

5.9 Viacnásobná lineárna regresia

V predchádzajúcich modeloch sme uvažovali iba jednu závislé premennú. Ukážeme si postup pre model

$$Tr_t = a_0 + a_1 t + a_2 u + a_3 v,$$

kde a_0, a_1, a_2 sú neznáme parametre a $t = 1, 2, \dots, n$ je časová premenná. Pre odhad parametrov pomocou metódy najmenších štvorcov je

$$S(a_0, a_1, a_2, a_3) = \sum_{i=1}^n (y_i - a_0 - a_1 t_i - a_2 u_i - a_3 v_i)^2$$

Podobne ako pri lineárnej funkcii, dostávame sústavu rovníc

$$\begin{aligned} a_0 n + a_1 \sum_{i=1}^n t_i + a_2 \sum_{i=1}^n u_i + a_3 \sum_{i=1}^n v_i &= \sum_{i=1}^n y_i \\ a_0 \sum_{i=1}^n t_i + a_1 \sum_{i=1}^n (t_i)^2 + a_2 \sum_{i=1}^n t_i u_i + a_3 \sum_{i=1}^n t_i v_i &= \sum_{i=1}^n t_i y_i \\ a_0 \sum_{i=1}^n u_i + a_1 \sum_{i=1}^n u_i t_i + a_2 \sum_{i=1}^n (u_i)^2 + a_3 \sum_{i=1}^n u_i v_i &= \sum_{i=1}^n u_i y_i \\ a_0 \sum_{i=1}^n v_i + a_1 \sum_{i=1}^n v_i t_i + a_2 \sum_{i=1}^n v_i u_i + a_3 \sum_{i=1}^n (v_i)^2 &= \sum_{i=1}^n v_i y_i \end{aligned}$$

Príklad 5.9 Analytik istého podniku potrebuje zostaviť model pre tvorbu zisku z , ktorý je závislý na predaji výrobku A v počte x kusov a výrobku B v počte y kusov. Analytik si zvolil jednoduchý model

$$z = a + b \cdot x + c \cdot y$$

a použil výsledky uvedené v tabuľke:

x	12	11	15	10	11	16	14	8	8	18
y	5	8	6	5	3	9	12	4	3	10
z	72	76	78	70	68	80	82	65	62	90

Riešenie:

Riešenie sústavy lineárnych rovníc:

$$\begin{aligned} a n + b \sum_{i=1}^n x_i + c \sum_{i=1}^n y_i &= \sum_{i=1}^n z_i \\ a \sum_{i=1}^n x_i + b \sum_{i=1}^n (x_i)^2 + c \sum_{i=1}^n x_i y_i &= \sum_{i=1}^n x_i z_i \\ a \sum_{i=1}^n y_i + b \sum_{i=1}^n y_i x_i + c \sum_{i=1}^n (y_i)^2 &= \sum_{i=1}^n y_i z_i \end{aligned}$$

získame pomocou tabuľky

x	y	z	x ²	x*y	y ²	y*z	x*z
12	5	72	144	60	25	360	864
11	8	76	121	88	64	608	836
15	6	78	225	90	36	468	1170
10	5	70	100	50	25	350	700
11	3	68	121	33	9	204	748
16	9	80	256	144	81	720	1280
14	12	82	196	168	144	984	1148
8	4	65	64	32	16	260	520
8	3	62	64	24	9	186	496
18	10	90	324	180	100	900	1620
123	65	743	1615	869	509	5040	9382

$$a=47,1649$$

$$b=1,5999$$

$$c=1,1487$$

$$z = a + b \cdot x + c \cdot y = 47,1649 + 1,5999x + 1,1487y .$$

Iný spôsob riešenia (pomocou EXCEL-u):

Riešenie pomocou Excelu, využitím príkazu *Nástroje - Analýza dát - Regresia*

Pole Y – označíme hodnoty z,

Pole X – označíme hodnoty x, y

Výstup z EXCEL-u:

VÝSLEDEK

	Koeficienty	Chyba stř. hodnoty	t stat	Hodnota P	Dolní 95%	Horní 95%
Hranice	47,1649	2,4704	19,0919	0,0000	41,3233	53,0065
Soubor X 1	1,5990	0,2810	5,6913	0,0007	0,9347	2,2634
Soubor X 2	1,1487	0,3052	3,7633	0,0070	0,4269	1,8705

Úlohy

1. Realitnú kanceláriu zaujíma, ako vplýva vek domu, veľkosť pozemku a veľkosť obytných priestorov na predajnú cenu domu. Náhodne vybrala 10 domov, ktoré sú v predaji, zistené údaje sú v tabuľke.

Na analýzu závislosti použite jednoduchý model $z = a + b \cdot x + c \cdot y + d \cdot u$.

dom	vek domu v rokoch	veľkosť pozemku v m ²	veľkosť obytného priestoru v m ²	predajná cena v mil. Sk
1	8	1265	185	3,5
2	5	1350	190	3,65
3	6	1900	200	3,8
4	2	2200	250	3,97
5	9	3255	225	4,25
6	10	1875	156	2,45
7	5	2156	185	2,73
8	8	3125	190	4,55
9	7	1110	130	1,89
10	6	1280	145	2,14

$$[z = -1,64 + 0,09x + 0,0002y + 0,02u]$$

2. Vedenie banky zaujíma, ako vplývajú náklady (v tis. Sk) na reklamu v televízii, v tlači a náklady na bilbordy na celkový zisk firmy. Za týmto účelom analyzovali údaje za posledných 12 mesiacov.

Popíšte túto závislosť jednoduchým modelom $z = a + b \cdot x + c \cdot y + d \cdot u$.

mesiac	reklama v TV	reklama v tlači	bilbordy	celkový zisk
1	120	56	48	850
2	85	45	56	920
3	96	40	35	1120
4	87	80	25	650
5	45	89	36	890
6	69	90	45	826
7	23	56	75	756
8	65	56	53	845
9	110	89	58	980
10	100	65	60	1020
11	114	70	65	1120
12	128	85	70	1250

$$[z = 469,65 + 3,19x - 0,31y + 4,02u]$$