je má tvar

$$f(x) = b_0 + b_1\, x$$

a odhady parametrů $\beta_0 \approx b_0$, $\beta_1 \approx b_1$ najdeme tzv. **metodou nejmenších čtverců** - hledáme, pro která β_0, β_1 je minimální součet čtverců

$$S = \sum_{i=1}^{n} (\beta_0 + \beta_1 x_i - y_i)^2.$$

Metodami mat. analýzy lze ověřit, že hledaná β_0, β_1 jsou řešením tzv. **soustavy normálních rovnic**

$$\beta_0 \cdot n + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i\, y_i$$

Regresní parabola

Pro funkci f tvaru

$$f(x) = b_0 + b_1\, x + b_2 x^2$$

odhady $\beta_0 \approx b_0$, $\beta_1 \approx b_1, \beta_2 \approx b_2$ metodou nejm. čtverců jsou ty, které minimalizují výraz

$$S = \sum_{i=1}^{n} (\beta_0 + \beta_1 x_i + \beta_2 x_i^2 - y_i)^2$$

a jsou řešením soustavy

$$\begin{aligned}
\beta_0 \cdot n + \beta_1 \sum_{i=1}^{n} x_i + \beta_2 \sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} y_i \\
\beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2 + \beta_2 \sum_{i=1}^{n} x_i^2 &= \sum_{i=1}^{n} x_i\, y_i \\
\beta_0 \sum_{i=1}^{n} x_i^2 + \beta_1 \sum_{i=1}^{n} x_i^3 + \beta_2 \sum_{i=1}^{n} x_i^4 &= \sum_{i=1}^{n} x_i^2\, y_i
\end{aligned}$$

(DODATKY K PPSTI)

Intenzita poruch, Weibullovo rozdělení

X...náh. veličina, doba do poruchy (součástky, zařízení)
X...spojitá, $f(x)$...hustota, $F(x)$...distribuční fce,
(Protože $x > 0$ je čas, je $f(x) = 0$ pro $x \leq 0$.)

Definujme **spolehlivostní funkci** $R(x)$ vztahem

$$R(x) = P(X > x) = 1 - F(x)\,, \tag{1}$$

intenzitu poruch $h(x)$ vztahem

$$h(x) = \frac{f(x)}{R(x)} = \frac{f(x)}{1 - F(x)}\,. \tag{2}$$

Výpočtem pro libovolné $x_0 > 0$ dostaneme:

$$P(X < x_0 + \Delta\,|X > x_0) = \frac{P(x_0 < X < x_0 + \Delta)}{P(X > x_0)} \approx$$

$$\approx \frac{f(x_0)\,\Delta}{1 - F(x_0)} = h(x_0)\,\Delta$$

K popisu životnosti se často hodí použít tzv. Weibullovo rozdělení, které je zobecněním exponenciálního rozdělení.

WEIBULLOVO rozdělení s parametry $c > 0$, $\delta > 0$:

X má *distribuční funkci*

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0] \\ 1 - e^{-\left(\frac{x}{\delta}\right)^c} & \text{pro } x \in (0, +\infty) \end{cases}$$

Píšeme: $X \sim W(c, \delta)$.

[$W(1, \delta)$ je zřejmě $Exp(\delta)$.]

Hustota ppsti:

$$f(x) = F'(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0] \\ \frac{c}{\delta^c}\, x^{c-1}\, e^{-\left(\frac{x}{\delta}\right)^c} & \text{pro } x \in (0, +\infty) \end{cases}$$

Je-li $X \sim W(c, \delta)$, pak intenzita poruch (viz (2)):

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{c}{\delta} \left(\frac{x}{\delta} \right)^{c-1}$$

Pro $c = 1$ [tj. je-li $X \sim Exp(\delta)$] dostaneme :

$h(x) = \frac{1}{\delta}$ (= **konstantní** funkce) pro $x > 0$.

Pro $0 < c < 1$ je $h(x)$ **klesající** funkce pro $x > 0$.

Pro $c > 1$ je $h(x)$ **rostoucí** funkce pro $x > 0$.

V praxi často:
nejprve $0 < c < 1$ (etapa častých poruch),
pak $c \approx 1$ (etapa ustáleného provozu),
nakonec $c > 1$ (etapa stárnutí).

Grafem $h(x)$ je pak tzv. **vanová křivka**.

Součet náhodných veličin, gama rozdělení

X_1 ... spojitá s hustotou ppsti $f_1(x_1)$ a distr. fcí $F_1(x_1)$

X_2 ... spojitá s hustotou ppsti $f_2(x_2)$ a distr. fcí $F_2(x_2)$

$\vec{X} = (X_1, X_2)$ ... se sdruženou hustotou ppsti $f(x_1,\, x_2)$

Předpokládejme, že X_1, X_2 jsou **nezávislé**.
[Pak $f(x_1,\, x_2) = f_1(x_1) \cdot f_2(x_2)$].

Nechť $\mathbf{Y} = \mathbf{X_1} + \mathbf{X_2}$.

Pak Y je náh. veličinou, označme:
$g(y)$ ... hustota Y,
$G(y)$ ... distrib. fce Y

$g(y) = ?$
$G(y) = ?$

Přímým výpočtem $G(y)$ a její derivací dle y dostaneme

$$g(y) = \int_{-\infty}^{+\infty} f_1(y - x_2) \cdot f_2(x_2)\, dx$$

Výraz vpravo je tzv. KONVOLUCE hustot $f_1,\, f_2$, píšeme $(\mathbf{f_1} * \mathbf{f_2})(\mathbf{y})$.

Platí $(f_1 * f_2)(y)$=$(f_2 * f_1)(y)$ (”symetrie”).

Součet n nezávislých náhodných veličin s rozdělením $Exp(\delta)$ má **gama rozdělení** s parametry n, δ.

GAMA rozdělení s parametry $m > 0$, $\delta > 0$:

X má *hustotu ppsti*

$$f(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0] \\ \frac{x^{m-1}}{\Gamma(m)\,\delta^m}\, e^{-\left(\frac{x}{\delta}\right)} & \text{pro } x \in (0, +\infty) \end{cases}$$

Píšeme: $X \sim \Gamma(m, \delta)$.

Poznámky:

1) **Gama funkce** $\Gamma(m)$ je definována takto:

$$\Gamma(m) = \int_0^{+\infty} t^{m-1} \cdot e^{-t}\, dt\,.$$

2) Volbou $m = \frac{\nu}{2}$, $\delta = 2$, kde $\nu \in N$, dostaneme χ^2**-rozdělení** (s ν stupni volnosti)