TECHNICKÁ UNIVERZITA V LIBERCI

Nové možnosti rozvoje vzdělávání na Technické univerzitě v Liberci

Specifický cíl A3: Tvorba nových profesně zaměřených studijních programů

NPO_TUL_MSMT-16598/2022

Téma 9: Příklad 2 – Korelační analýza

Ing. Vladimíra Hovorková Valentová, Ph.D.

Financováno Evropskou unií NextGenerationEU

Národní plán obnovy

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

Zadání příkladu:

Máme zadány tyto údaje o proměnných *x* a *y*:

x_i	5	5	7	8	9	9	11	12	14	15
y_i	25	23	22	22	20	19	19	17	16	16

Vypočtěte rovnice sdružených regresních přímek a interpretujte hodnoty obou sdružených regresních koeficientů. Dále vypočítejte hodnotu korelačního koeficientu a koeficientu determinace – obě hodnoty interpretujte. Ověřte významnost koeficientu korelace pomocí vhodného testu na hladině významnosti 5 %.

Řešení příkladu:

Vzhledem k vyšší časové náročnosti všech výpočtů je příklad řešen jen prostřednictvím programu SPSS. Data o proměnných x a y zadáme do dvou samostatných sloupců:

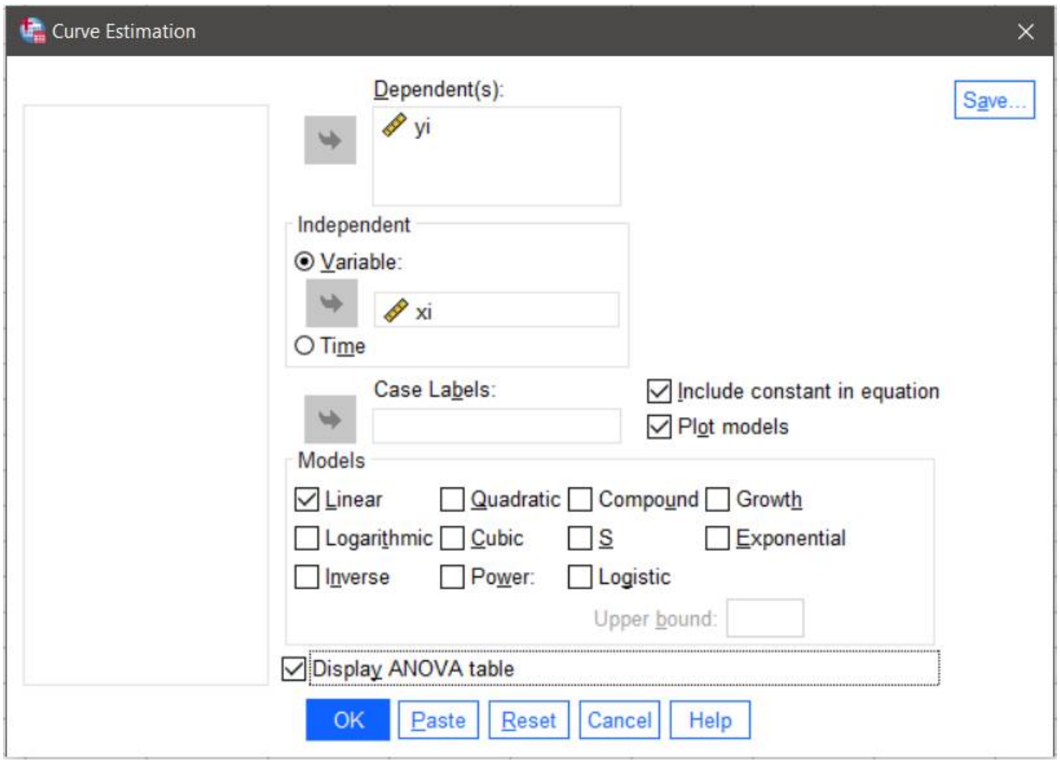
xi	yi
5	25
5	23
7	22
8	22
9	20
9	19
11	19
12	17
14	16
15	16

Pro výpočet rovnic sdružených regresních přímek zvolíme Analyze – Regression – Curve Estimation. Pokud budeme nejprve konstruovat přímku, která popisuje závislost proměnné y na proměnné x (tj. y vystupuje jako závisle proměnná, x jako nezávisle proměnná), vyplníme vstupní panel následovně:

Financováno Evropskou unií NextGenerationEU

Národní plán obnovy

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY



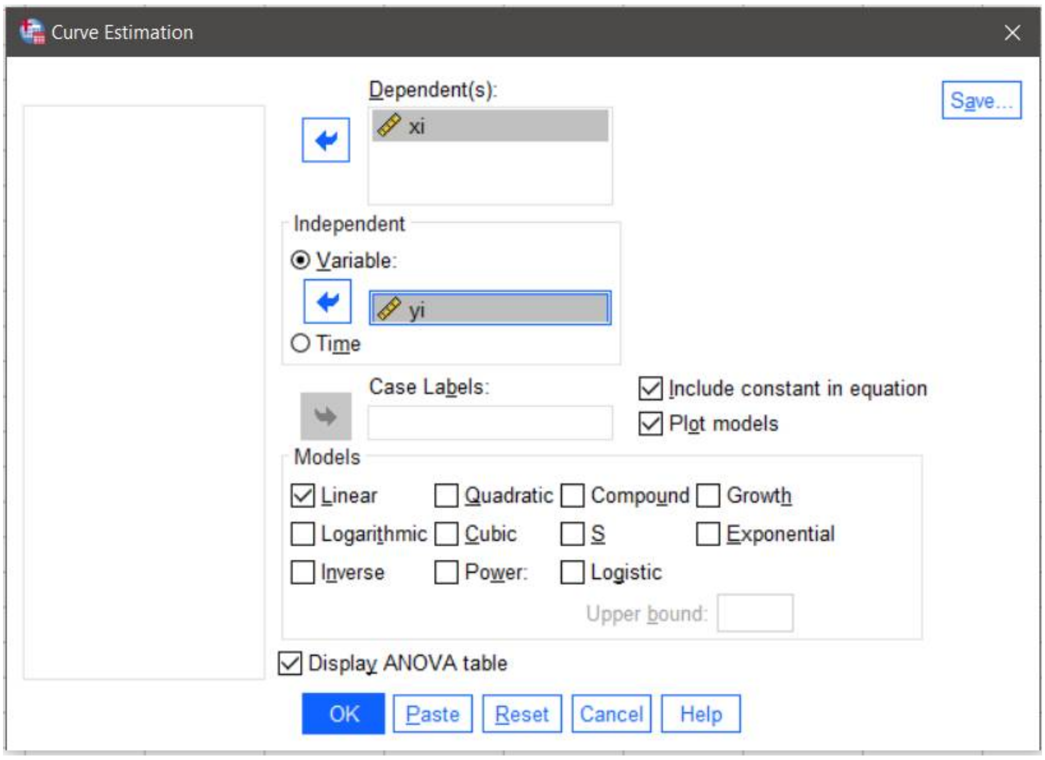
Odhady parametrů přímky popisující závislost y na x najdeme v této tabulce (sloupec označený jako B), která je součástí výstupu:

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
xi	-,853	,083	-,964	-10,220	<,001
(Constant)	27,999	,839		33,380	<,001

Na základě těchto údajů zkonstruujeme regresní přímku: $\boldsymbol{Y = 27{,}999 - 0{,}853x}$.

Nyní přistoupíme ke konstrukci regresní přímky, která popisuje závislost proměnné x na proměnné y (tj. x vystupuje v roli závisle proměnné, y v roli nezávisle proměnné). Opět vybereme posloupnost procedur Analyze – Regression – Curve Estimation a vstupní panel vyplníme takto:



Odhady parametrů přímky popisující závislost x na y najdeme v této tabulce (sloupec označený jako B), která je součástí výstupu:

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
yi	-1,090	,107	-,964	-10,220	<,001
(Constant)	31,181	2,144		14,542	<,001

Na základě těchto údajů zkonstruujeme regresní přímku: $\boldsymbol{X = 31,181 - 1,090y}$.

Rovnice sdružených regresních přímek, které popisují vzájemnou závislost proměnných x a y, jsou tedy: $\boldsymbol{Y = 27,999 - 0,853x}$

$\boldsymbol{X = 31,181 - 1,090y}$

Hodnotu *koeficientu determinace* najdeme ve výstupu předchozích procedur v této tabulce, je označený jako R square:

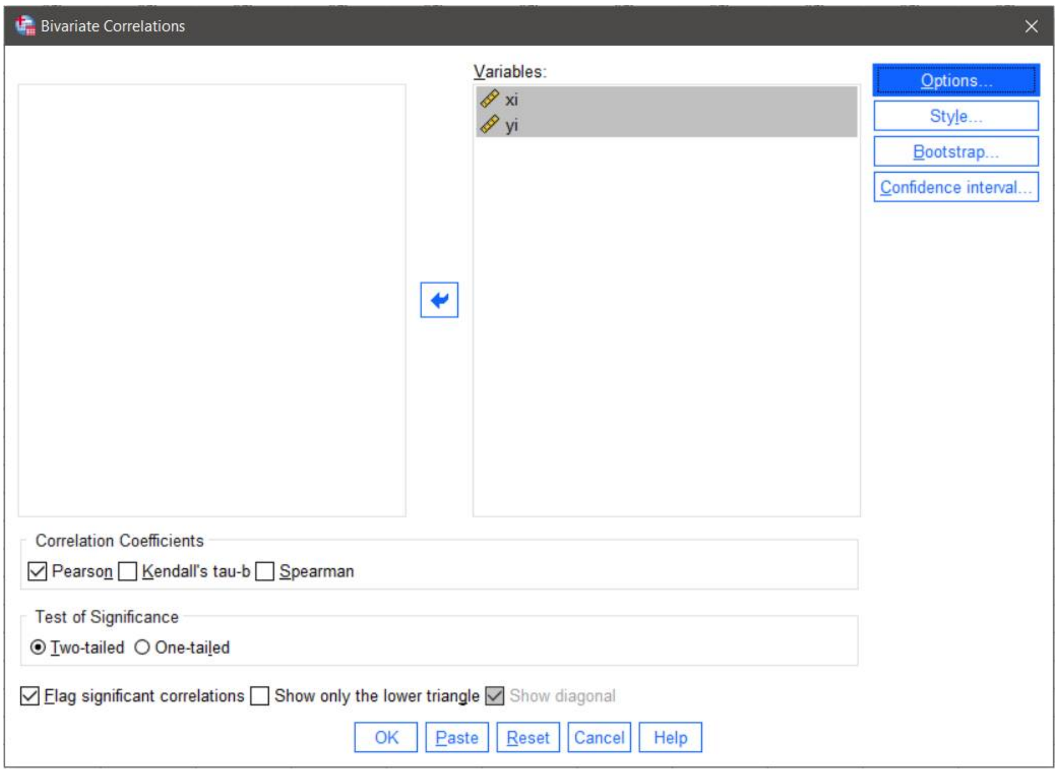
Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
,964	,929	,920	,982

The independent variable is yi.

$r_{yx}^2 = 0{,}926 \rightarrow$ 92,6 % z celkové variability závisle proměnné je možné vysvětlit příslušnou regresní přímkou,

Hodnotu *korelačního koeficientu* a test jeho významnosti najdeme v posloupnosti procedur Analyze – Correlate – Bivariate. Vstupní panel je potřeba vyplnit takto:



Výstupem této procedury je následující korelační matice, která obsahuje všechny potřebné hodnoty:

Correlations

		xi	yi
xi	Pearson Correlation	1	-,964**
	Sig. (2-tailed)		<,001
	N	10	10
yi	Pearson Correlation	-,964**	1
	Sig. (2-tailed)	<,001	
	N	10	10

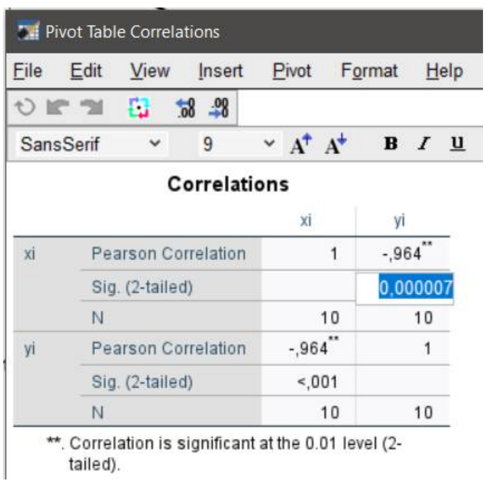
**. Correlation is significant at the 0.01 level (2-tailed).

Hodnota výběrového korelačního koeficientu r_{yx} je rovna -0,964 a najdeme ji v řádku označeném jako Pearson Correlation (zelená barva). Abychom zjistili, zda je koeficient korelace statisticky významný i v základním souboru, je potřeba provést test jeho významnosti:

1) $H_0: \rho_{yx} = 0$

   $H_1: \rho_{yx} \neq 0$

2) Sig. = 0,000007 (Sig. najdeme v tabulce výše, označena žlutě – vidíme, že je velmi malá, menší než 0,001 a pokud chceme znát její konkrétní hodnotu, na tabulku 2x klikneme a potom 2x klikneme i na hodnotu Sig.)



Correlations

		xi	yi
xi	Pearson Correlation	1	-,964**
	Sig. (2-tailed)		0,000007
	N	10	10
yi	Pearson Correlation	-,964**	1
	Sig. (2-tailed)	<,001	
	N	10	10

**. Correlation is significant at the 0.01 level (2-tailed).

3) Sig. < α → zamítáme H_0, přijímáme H_1.

Na hladině významnosti 5 % jsme prokázali, že proměnné jsou lineárně závislé (korelované). Závislost je velmi silná, nepřímá (to nám říká hodnota koeficientu korelace).

Financováno Evropskou unií NextGenerationEU

Národní plán obnovy

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY