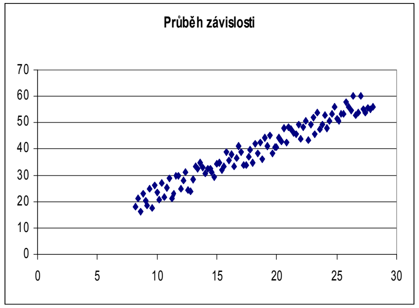
Regresní analýza

Cíl regresní analýzy: vystižení závislosti hodnot znaku Y na hodnotách znaku X. Při tom je nutné vyřešit dva problémy:

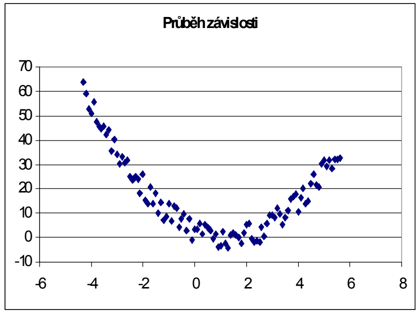
- jaký typ funkce použít k vystižení dané závislosti
- jak stanovit konkrétní parametry zvoleného typu funkce?

Typ funkce určíme buď logickým rozborem zkoumané závislosti nebo se snažíme ho odhadnout pomocí dvourozměrného tečkového diagramu.



přímka

$$y = \beta_0 + \beta_1 x$$



parabola

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

Regresní přímka

Zde se omezíme na lineární závislost $y = \beta_0 + \beta_1 x$.

Odhady b_0 a b_1 neznámých regresních parametrů β_0, β_1 získáme na základě datového souboru $\begin{pmatrix} x_1 & y_1 \\ \dots & \dots \\ x_n & y_n \end{pmatrix}$ metodou nejmenších čtverců. Požadujeme, aby výraz $\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$ nabýval svého minima vzhledem k β_0 a β_1. Tento výraz je minimální, jsou-li jeho první derivace podle β_0 a β_1 nulové. Stačí tyto derivace spočítat, položit je rovny 0 a řešit systém dvou rovnic o dvou neznámých, tzv. systém normálních rovnic.

Nechť je dán dvourozměrný datový soubor $\begin{pmatrix} x_1 & y_1 \\ \dots & \dots \\ x_n & y_n \end{pmatrix}$ a přímka $y = \beta_0 + \beta_1 x$.

Výraz $q(\beta_0, \beta_1) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$ se nazývá rozptyl hodnot znaku Y kolem přímky $y = \beta_0 + \beta_1 x$,

přímka $y = b_0 + b_1 x$, jejíž parametry minimalizují rozptyl $q(\beta_0, \beta_1)$ v celém dvourozměrném prostoru, se nazývá regresní přímka znaku Y na znak X,

$\hat{y}_i = b_0 + b_1 x_i$, $i = 1, \dots, n$ … regresní odhad i-té hodnoty znaku Y,

$r_{12}^2 = ID^2$ … index determinace (Index determinace udává, jakou část variability hodnot znaku Y vystihuje regresní přímka. Nabývá hodnot z intervalu $\langle 0, 1 \rangle$. Čím je bližší 1, tím lépe vystihuje regresní přímka závislost Y na X.)

Odvození odhadů regresních parametrů

Systém normálních rovnic získáme derivováním výrazu

$q(\beta_0,\beta_1)=\frac{1}{n}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)^2$ parciálně podle β_0 a β_1:

$$\frac{\partial q(\beta_0,\beta_1)}{\partial \beta_0}=\frac{2}{n}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)(-1)=0$$

$$\frac{\partial q(\beta_0,\beta_1)}{\partial \beta_1}=\frac{2}{n}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)(-x_i)=0$$

Systém normálních rovnic:

$$\sum y_i = nb_0 + b_1 \sum x_i$$
$$\sum y_i x_i = b_0 \sum x_i + b_1 \sum x_i^2$$

Řešením tohoto systému získáme odhady

$$b_0=\frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2},\quad b_1=\frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2-\left(\sum_{i=1}^{n} x_i\right)^2}$$

$$y = m_2 + \frac{s_{12}}{s_1^2}(x-m_1)$$

Po jednoduchých úpravách dospějeme ke tvaru $b_1=\frac{s_{12}}{s_1^2}$, kde s_{12} je kovariance znaků X, Y a s_1^2 je rozptyl znaku X. Dále dostáváme $b_0=m_2-b_1 m_1$, tedy regresní přímku můžeme vyjádřit ve tvaru $y=m_2+\frac{s_{12}}{s_1^2}(x-m_1)$.

Úsek b_0 regresní přímky udává, jaký je regresní odhad hodnoty znaku Y, nabývá-li znak X hodnoty 0.
Směrnice b_1 udává, o kolik jednotek se změní hodnota znaku Y, změní-li se hodnota znaku X o jednotku. Je-li $b_1 > 0$, dochází s růstem X k růstu Y a hovoříme o přímé závislosti hodnot znaku Y na hodnotách znaku X. Je-li $b_1 < 0$, dochází s růstem X k poklesu Y a hovoříme o nepřímé závislosti hodnot znaku Y na hodnotách znaku X.

Příklad

Příklad: Pro datový soubor obsahující údaje o mezi plasticity (znak X) a mezi pevnosti oceli (znak Y)

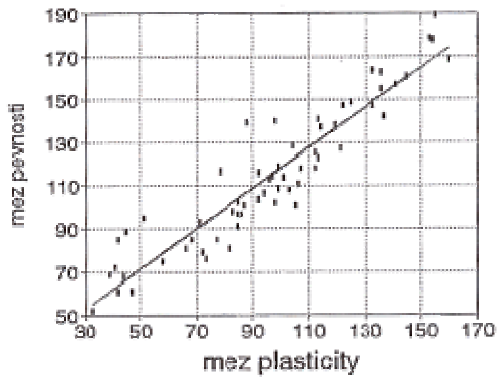
a) Určete regresní přímku meze pevnosti na mez plasticity.
b) Zakreslete regresní přímku do dvourozměrného tečkového diagramu.
c) Jak se změní mez pevnosti, vzroste-li mez plasticity o jednotku?
d) Najděte regresní odhad meze pevnosti pro mez plasticity = 60.
e) Vypočtěte index determinace a interpretujte ho.

Přitom již víme, že $m_1 = 95{,}5$, $m_2 = 114{,}4$, $s_1 = 32{,}4$, $s_2 = 32{,}5$, $s_{12} = 985{,}76$, $r_{12} = 0{,}936$.

Řešení:

ad a) $b_1 = \frac{s_{12}}{s_1^2} = \frac{985{,}76}{1052{,}4} = 0{,}937$, $b_0 = m_2 - b_1 m_1 = 114{,}4 - 0{,}937 \cdot 95{,}9 = 24{,}5$, $y = 24{,}5 + 0{,}937x$.

ad b)



ad c) Mez pevnosti vzroste o 0,937 kpcm^{-2} – viz parametr b_1 vypočtený v bodě (a)

ad d) $\hat{y} = 24{,}5 + 0{,}937 \times 60 = 80{,}72$.

ad e) $ID^2 = r_{12}^2 = 0{,}936^2 = 0{,}876$. Znamená to, že 87,6% variability hodnot meze pevnosti je vysvětleno regresní přímkou.

Příklad

x_i	y_i	$x_i{*}y_i$	x_i^2
100	120	12 000	10 000
90	105	9 450	8 100
86	95	8 170	7 396
94	100	9 400	8 836
120	135	16 200	14 400
135	140	18 900	18 225
79	102	8 058	6 241
62	98	6 076	3 844
110	125	13 750	12 100
125	134	16 750	15 625
1 001	**1 154**	**118 754**	**104 767**

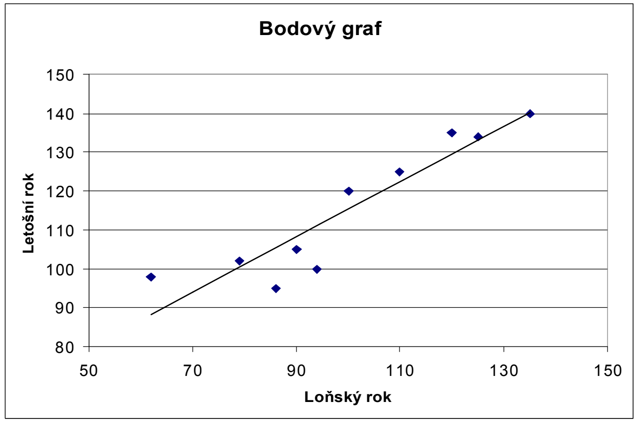
$$\sum y_i = nb_0 + b_1 \sum x_i$$

$$\sum y_i x_i = b_0 \sum x_i + b_1 \sum x_i^2$$

$$1154 = 10 \cdot b_0 + 1001 \cdot b_1$$

$$118754 = 1001 \cdot b_0 + 104767 \cdot b_1$$

$$y = 44{,}41 + 0{,}709 \cdot x$$



Příklad

Index determinace lze vyjádřit ve tvaru:

$$ID^2 = \frac{\sum \hat{y}_i^2 - \frac{1}{n}\cdot\left(\sum y_i\right)^2}{\sum y_i^2 - \frac{1}{n}\cdot\left(\sum y_i\right)^2}$$

x_i	y_i	$\hat{y}_i$	y_i^2	$\hat{y}_i^2$
100	120	115	14 400	13 301
90	105	108	11 025	11 715
86	95	105	9 025	11 109
94	100	111	10 000	12 337
120	135	130	18 225	16 773
135	140	140	19 600	19 642
79	102	100	10 404	10 088
62	98	88	9 604	7 811
110	125	122	15 625	14 987
125	134	133	17 956	17 704
1 001	**1 154**	**1 154**	**135 864**	**135 468**

$$ID^2 = \frac{135468 - \frac{1}{10}\cdot 1154^2}{135864 - \frac{1}{10}\cdot 1154^2} = 0{,}853$$

Maticové vyjádření MNČ

$$\boxed{b = \left(X^T \cdot X\right)^{-1} \cdot X^T \cdot y}$$

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

b	sloupcový vektor 2 neznámých parametrů regresní funkce,
X	matice rozměru $n \times 2$, tvořená konstantou 1 a hodnotami znaku X
y	sloupcový vektor n hodnot znaku Y

Příklad

Nalezněte koeficienty regresní přímky:

$$\mathbf{y} = \begin{bmatrix} 120 \\ 105 \\ 95 \\ 100 \\ 135 \\ 140 \\ 102 \\ 98 \\ 125 \\ 134 \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & 100 \\ 1 & 90 \\ 1 & 86 \\ 1 & 94 \\ 1 & 120 \\ 1 & 135 \\ 1 & 79 \\ 1 & 62 \\ 1 & 110 \\ 1 & 125 \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 100 & 90 & 86 & 94 & 120 & 135 & 79 & 62 & 110 & 125 \end{bmatrix}$$

Příklad

$$g = X^T * y = \begin{bmatrix} 1154 \\ 118754 \end{bmatrix} \qquad A = X^T * X = \begin{bmatrix} 10 & 1001 \\ 1001 & 104767 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 2{,}2941 & -0{,}0219 \\ -0{,}0219 & 0{,}0002 \end{bmatrix}$$

$$b = A^{-1} * g = \begin{bmatrix} 44{,}414 \\ 0{,}709 \end{bmatrix}$$

Sdružené regresní přímky

V některých situacích má smysl zkoumat nejenom závislost znaku Y na znaku X, ale též závislost X na Y. V takovém případě hledáme druhou regresní přímku a souhrnně hovoříme o sdružených regresních přímkách.
Regresní přímkou znaku X na znak Y nazveme tu přímku x = $\bar{b}_0 + \bar{b}_1 y$, jejíž parametry minimalizují rozptyl q($\bar{\beta}_0, \bar{\beta}_1$) = $\sum_{i=1}^{n}(x_i - \bar{\beta}_0 - \bar{\beta}_1 y_i)^2$ v celé rovině. Nazývá se též druhá regresní přímka. Regresní přímka znaku Y na znak X a regresní přímka znaku X na znak Y se nazývají sdružené regresní přímky.

Rovnice regresní přímky znaku X na znak Y má tvar:

$$\boxed{x = m_1 + \frac{s_{12}}{s_2^2}\left(x - m_2\right)}$$

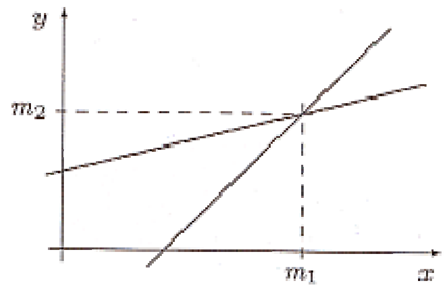
Vlastnosti sdružených regresních přímek

Sdružené regresní přímky se protínají v bodě (m_1,m_2).
Pro regresní parametry $b_1, \overline{b}_1$ platí: $b_1\overline{b}_1 = r_{12}^2$.
Rovnice sdružených regresních přímek můžeme psát ve tvaru

$y = m_2 + r_{12}\frac{s_2}{s_1}(x - m_1)$, $y = m_2 + \frac{1}{r_{12}}\frac{s_2}{s_1}(x - m_1)$ (je-li $r_{12} \neq 0$).

Regresní přímky svírají tím menší úhel, čím méně se od sebe liší r_{12} a $\frac{1}{r_{12}}$.

Regresní přímky splynou, je-li $r_{12}^2 = 1$. K tomu dojde právě tehdy, existuje-li mezi X a Y úplná lineární závislost. Všechny body (x_i, y_i), $i = 1, \dots, n$ leží na jedné přímce, tedy ze znalosti x_i můžeme přesně vypočítat y_i, $i = 1, \dots, n$.
Jsou-li znaky X, Y nekorelované, pak mají sdružené regresní přímky rovnice $y = m_2$, $x = m_1$ a jsou na sebe kolmé.
Označíme-li α úhel, který svírají sdružené regresní přímky, pak platí:
$\cos\alpha = 0$, právě když mezi X a Y neexistuje žádná lineární závislost,
$\cos\alpha = 1$, právě když mezi X a Y existuje úplná přímá lineární závislost,
$\cos\alpha = -1$, právě když mezi X a Y existuje úplná nepřímá lineární závislost.



Příklad

Příklad: Pro datový soubor obsahující údaje o mezi plasticity (znak X) a mezi pevnosti oceli (znak Y):

a) Určete regresní přímku meze plasticity na mez pevnosti.

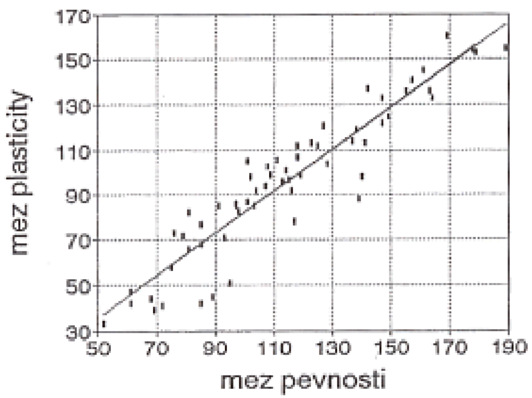
b) Zakreslete tuto druhou regresní přímku do dvourozměrného tečkového diagramu.

Přitom již víme, že $m_1 = 95{,}5$, $m_2 = 114{,}4$, $s_1 = 32{,}4$, $s_2 = 32{,}5$, $s_{12} = 985{,}76$, $r_{12} = 0{,}936$.

Řešení:

ad a) $\bar{b}_1 = \frac{s_{12}}{s_2{}^2} = \frac{985{,}76}{1057{,}21} = 0{,}932, \bar{b}_0 = m_1 - \bar{b}_1 m_2 = 95{,}9 - 0{,}932 \times 114{,}4 = -10{,}7$, tedy x = -10,7 + 0,932y.

ad b)



Příklad

Poptávka po vepřovém mase	154	164	123	181	193	105	143	167	158	62
Poptávka po hovězím mase	103	116	98	175	165	90	103	140	113	49

- Sestrojte sdružené regresní přímky.
- Vypočtěte koeficient korelace.

Příklad

x_i	y_i	x_i*y_i	x_i^2	y_i^2
154	103	15 862	23 716	10 609
164	116	19 024	26 896	13 456
123	98	12 054	15 129	9 604
181	175	31 675	32 761	30 625
193	165	31 845	37 249	27 225
105	90	9 450	11 025	8 100
143	103	14 729	20 449	10 609
167	140	23 380	27 889	19 600
158	113	17 854	24 964	12 769
62	49	3 038	3 844	2 401
1 450	**1 152**	**178 911**	**223 922**	**144 998**

$$m_1 = \frac{1}{10} \cdot 1450 = 145 \qquad m_2 = \frac{1}{10} \cdot 1152 = 115{,}2$$

$$s_1^2 = \frac{1}{10} \cdot 223922 - 145^2 = 1367{,}2$$

$$s_2^2 = \frac{1}{10} \cdot 144998 - 115{,}2^2 = 1228{,}76$$

$$s_{12} = \frac{178911}{10} - 145 \cdot 115{,}2 = 1187{,}1$$

$$b_1 = \frac{1187{,}1}{1367{,}2} = 0{,}868$$

$$\overline{b}_1 = \frac{1187{,}1}{1228{,}76} = 0{,}966$$

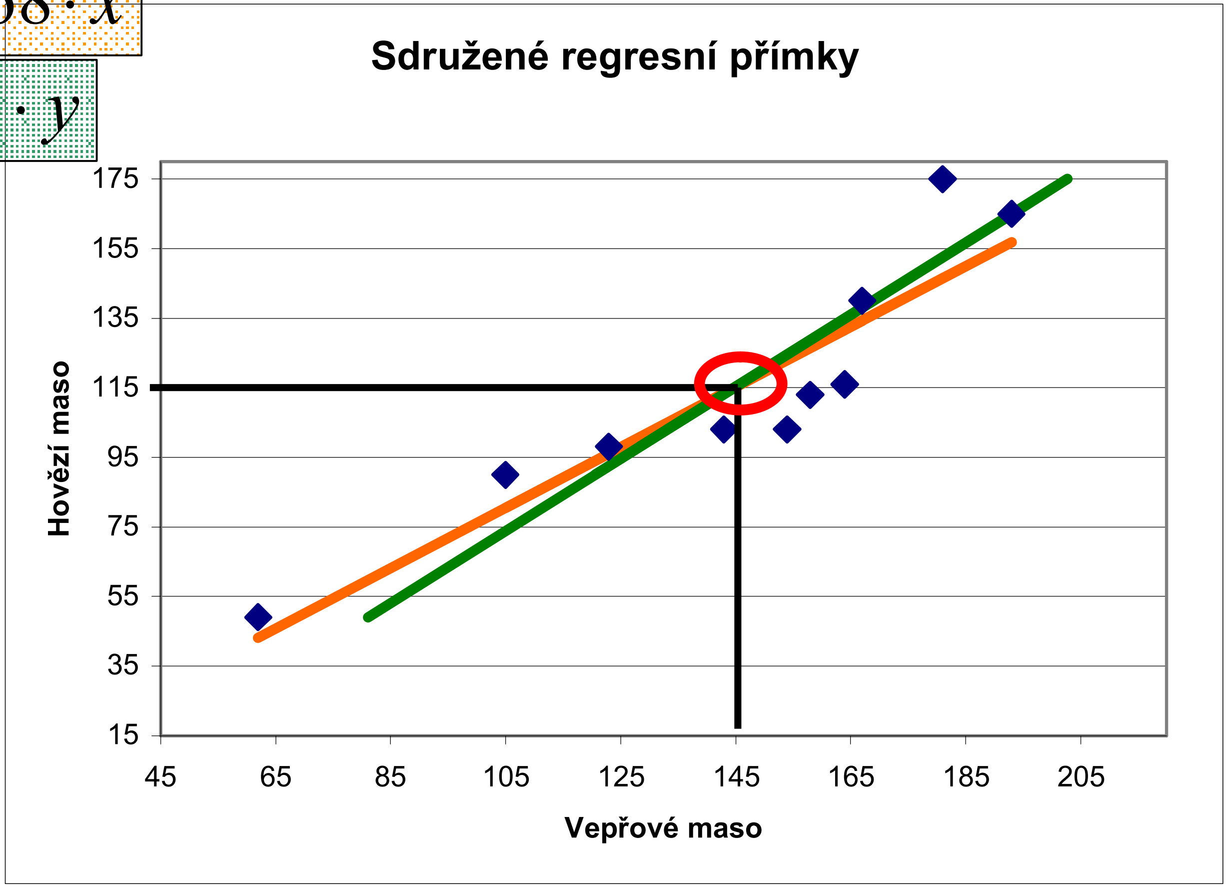
$$b_0 = 115{,}2 - 0{,}868 \cdot 145 = -10{,}66$$

$$\overline{b}_0 = 145 - 0{,}966 \cdot 115{,}2 = 33{,}72$$

Příklad

$$y = -10{,}66 + 0{,}868 \cdot x$$

$$x = 33{,}72 + 0{,}966 \cdot y$$



Příklad

$$r_{12} = \frac{\sum x_i y_i - n m_1 m_2}{\sqrt{\left[\sum x_i^2 - n m_1^2\right] \cdot \left[\sum y_i^2 - n m_2^2\right]}}$$

$$r_{12} = \frac{s_{12}}{s_1 \cdot s_2}$$

$$r_{12} = \mathrm{sgn}(\mathrm{b}_1)\sqrt{b_1 \cdot \overline{b_1}}$$

Příklad

$$r = \frac{178911 - 10 \cdot 145 \cdot 115{,}2}{\sqrt{\left[223922 - 10 \cdot 145^2\right] \cdot \left[144998 - 10 \cdot 115{,}2^2\right]}} = 0{,}916$$

$$r = \frac{1187{,}1}{36{,}976 \cdot 35{,}054} = 0{,}916$$

$$r = \sqrt{0{,}868 \cdot 0{,}966} = 0{,}916$$

Příklad

x_i	y_i
154	103
164	116
123	98
52	175
193	165
105	90
143	103
167	140
158	113
191	49

- Sestrojte sdružené regresní přímky.
- Vypočtěte koeficient korelace.
- Porovnejte výsledky z výsledky předchozího příkladu.

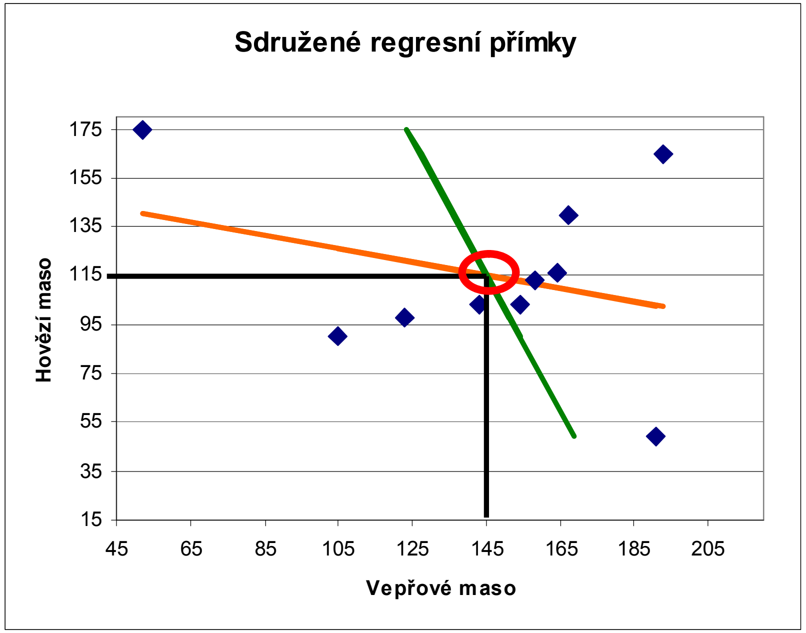
Příklad

x_i	y_i	x_i*y_i	x_i^2	y_i^2
154	103	15 862	23 716	10 609
164	116	19 024	26 896	13 456
123	98	12 054	15 129	9 604
52	175	9 100	2 704	30 625
193	165	31 845	37 249	27 225
105	90	9 450	11 025	8 100
143	103	14 729	20 449	10 609
167	140	23 380	27 889	19 600
158	113	17 854	24 964	12 769
191	49	9 359	36 481	2 401
1 450	**1 152**	**162 657**	**226 502**	**144 998**

průměry	145,0
	115,2
rozptyly	1625,2
	1228,76
směrodatné odchylky	40,3138
	35,0537

$$y = 154{,}31 - 0{,}270 \cdot x$$

$$x = 186{,}09 - 0{,}357 \cdot y$$



$$r = -\sqrt{0{,}270 \cdot 0{,}357} = -0{,}310$$

Příklad

Rozhodněte zda následující dvojice přímek mohou být sdruženými regresními přímkami:

$$A)\ y = 13 - 2x \qquad x = 2{,}5$$

$$B)\ y = 13 - 2x \qquad x = 0{,}4y$$

$$C)\ y = 13 - 2x \qquad x = 8 - y$$

$$D)\ y = 13 - 2x \qquad x = 6{,}5 - 0{,}5y$$

$$E)\ y = 13 - 2x \qquad x = -2 - 0{,}4y$$

$$F)\ y = 13 - 2x \qquad x = -0{,}5y$$

Příklad

$$A)\ y = 13 - 2x \qquad B)\ y = 13 - 2x \qquad C)\ y = 13 - 2x$$
$$x = 2{,}5 \qquad x = 0{,}4y \qquad x = 8 - y$$
$$D)\ y = 13 - 2x \qquad E)\ y = 13 - 2x \qquad F)\ y = 13 - 2x$$
$$x = 6{,}5 - 0{,}5y \qquad x = -2 - 0{,}4y \qquad x = -0{,}5y$$

1. b_1 i $\overline{b_1}$ mají stejná znaménka
2. je-li jeden roven nule, pak je 0-vý i druhý
3. $r \in [-1, 1]$,tj. $b_1 \cdot \overline{b_1} \in [0, 1]$
4. pro r=1/r=-1 platí $\overline{b_0} = -\dfrac{b_0}{b_1}$

A) NE(2)
B) NE(1)
C) NE(3)
D) ANO
E) ANO
F) NE(4)

Počet pravděpodobnosti -úvod

Počet pravděpodobnosti se zabývá studiem zákonitostí v náhodných pokusech. Matematickými prostředky modeluje situace, v nichž hraje roli náhoda. Pod pojmem náhoda rozumíme působení faktorů, které se živelně mění při různých provedeních téhož pokusu a nepodléhají naší kontrole.

Počet pravděpodobnosti jako vědecká disciplína se začal vytvářet **v 17. století** a jeho počátky jsou spjaty se jmény Blaise Pascala, Pierra de Fermata, Christiana Huygense (studovali hazardní hry, zformulovali takové pojmy, jako je pravděpodobnost a střední hodnota, odvodili jejich vlastnosti) a především Jakoba Bernoulliho (dokázal zákon velkých čísel).

V 18. století: Abraham de Moivre a Pierre Simeon Laplace – formulace jedné z forem centrální limitní věty, Georges Buffon odvodil binomickou větu, zavedl diferenciální a integrální počet do teorie pravděpodobnosti, Thomas Bayes odvodil způsob výpočtu aposteriorních pravděpodobností pomocí apriorních pravděpodobností (Bayesův vzorec).

V 19. století: Petrohradská matematická škola – dala teorii pravděpodobnosti pevný logický a matematický základ (Viktor Jakovlevič Buňakovskij, Pafnutij Lvovič Čebyšev, Andrej Andrejevič Markov, Alexandr Michailovič Ljapunov), Karl Fridirich Gauss (mj. vyvinul netodu zpracování experimentálních údajů známou pod názvem metoda nejmenších čtverců), Siméon Denis Poisson (zobecnil Bernoulliho zákon velkých čísel a odvodil speciální zákon rozložení pravděpodobností - Poissonův zákon rozložení).

Ve 20. století: Andrej Nikolajevič Kolmogorov (axiomatická teorie pravděpodobnosti), Norbert Wiener , William Feller (rozvoj reorie stochastických procesů).

Odkaz na zajímavou webovou stránku:

http://www-groups.dcs.st-and.ac.uk/~history

http://turnbull.mcs.st-and.ac.uk/~history/HistTopics/Statistics.html

Základní prostor

1.1. Definice (definice pokusu): Pokusem rozumíme jednorázové uskutečnění konstantně vymezeného souboru definičních podmínek. Předpokládáme, že pokus můžeme mnohonásobně nezávisle opakovat za dodržení definičních podmínek (ostatní podmínky se mohou měnit, proto různá opakování pokusu mohou vést k různým výsledkům). Dále předpokládáme, že opakováním pokusu vzniká opět pokus.
Deterministickým pokusem nazýváme takový pokus, jehož každé opakování vede k jedinému možnému výsledku.
Náhodným pokusem nazýváme takový pokus, jehož každé opakování vede k právě jednomu z více možných výsledků, které jsou vzájemně neslučitelné.
Příklad deterministického pokusu: při tlaku 1015 hPa zahříváme vodu na 100 °C. Jediným možným výsledkem je var vody.
Příklady náhodných pokusů: hod hrací kostkou, hod mincí, vylosování čísla z osudí apod.

1.2. Definice (definice základního prostoru): Neprázdnou množinu možných výsledků náhodného pokusu značíme Ω a nazýváme ji základní prostor. Možné výsledky značíme ω_t, kde $t \in T$, T je indexová množina.

Příklad

1.3. Příklad

a) Náhodný pokus spočívá v hodu kostkou. Možný výsledek ω_i znamená polohu kostky číslem i nahoru, i = 1, …, 6. Základní prostor $\Omega = \{\omega_1, \ldots, \omega_6\}$, počet možných výsledků $m(\Omega) = 6$.

b) Náhodný pokus spočívá v hodu dvěma kostkami. Možný výsledek je uspořádaná dvojice $[\omega_i, \omega_j]$, i, j = 1, …, 6. Základní prostor $\Omega = \{[\omega_1, \omega_1], [\omega_1, \omega_2], \ldots, [\omega_1, \omega_6], \ldots, [\omega_6, \omega_6]\}$, počet možných výsledků $m(\Omega) = 6^2 = 36$.

c) Náhodný pokus spočívá v opakovaném házení mincí tak dlouho, dokud nepadne první líc. Potom základní prostor $\Omega = \{\omega_1, \omega_2, \omega_3 \ldots\}$, kde ω_1 znamená, že hned v prvním hodu padl líc, ω_2 znamená, že až ve druhém hodu padl líc, ω_3 znamená, že až ve třetím hodu padl líc atd. Sybmolicky lze zapsat $\omega_1 = [L]$, $\omega_2 = [R, L]$, $\omega_3 = [R, R, L]$, … Tedy základní prostor Ω má nekonečně spočetně mnoho možných výsledků.

Jevové pole

1.4. Definice (definice jevového pole): Systém podmnožin $\mathcal{A}$ základního prostoru Ω, který splňuje následující tři axiómy:

J5: $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 - A_2 \in \mathcal{A}$,

J6: $\Omega \in \mathcal{A}$,

J8: $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

se nazývá jevové pole. Jestliže $A \in \mathcal{A}$, pak řekneme, že A je jev. Dvojice $(\Omega, \mathcal{A})$ se nazývá měřitelný prostor.

(Axióm J5 nám říká, že jevové pole obsahuje s každými dvěma množinami i jejich množinový rozdíl. Axióm J6 říká, že jevové pole obsahuje celý základní prostor a konečně axióm J8 říká, že když jevové pole obsahuje každou ze spočetné posloupnosti množin, obsahuje i jejich spočetné sjednocení. Znamená to, že systém $\mathcal{A}$ je uzavřený vzhledem k množinovým operacím.

Protože jevy jsou množiny, pro operace s nimi platí stejné zákony jako pro operace s množinami - komutativní zákon, asociativní zákon, de Morganova pravidla.)

Množinové a pravděpodobnostní pojmy

1.5. Poznámka (slovník množinových a pravděpodobnostních pojmů)

Ω se nazývá jistý jev, $\varnothing$ se nazývá nemožný jev

$\omega \in A$ znamená, že možný výsledek ω je příznivý nastoupení jevu A

$A \subseteq B$ znamená, že jev A má za důsledek jev B

$A \cup B$ znamená nastoupení aspoň jednoho z jevů A, B

$A \cap B$ znamená společné nastoupení jevů A, B

$A - B$ znamená nastoupení jevu A za nenastoupení jevu B

$\overline{A} = \Omega - A$ znamená jev opačný k jevu A

$A \cap B = \varnothing$ znamená, že jevy A, B jsou neslučitelné.

Příklad

1.6. Příklad: Je dán systém složený ze dvou bloků, který jednorázově použijeme. Nechť jev A_i znamená bezporuchovou funkci i-tého bloku, i = 1, 2. Pomocí jevů A_1, A_2 vyjádřete jevy:

a) bezporuchová funkce aspoň jednoho bloku: $A_1 \cup A_2$

b) bezporuchová funkce obou bloků: $A_1 \cap A_2$

c) porucha aspoň jednoho bloku: $\overline{A_1} \cup \overline{A_2}$

d) porucha obou bloků: $\overline{A_1} \cap \overline{A_2}$

e) porucha právě jednoho bloku: $(\overline{A_1} \cap A_2) \cup (A_1 \cap \overline{A_2})$

Jevové pole - poznámky

1.7. Poznámka: Systém axiómů jevového pole je bezesporný (tj. na každém základním prostoru lze sestrojit aspoň jedno jevové pole) a neúplný (tzn., že na každém aspoň dvouprvkovém základním prostoru lze vytvořit jevových polí více).

Neúplnost systému axiómů jevového pole je výhodná, protože umožňuje rozlišovat výsledky náhodného pokusu s různým stupněm podrobnosti.

Např. jevové pole $\mathcal{A}_{min} = \{\Omega, \emptyset\}$ se nazývá minimální jevové pole a charakterizuje krajně „tupozrakého“ pozorovatele, který rozliší pouze jev jistý a jev nemožný.

Jevové pole $\mathcal{A}_1 = \{\Omega, \emptyset, A, \overline{A}\}$ již dovolí rozeznat, zda nastal jev A nebo jev opačný $\overline{A}$.

Tak můžeme konstruovat stále bohatší jevová pole, až dostaneme maximální jevové pole $\mathcal{A}_{max} = \{A; A \subseteq \Omega\}$. To charakterizuje krajně „bystrozrakého“ pozorovatele, který rozliší jevy do všech podrobností. Pro libovolné jevové pole $\mathcal{A}$ ovšem platí: $\mathcal{A}_{min} \subseteq \mathcal{A} \subseteq \mathcal{A}_{max}$.

Příklad

1.8. Příklad: Sestrojte všechna možná jevová pole na základním prostoru $\Omega = \{\omega_1, \omega_2, \omega_3\}$.

Řešení:

$\mathcal{A}_1 = \{\Omega, \emptyset\} (= \mathcal{A}_{\min})$

$\mathcal{A}_2 = \{\Omega, \emptyset, \{\omega_1\}, \{\omega_2, \omega_3\}\}$

$\mathcal{A}_3 = \{\Omega, \emptyset, \{\omega_2\}, \{\omega_1, \omega_3\}\}$

$\mathcal{A}_4 = \{\Omega, \emptyset, \{\omega_3\}, \{\omega_1, \omega_2\}\}$

$\mathcal{A}_5 = \{\Omega, \emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}\} (= \mathcal{A}_{\max})$

Jevové pole - vlastnosti

1.9. Věta (vlastnosti jevového pole): Nechť $(\Omega, \mathcal{A})$ je měřitelný prostor. Pak jevové pole $\mathcal{A}$ má následujících 9 vlastností:

J1: $\mathcal{A} \neq \emptyset$,

J2: $\emptyset \in \mathcal{A}$,

J3: $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cup A_2 \in \mathcal{A}$,

J4: $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cap A_2 \in \mathcal{A}$,

J5: $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 - A_2 \in \mathcal{A}$ (axióm),

J6: $\Omega \in \mathcal{A}$ (axióm),

J7: $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$,

J8: $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ (axióm),

J9: $A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{A}$.

Důkaz: J1 plyne z J6.

J2 plyne z J5 a J6, protože $\Omega - \Omega = \emptyset$.

J3 plyne z J2 J8 speciální volbou $A_3 = \emptyset, A_4 = \emptyset, \ldots$ Pak $\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2$.

J7 plyne J5 a J6, protože $\overline{A} = \Omega - A$.

J9 odvodíme z J7 a J8 užitím de Morganových pravidel $\bigcap_{i=1}^{\infty} A_i = \overline{\bigcup_{i=1}^{\infty} \overline{A_i}}$:

$A_1, A_2, \ldots \in \mathcal{A} \Rightarrow \overline{A_1}, \overline{A_2}, \ldots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} \overline{A_i} \in \mathcal{A} \Rightarrow \overline{\bigcup_{i=1}^{\infty} \overline{A_i}} \in \mathcal{A}$, ovšem $\overline{\bigcup_{i=1}^{\infty} \overline{A_i}} = \bigcap_{i=1}^{\infty} A_i$.

J4 plyne z J9 speciální volbou $A_3 = \Omega, A_4 = \Omega, \ldots$

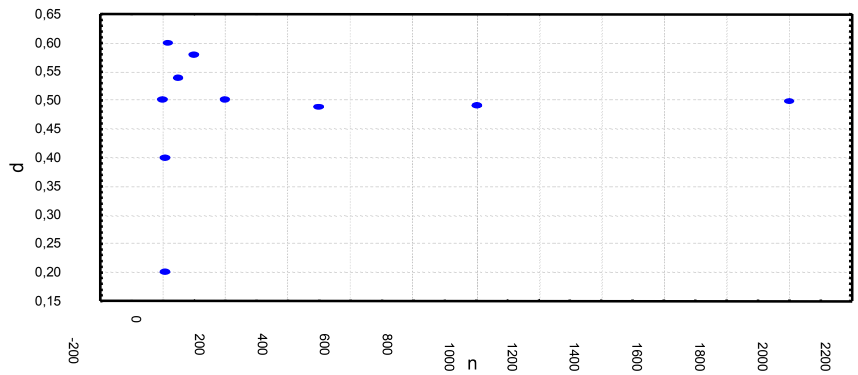
Pravděpodobnostní prostor

2.1. Motivace: Provádíme opakovaně nezávisle týž náhodný pokus a v každém pokusu sledujeme nastoupení jevu A, kterému říkáme úspěch. Označme n celkový počet pokusů a N(A) počet těch pokusů, kdy nastal úspěch. S rostoucím n pozorujeme, že relativní četnost úspěchu $\frac{N(A)}{n}$ se blíží číslu P(A), které považujeme za pravděpodobnost úspěchu. (Tento poznatek je znám jako **empirický zákon velkých čísel**).

Ilustrace empirického zákona velkých čísel

Provádíme n nezávislých hodů mincí. Padnutí líce považujeme za úspěch. Budeme sledovat závislost relativní četnosti úspěchu na počtu pokusů. (Počet pokusů volíme 2, 5, 10, 20, 50, 100, 200, 500,1000, 2000.)

n	2	5	10	20	50	100	200	500	1000	2000
p	0,5	0,2	0,4	0,6	0,54	0,58	0,5	0,488	0,49	0,4975



Axiomatická teorie pravděpodobnosti

Vzniká otázka, jak zavést pravděpodobnost, aby byla „zidealizovaným“ protějškem relativní četnosti. Zdálo by se vhodné zavést pravděpodobnost takto:

$$P(A) = \lim_{n\to\infty} \frac{N(A)}{n}.$$

Jde o tzv. **statistickou definici pravděpodobnosti**. Z matematického hlediska tato definice není v pořádku, protože počet pokusů je vždy konečný a nelze se přesvědčit o existenci uvedené limity. Proto ve 30. letech 20. století ruský matematik A. A. Kolmogorov (1903– 1987) vybudoval **axiomatickou teorii pravděpodobnosti.**

Axiomatická teorie pravděpodobnosti zavádí pravděpodobnost jako funkci, která každému jevu přiřazuje číslo mezi 0 a 1 a přitom je zidealizovaným protějškem relativní četnosti. Má tedy všechny vlastnosti relativní četnosti a kromě toho některé další vlastnosti, které vyplývají z vnitřních potřeb matematické teorie.

Pravděpodobnost - definice

2.2. Definice: Nechť $(\Omega, \mathcal{A})$ je měřitelný prostor. Reálná množinová funkce $P: \mathcal{A} \to R$ se nazývá pravděpodobnost, když splňuje následující 3 axiómy:

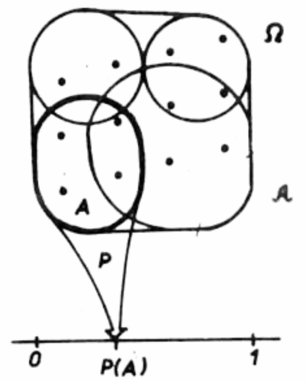
P2: $\forall A \in \mathcal{A}: P(A) \geq 0$ (nezápornost)

P10: $P(\Omega) = 1$ (normovanost)

P15: $A_1, A_2, \ldots \in \mathcal{A}$ jsou neslučitelné $\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ (spočetná aditivita)

Trojice $(\Omega, \mathcal{A}, P)$ se nazývá pravděpodobnostní prostor. (Je to matematický model jednorázového provedení náhodného pokusu.)

Ilustrace pravděpodobnostního prostoru



2.3. Poznámka: Systém axiómů pravděpodobnosti je bezesporný (tj. na každém měřitelném prostoru lze sestrojit pravděpodobnost) a neúplný (tj. na každém měřitelném prostoru, jehož jevové pole není minimální, lze sestrojit pravděpodobností více).

Pravděpodobnost - vlastnosti

2.4. Věta (vlastnosti pravděpodobnosti): Nechť (Ω, $\mathcal{A}$, P) je pravděpodobnostní prostor, $A, A_1, A_2, \ldots \in \mathcal{A}$ libovolné jevy. Pak pravděpodobnost P má následujících 17 vlastností:

P1: $P(\emptyset) = 0$

P2: $P(A) \geq 0$ (nezápornost – axióm)

P3: $P(A_1 \cup A_2) + P(A_1 \cap A_2) = P(A_1) + P(A_2)$

P4: $1 + P(A_1 \cap A_2) \geq P(A_1) + P(A_2)$

P5: $P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$ (subaditivita)

P6: $A_1 \cap A_2 = \emptyset \Rightarrow P(A_1 \cup A_2) = P(A_1) + P(A_2)$ (aditivita)

P7: $P(A_2 - A_1) = P(A_2) - P(A_1 \cap A_2)$

P8: $A_1 \subseteq A_2 \Rightarrow P(A_2 - A_1) = P(A_2) - P(A_1)$ (subtraktivita)

P9: $A_1 \subseteq A_2 \Rightarrow P(A_2) \leq P(A_1)$ (monotonie)

P10: $P(\Omega) = 1$ (normovanost – axióm)

P11: $P(A) + P(\bar{A}) = 1$ (komplementarita)

P12: $P(A) \leq 1$

P13: $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$ (spočetná subaditivita)

P14: $A_1, A_2, \ldots \in \mathcal{A}$ jsou neslučitelné $\Rightarrow \sum_{i=1}^{\infty} P(A_i) < \infty$ (absolutní konvergence)

P15: $A_1, A_2, \ldots \in \mathcal{A}$ jsou neslučitelné $\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ (spočetná aditivita – axióm)

P16: $A_1 \subseteq A_2 \subseteq \ldots \in \mathcal{A} \Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i \to \infty} P(A_i)$ (spojitost pravděpodobnosti zdola)

P17: $A_1 \supseteq A_2 \supseteq \ldots \in \mathcal{A} \Rightarrow P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i \to \infty} P(A_i)$ (spojitost pravděpodobnosti shora)

Pravděpodobnost - vlastnosti

P14 Položme $A_0 = \overline{\bigcup_{i=1}^{\infty} A_i}$. Pak jevy $A_0, A_1, A_2, \ldots$ jsou neslučitelné a jejich sjednocením je celý základní prostor, tedy podle axiómu P10 dostáváme: $1 = P(\Omega) = P(\bigcup_{i=0}^{\infty} A_i) = \sum_{i=0}^{\infty} P(A_i)$, přičemž poslední rovnost vyplývá z axiómu P15. $\sum_{i=0}^{\infty} P(A_i)$ tedy absolutně konverguje, tudíž bude konvergovat také $\sum_{i=1}^{\infty} P(A_i)$, kde jsme vynechali první člen.

P1 Položme $A_1 = \emptyset, A_2 = \emptyset, \ldots$ Pak $\bigcup_{i=1}^{\infty} A_i = \emptyset$, tedy podle axiómu P15 $0 = P(\emptyset) = P(\bigcup_{i=1}^{\infty} \emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$, což je možné jen tak, že $P(\emptyset) = 0$.

P6 V axiómu P15 položíme $A_3 = \emptyset$, $A_4 = \emptyset, \ldots$, tedy $P(\bigcup_{i=1}^{\infty} A_i) = P(A_1 \cup A_2) = \sum_{i=1}^{\infty} P(A_i) = P(A_1) + P(A_2)$.

P11 Plyne z vlastnosti P6 a axiómu P10: $P(A \cup \overline{A}) = P(\Omega) = 1 = P(A) + P(\overline{A})$.

P12 Plyne okamžitě z axiómu P2 a vlastnosti P11.

Pravděpodobnost - vlastnosti

Pro důkaz vlastností P3, P4 a P5 jevy $A_1 \cup A_2$, A_1 a A_2 rozložíme na součet disjunktních sčítanců:

$A_1 \cup A_2 = (A_1 \setminus A_2) \cup (A_1 \cap A_2) \cup (A_2 \setminus A_1)$

$A_1 = (A_1 \setminus A_2) \cup (A_1 \cap A_2)$

$A_2 = (A_2 \setminus A_1) \cup (A_1 \cap A_2)$

P3 Podle P6 dostáváme: $P(A_1 \cup A_2) + P(A_1 \cap A_2) = P(A_1 \setminus A_2) + P(A_1 \cap A_2) + P(A_2 \setminus A_1) + P(A_1 \cap A_2) = P(A_1) + P(A_2)$.

Protože podle P12 je $P(A_1 \cup A_2) \leq 1$ a podle P2 je $P(A_1 \cap A_2) \geq 0$, dostáváme z P3 okamžitě P4 a P5.

P7 Opět vyjádříme A_2 jako sjednocení neslučitelných jevů: $A_2 = (A_2 \setminus A_1) \cup (A_1 \cap A_2)$. Podle P3 pak dostaneme: $P(A_2) = P(A_2 \setminus A_1) + P(A_1 \cap A_2)$, tedy $P(A_2 \setminus A_1) = P(A_2) - P(A_1 \cap A_2)$.

P8 Jelikož $A1 \subseteq A_2$, platí $A_1 \cap A_2 = A_1$ a P8 plyne z P7.

P9 Plyne z P8, protože podle P2 je $P(A_2 \setminus A_1) \geq 0$, tudíž $P(A_2) - P(A_1) \geq 0$, tj. $P(A_1) \leq P(A_2)$.

Pravděpodobnost - vlastnosti

P13 Položíme $\bigcup_{i=1}^{\infty} A_i = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cup A_2)) \cup \ldots$ Tím jsme dostali sjednocení posloupnosti neslučitelných jevů a aplikujeme axióm P15 a vlastnost P7: $P(\bigcup_{i=1}^{\infty} A_i) = P(A_1) + P(A_2 \setminus A_1) + P(A_3 \setminus (A_1 \cup A_2)) + \ldots \leq P(A_1) + P(A_2) + P(A_3) + \ldots = \sum_{i=1}^{\infty} P(A_i)$.

P16 Jev $\bigcup_{i=1}^{\infty} A_i$ vyjádříme jako sjednocení neslučitelných jevů. Z předpokladu $A_1 \subseteq A_2 \subseteq \ldots$ plyne $\bigcup_{i=1}^{\infty} A_i = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \ldots \cup (A_i \setminus A_{i-1}) \cup \ldots$, tedy podle axiómu P15 a vlastnosti P8 dostáváme: $P(\bigcup_{i=1}^{\infty} A_i) = P(A_1) + P(A_2 \setminus A_1) + P(A_3 \setminus A_2) + \ldots + P(A_i \setminus A_{i-1}) + \ldots = P(A_1) + [P(A_2) - P(A_1)] + [P(A_3) + P(A_2)] + \ldots + [P(A_i) + P(A_{i-1})] + \ldots = \lim_{i \to \infty} P(A_i)$.

P17 Podle vlastnosti P16 dostáváme $P(\overline{\bigcup_{i=1}^{\infty} A_i}) = \lim_{i \to \infty} P(\overline{A_i})$. Z de Morganových pravidel plyne $P(\bigcap_{i=1}^{\infty} A_i) = P(\overline{\bigcup_{i=1}^{\infty} \overline{A_i}}) = 1 - P(\bigcup_{i=1}^{\infty} \overline{A_i}) = 1 - \lim_{i \to \infty} P(\overline{A_i}) = 1 - \lim_{i \to \infty} [1 - P(A_i)] = \lim_{i \to \infty} P(A_i)$.

Pravděpodobnost - vlastnosti

2.5.Věta (další vlastnosti pravděpodobnosti): Nechť $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor, $A_1, A_2, \ldots, A_n \in \mathcal{A}$ libovolné jevy. Pak platí:

a) $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} P(A_i \cap A_j) + \sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\sum_{k=j+1}^{n} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n)$

(Pro neslučitelné jevy $A_1, \ldots, A_n$ dostáváme $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.) (Věta o sčítání pravděpodobností)

b) $\max\limits_{1 \le i \le n} P(A_i) \le P\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} P(A_i)$

c) $1 - n + \sum_{i=1}^{n} P(A_i) \le P\left(\bigcap_{i=1}^{n} A_i\right) \le \min\limits_{1 \le i \le n} P(A_i)$ (nerovnost vlevo se nazývá Bonferroniho nerovnost)

Pravděpodobnost - vlastnosti

Důkaz:

ad a) Vlastnost vyjadřuje princip inkluze a exkluze. Tvrzení o neslučitelných jevech plyne z axiómu P15, kde položíme $A_{n+1} = \emptyset, A_{n+2} = \emptyset, \ldots$

ad b) Levá strana: Plyne z monotonie P9. Pro $\forall i \in \{1,\ldots,n\}$ je $A_i \subseteq \bigcup_{j=1}^{n} A_j$, tedy pro $\forall i \in \{1,\ldots,n\}$ platí $P(A_i) \leq P\left(\bigcup_{j=1}^{n} A_j\right)$.

Tvrzení musí platit i pro ten index i, pro který je $P(A_i)$ maximální.

Pravá strana: Plyne ze spočetné subaditivity P13, kde položíme $A_{n+1} = \emptyset, A_{n+2} = \emptyset, \ldots$

ad c) Levá strana: $P\left(\bigcap_{i=1}^{n} A_i\right) = P\left(\overline{\bigcup_{i=1}^{n} \overline{A_i}}\right) = 1 - P\left(\bigcup_{i=1}^{n} \overline{A_i}\right) \geq 1 - \sum_{i=1}^{n} P\left(\overline{A_i}\right) = 1 - \sum_{i=1}^{n} [1 - P(A_i)] = 1 - n + \sum_{i=1}^{n} P(A_i)$

Pravá strana: Plyne z monotonie P9. Pro $\forall i \in \{1,\ldots,n\}$ je $A_i \supseteq \bigcap_{j=1}^{n} A_j$, tedy pro $\forall i \in \{1,\ldots,n\}$ platí

$P(A_i) \geq P\left(\bigcap_{j=1}^{n} A_j\right)$. Tvrzení musí platit i pro ten index i, pro který je $P(A_i)$ minimální.

Příklad

2.6. Příklad: Je dán systém složený ze dvou bloků. Jev A_i značí bezporuchovou funkci i-tého bloku, i = 1, 2. Je známo, že $P(A_i) = \vartheta_i$, i = 1, 2.

a) Odhadněte pravděpodobnost správné funkce celého systému, jsou-li bloky zapojeny

α) sériově, β) paralelně.

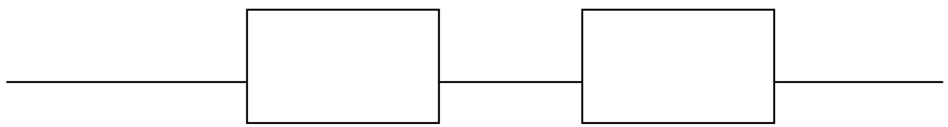
b) Předpokládejme navíc, že $P(A_1 \cap A_2) = \vartheta_{12}$. Vypočtěte nyní pravděpodobnost správné funkce celého systému, jsou-li bloky zapojeny

α) sériově, β) paralelně.

Řešení:

ad a)

Případ sériového zapojení



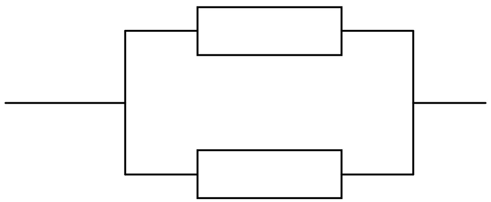
$P(A_1 \cap A_2)$ lze shora i zdola odhadnout pomocí věty 2.5. (c), kde n = 2:

$$1 - 2 + P(A_1) + P(A_2) \le P(A_1 \cap A_2) \le \min\{P(A_1), P(A_2)\}$$

$$\vartheta_1 + \vartheta_2 - 1 \le P(A_1 \cap A_2) \le \min\{\vartheta_1, \vartheta_2\}$$

Příklad

Případ paralelního zapojení



$P(A_1 \cup A_2)$ lze shora i zdola odhadnout pomocí věty 2.5. (b), kde n = 2:

$$\max\{P(A_1), P(A_2)\} \le P(A_1 \cup A_2) \le P(A_1) + P(A_2)$$

$$\max\{\vartheta_1, \vartheta_2\} \le P(A_1 \cup A_2) \le \vartheta_1 + \vartheta_2$$

ad b)

Případ sériového zapojení: $P(A_1 \cap A_2) = \vartheta_{12}$

Případ paralelního zapojení: Podle vlastnosti P3 dostáváme: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = \vartheta_1 + \vartheta_2 - \vartheta_{12}$