

1. del: Korelacija

Težke kovine.

Oglejte si datoteko **waste.sav**. V njej so zbrani podatki o številu ton odpadkov, ki so jih male področne enote v ZDA zbrale za reciklažo. Odpadki so ločeni glede na tip.

- Narišite razsevne grafikone, ki prikazujejo povezanost med količinami posameznih odpadkov (postopek **Graphs->Legacy Dialogs->Scatter->Matrix Scatter...**). Kaj lahko na podlagi slike sklepate o povezanosti med spremenljivkami?
- Ali so spremenljivke paroma povezane? Ali bi povezanost lahko opisali s premico (linearna povezanost)?
- Izračunajte korelacijske koeficiente za vsak par spremenljivk! (postopek **Analyze -> Correlate -> Bivariate ->...**; vse tri spremenljivke prenesite v okno *variables*)

Correlations

		plastic	paper	organics	glass
plastic	Pearson Correlation	1			
	Sig. (2-tailed)				
	N	46	46	46	46
paper	Pearson Correlation		1		
	Sig. (2-tailed)				
	N	46	46	46	46
organics	Pearson Correlation			1	
	Sig. (2-tailed)				
	N	46	46	46	46
glass	Pearson Correlation				1
	Sig. (2-tailed)				
	N	46	46	46	46

$r_{\text{plastic,paper}}$	=
$r_{\text{plastic,organics}}$	=
$r_{\text{plastic,glass}}$	=
$r_{\text{paper,organics}}$	=
$r_{\text{paper,glass}}$	=
$r_{\text{organics,glass}}$	=

4. Interpretacija

a) korelacijskega koeficienta za plastiko in papir:

b) korelacijskega koeficienta za plastiko in steklo:

c) korelacijskega koeficienta za papir in organske odpadke:

5. Interpretacija

a) testa statistične značilnosti korelacijskega koeficienta za za plastiko in papir

$$H_0 = \underline{\hspace{2cm}}$$

$$p = \underline{\hspace{2cm}}$$

b) testa statistične značilnosti korelacijskega koeficienta za za plastiko in steklo

$$H_0 = \underline{\hspace{2cm}}$$

$$p = \underline{\hspace{2cm}}$$

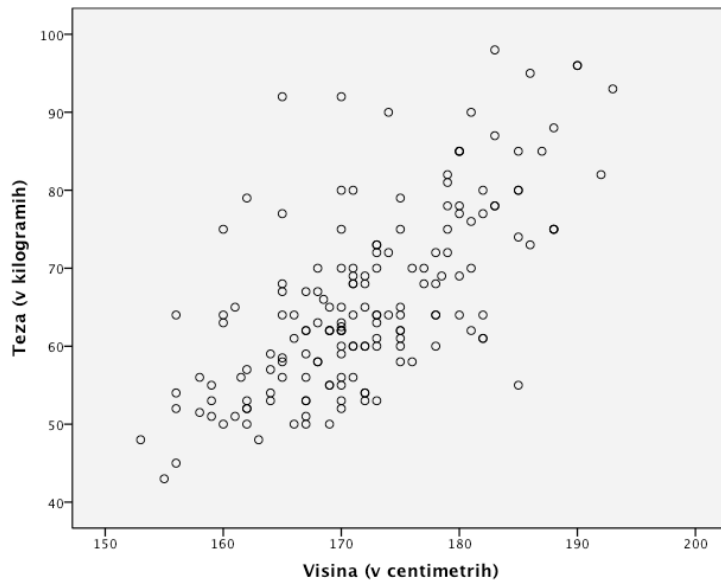
c) testa statistične značilnosti korelacijskega koeficienta za za papir in organske odpadke

$$H_0 = \underline{\hspace{2cm}}$$

2. del : Linearna regresija

1. Zaženite program SPSS in odprite datoteko **podatkiVet.sav**.

- S postopkom **Graphs** → **Legacy Dialogs** → **Scatter...** (privzeta možnost **Simple** → gumb **Define**) narišite razsevni diagram, pri čemer naj bo odvisna spremenljivka **teža** študentov (na ordinatni osi), neodvisna spremenljivka pa **višina** (na abscisni osi)! V diagram vrišite regresijsko premico (z dvoklikom v diagram odprite **Chart Editor**; s klikom na eno od točk označite podatkovni niz; dodajte regresijsko premico s klikom na ikono **Add fit line** ali preko menija **Chart** → **Add Chart Element** → **Fit Line at Total** → gumb **Close**)!



2. Približna ocena naklona premice in presečišča z y osjo na podlagi razsevnega diagrama:

$$a \approx \underline{\hspace{2cm}}$$
$$b \approx \underline{\hspace{2cm}}$$

3. Kakšna se vam zdi razpršenost točk okoli regresijske premice (obkrožite):

majhna zmerna velika

4. Kakšen korelacijski koeficient torej pričakujete?

5. S postopkom **Analyze** → **Regression** → **Linear...** za iste podatke izvedite linearno regresijsko analizo (v okence **Dependent** spravite spremenljivko teža, v okence **Independent(s)** pa višina) ter dobljeni vrednosti regresijske konstante in regresijskega koeficienta primerjajte s približno ocenjenima!

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1			,455	8,610

a. Predictors: (Constant), Višina (v centimetrih)

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8126,559	1	8126,559	109,610	,000 ^a
	Residual	9564,113	129	74,140		
	Total	17690,672	130			

a. Predictors: (Constant), Višina (v centimetrih)

b. Dependent Variable: Teža (v kilogramih)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)					
	Višina (v centimetrih)					

a. Dependent Variable: Teža (v kilogramih)

- Zapišite enačbo regresijske premice:

a = _____

b = _____

enačba regresijske premice: _____

- interpretacija regresijskega koeficienta:
- interpretacija testa statistične značilnosti regresijskega koeficienta:

H_0 : _____

SE_b = _____

t = _____

p = _____

- interpretacija korelacijskega koeficienta:

r = _____

- interpretacija determinacijskega koeficienta:

R^2 = _____

6. Vsebinski povzetek rezultatov.

7. Višina dveh študentov se razlikuje za 5 cm. Kakšno razliko pričakujete pri teži?

8. Kakšno težo pričakujete za študenta, ki je visok 190cm? Kaj pa za študenta, ki je visok 140cm? Komentirajte!

3. del : Multipla linearna regresija

1. V SPSSu preverite, kako sta višina in številka čevljev povezani s težo. Izvedite linearno regresijo.

- Ali je model statistično značilen?
- Interpretirajte korelacijski koeficient in koeficient determinacije ter ju primerjajte s tisto iz zgornje linearne regresije. V čem je razlika?

- Zapišite enačbo regresijske premice:

$$\begin{aligned} a &= \underline{\hspace{2cm}} \\ b_{\text{višina}} &= \underline{\hspace{2cm}} \\ b_{\text{štČev}} &= \underline{\hspace{2cm}} \end{aligned}$$

enačba regresijske premice: _____

- Interpretirajte regresijske koeficiente in pripadajoče statistične teste:

Dodatna naloga

Vzvodne točke.

Odprite spletno stran

http://bcs.whfreeman.com/ips4e/cat_010/applets/CorrelationRegression.html

(oziroma v Google vpišite »correlation regression applet« in ta zadetek bo eden prvih)

- a. Izberite kvadratega **Show least-square line**. Z miško dodajte 5 točk, ki med seboj linearno korelirajo, v diagram in si oglejte ocenjeno premico.
- b. Dodajte točko, ki zelo odstopa od trenda. Kaj se zgodi z regresijsko premico? Razložite!

- c. Izdelajte razsevni diagram, ki ima vrednost korelacijskega koeficienta:
 - a. $r = 1$,
 - b. $r = -1$,
 - c. $r = 0$
- d. V levem spodnjem kotu koordinatne ravnine narišite oblak točk, ki ne kažejo nikakršnega trenda. Potem narišite eno točko v desnem zgornjem kotu. Ali dobljeni razsevni diagram kaže linearno povezanost? Zakaj (glej točko a.)?

- e. Ali visoka vrednost r vedno pomeni, da so podatki linearno povezani? (narišite protiprimer)

Povzetek

Oglejte si *korelacijski koeficient* in *regresijski koeficient* pri naslednjih 4 grafih in komentirajte.

