

Univerza v Ljubljani
Fakulteta *za upravo*



Prikaz korelacije in regresije za populacijo otrok v vrtcih med leti 1995 in 2005

Predmet: Statistika
Mentor:
Študent:
Številka indeksa:

Ljubljana, avgust 2006

KAZALO

1 Predstavitev podatkov.....	1
2 Tabela izračunov.....	1
3 Računanje vrednosti.....	1
3.1 Varianca x, varianca y, kovarianca.....	1
3.2 Regresijska koeficienta in regresijski konstanti.....	1
3.3 Regresijski premici.....	2
3.4 Standardna napaka ocene.....	2
3.5 Korelacijski in determinacijski koeficient.....	2
3.6 Maksimalna in minimalna vrednost po premici.....	2
3.7 Prikaz gibanja pojava.....	3
3.7.1 Slika 1: Gibanje pojava (premica x).....	3
3.7.2 Slika 2: Gibanje pojava (premica y).....	3
3.8 Uporabljene funkcije v Microsoft Excelu.....	3
3.9 Priloga 1: Izvorna tabela podatkov	4

1 Predstavitev podatkov

Na strani www.stat.si sem preko iskalnika poiskal statistični letopis 2005 (http://www.stat.si/letopis/index_letopis.asp). Izbral sem si šesto poglavje, imenovano Izobraževanje, v njem pa tabelo 6.4 Vrtci.

2 Tabela izračunov

leto	zavodi (x)	otroci (y)	xi-X	yi-Y	(xi-X) ²	(yi-Y)(xi-X)	(yi-Y) ²
1995/96	793	66553	4,25	5680,38	18,06	24141,59	32266660,14
1998/99	802	62848	13,25	1975,38	175,56	26173,72	3902106,391
1999/00	808	64151	19,25	3278,38	370,56	63108,72	10747742,64
2000/01	814	63328	25,25	2455,38	637,56	61998,22	6028866,391
2001/02	801	61803	12,25	930,38	150,06	11397,09	865597,6406
2002/03	773	58968	-15,75	-1904,63	248,06	29997,84	3627596,391
2003/04	767	54515	-21,75	-6357,63	473,06	138278,34	40419395,64
2004/05	752	54815	-36,75	-6057,63	1350,56	222617,72	36694820,64
	788,75	60872,625	0	0	3423,5	577713,25	134552785,9
	\bar{X}	\bar{Y}					

3 Računanje vrednosti

3.1 Varianca x, varianca y, kovarianca

$$\sigma_x = \text{varianca (x)} = 427,938$$

$$\sigma_y = \text{varianca (y)} = 16819098,234$$

$$C_{yx} = \text{kovarianca (Cxy)} = 72214,156$$

Iz izračunanih vrednosti izračunam regresijski koeficient in regresijski konstanti za obe regresijski premici.

3.2 Regresijska koeficienta in regresijski konstanti

$$\alpha_1 = -72228,390$$

$$\beta_1 = 168,749$$

$$\alpha_2 = 527,388$$

$$\beta_2 = 0,004$$

Regresijski koeficient prve regresijske premice (beta 1) nam pove, da se število zavodov poveča za eno enoto, če se število otrok poveča za 168,749. Regresijska konstanta prve regresijske premice (alfa 1) pa nam pove, da če ne bi bilo vrtcev, bi bilo -72228,390 otrok, kar je nesmiselno.

Regresijski koeficient druge regresijske premice (beta 2) nam pove, da se število zavodov poveča za 0,004 (kar je teoretično nemogoče), če bi rad v vrtec vstopil nov otrok. Regresijska konstanta druge regresijske premice (alfa 2) pa nam pove, da če ne bi bilo otrok, bi bilo število vrtcev dobrih 527.

3.3 Regresijski premici

$$y_i = -72228,3903 + 168,749 \cdot x_i$$
$$x_i = 527,38847 + 0,00429 \cdot y_i$$

3.4 Standardna napaka ocene

$$\sigma_{\varepsilon y} = \text{Standardna napaka ocene (y)} = 4101,109$$
$$\sigma_{\varepsilon x} = \text{Standardna napaka ocene (x)} = 20,687$$

Število otrok se je od ocenjene regresijske premice odklanjalo za dobrih 4101 otrok. Nepojasneni vplivi znašajo dobrih 21 otrok.

3.5 Korelacijski in determinacijski koeficient

$$\rho_{xy} = \frac{C_{xy}}{\sigma_y \cdot \sigma_x}$$

$$\text{Korelacijski koeficient} = 0,851$$

Izračunana vrednost kaže, da je odvisnost števila zavodov (vrtcev) od otrok pozitivna in močna.

$$\rho_{xy}^2 = \frac{\sigma_{xy}^2}{\sigma_y^2} = \frac{C_{yx}^2}{\sigma_y^2 \cdot \sigma_x^2} = (\rho_{xy})^2$$

$$\text{Determinacijski koeficient} = 0,725$$

Vrednost determinacijskega koeficienta pa pomeni, da je delež pojasnjene variance 72,5 %. To pomeni, da lahko 72,5 % variabilnosti števila zavodov pojasnimo z vplivom števila otrok, ostalih 27,5 % variabilnosti števila zavodov pa je rezultat drugih vplivov.

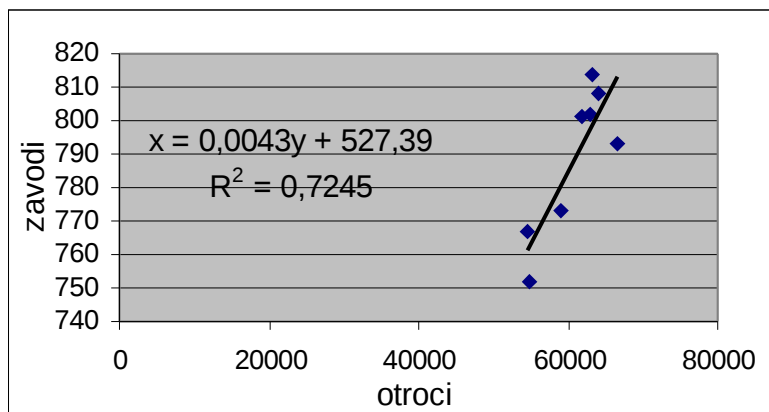
3.6 Maksimalna in minimalna vrednost po premici

$$y_i = -72228,3903 + 168,749 \cdot x_i$$
$$x_i = 527,38847 + 0,00429 \cdot y_i$$

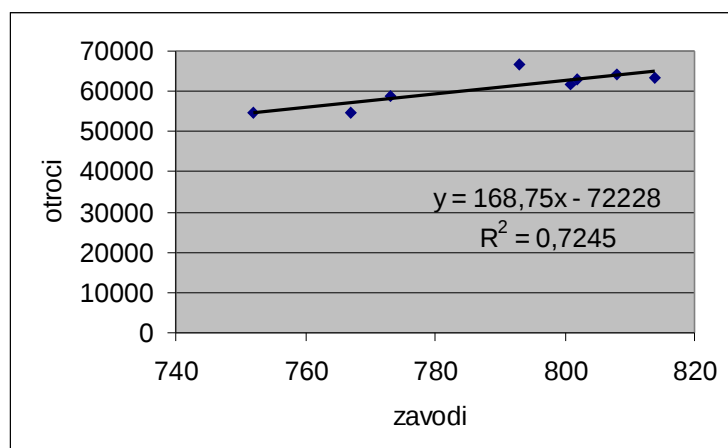
	x (zavodi)	y (otroci)
max	814	65133,545
min	752	54671,088

3.7 Prikaz gibanja pojava

3.7.1 Slika 1: Gibanje pojava (premica x)



3.7.2 Slika 2: Gibanje pojava (premica y)



3.8 Uporabljene funkcije v Microsoft Excelu

Ime funkcije	Primer dobesedne navedbe iz gradiva
AVERAGE [za aritmetično sredino]	=AVERAGE(D21:D28)
SUM	=SUM(J21:J28)
CORREL [za korelacijski koeficient]	=CORREL(D21:D28;E21:E28)
SQRT [za standardno napako ocene]	=SQRT(E33)
MAX [za najvišjo vrednost iz tabele]	=MAX(D21:D28)
MIN [za najmanjšo vrednost iz tabele]	=MIN(D21:D28)

3.9 Priloga 1: Izvorna tabela podatkov

http://www.stat.si/letopis/index_letopis

	Vrtci		otroci			zaposleno osebje za nego in vzgojo		
	zavodi	Oddelki	skupaj	deklice	dečki	skupaj	Ženske	moški
1995/96	793	3500	66553	31559	34994	6672	6555	117
1998/99	802	3455	62848	29226	33622	7012	6935	77
1999/00	808	3523	64151	30639	33512	7329	7272	57
2000/01	814	3531	63328	30350	32978	7163	7022	141
2001/02	801	3477	61803	29293	32510	7099	6956	143
2002/03	773	3342	58968	28177	30791	6949	6829	120
2003/04	767	3243	54515	25815	28700	6729	6636	93
2004/05	752	3262	54815	26033	28782	6762	6716	46

