

LEKCE09

MĚŘENÍ (SÍLY) ASOCIACE MEZI DVĚMA SPOJITÝMI PROMĚNNÝMI: KORELAČNÍ KOEFICIENTY A GRAFY

vzorový výsledek cvičení

C9.1:

- Jak velkou byste očekávali souvislost mezi důležitostí Boha (q33) a četností modliteb k Bohu mimo církevní obřady (q36)?
- Potvrdilo se vám toto očekávání? Jaký koeficient jste pro výpočet použili a proč?
- Pokud ano, o čem to svědčí?
- Co znamená, že má vypočtený korelační koeficient záporné znaménko?
- Je tato korelace signifikantní?

Řešení:

Correlations

			Q33 Bůh - důležitost v životě	Q36 Modlení k Bohu mimo obřady
Spearman's rho	Q33 Bůh - důležitost v životě	Correlation	1,000	-,784*
		Coefficient		
		Sig. (2-tailed)	,	,000
		N	1880	1857
	Q36 Modlení k Bohu mimo obřady	Correlation	-,784*	1,000
		Coefficient		
		Sig. (2-tailed)	,000	,
		N	1857	1911

** . Correlation is significant at the .01 level (2-tailed).

- Očekávali jsme vysokou korelaci a
- ona se potvrdila. Spearmana, jedna z proměnných je ordinální povahy.
- Svědčí to o tom, že respondenti jsou ve svých odpovědích konsistentní.
- Záporné znaménko znamená v daném případě kladnou korelaci: čím je Bůh důležitější, tím častěji se respondent modlí k Bohu i mimo církevní obřady (a naopak).

Poznámka: Pozor, znaménko znamená směr vztahu, ale formální směr vztahu. Ve skutečnosti záporné znaménko tohoto vztahu říká pouze: **KDYŽ ROSTE HODNOTA JEDNÉ ŠKÁLY, KLESÁ HODNOTA DRUHÉ.** Musíme vzít vždy v potaz natočení škál. U q33 důležitost boha na škále 1-10 stoupá a u q36 frekvence modlení na škále 1-7 klesá. Kdybychom změnil pořadí u jedné ze škál (např. 1 by reprezentovala nejvyšší a ne nejnižší význam boha), bylo by výsledné znaménko kladné. **PROTO JE VĚTA „čím je Bůh důležitější, tím častěji se respondent modlí k Bohu i mimo církevní obřady“ JIŽ INTERPRETACÍ ZJIŠTĚNÉHO VZTAHU.**

- Ano, korelace je signifikantní - viz sig. (2-tailed) = .000.

LEKCE 9: MĚŘENÍ (SÍLY) ASOCIACE MEZI DVĚMA SPOJITÝMI PROMĚNNÝMI: KORELAČNÍ KOEFICIENTY A GRAFY

C9.2:

- Do jaké míry souvisí postoje v baterii *q54a – q54h* se vzděláním respondenta (*ISCED1*)? Výpočet udělejte jako obdélníkovou matici, v níž vzdělání je nezávisle proměnnou. Formát výstupního výpočtu přetvořte tak, aby vzdělání bylo umístěno ve sloupci a závisle proměnné v řádcích.
- Jaký koeficient jste pro výpočet použili a proč?
- S jakou proměnnou v dané baterii vzdělání nejvíce koreluje? Jaký je příslušný koeficient determinace?
- A s jakou proměnnou naopak vzdělání koreluje nejméně?
- Proč jsou některé korelace kladné a některé záporné?

Řešení:

- Výsledkem je tato matice.

Correlations

	Spearman's rho		
	ISCED1		
	Correlation Coefficient	Sig. (2-tailed)	N
Q54A Odpovědnost za jednotlivce má:	-,192**	,000	1860
Q54B Nezaměstnaní nabízené zaměstnání:	,024	,312	1850
Q54C Konkurence je:	-,150**	,000	1848
Q54D Stát by měl:	-,170**	,000	1834
Q54E Příjmy by měly být:	,205**	,000	1851
Q54F Rozšiřovat/ zachovat by se mělo obchod. a prům.:	-,187**	,000	1823
Q54G Za důchodové zabezp. má být zodpovědný:	-,173**	,000	1851
Q54H Za bydlení má být zodpovědný:	-,135**	,000	1860

** . Correlation is significant at the .01 level (2-tailed).

- Použili jsme Spearmanův koeficient pořadové korelace. Nejde o kardinální proměnné (nelze použít Pearsonův korelační koeficient) a tabulky nejsou ani čtvercové, neboť vzdělání má jiný počet variant (nelze použít Kendallovo taub) než ostatní proměnné.
- Nejvíce je závislý na vzdělání názor o vyrovnanosti příjmů (*q54e*), Spearmanův koeficient je **0,205**. Je kladný, takže čím vyšší vzdělání (je seřazeno od nejnižšího k nejvyššímu), tím více se respondenti přiklánějí k názoru, že by se mělo více podnitit úsilí jednotlivce a naopak, čím nižší vzdělání, tím více se respondenti přiklánějí k názoru, že příjmy by měly být vyrovnanější.
- Nejméně je závislý na vzdělání názor o vyrovnanosti příjmů (*q54e*), Spearmanův koeficient je **0,024** (je to tak málo, že de facto o žádný vztah nejde, rozložení odpovědí v jednotlivých vzdělanostních kategoriích je přibližně stejné).
- Protože vztah má v konkrétních případech různý směr (pozitivní = čím vyšší vzdělání, tím více roste škálová hodnota druhé škály; negativní = čím nižší roste vzdělání, tím více roste škálová

LEKCE 9: MĚŘENÍ (SÍLY) ASOCIACE MEZI DVĚMA SPOJITÝMI PROMĚNNÝMI: KORELAČNÍ KOEFICIENTY A GRAFY

C9.3: V otázce q62 lidé posuzovali svůj postoj k různým typům politických systémů. Jak spolu tyto postoje korelovaly (zvolte dobře koeficient)? Popište základní zjištění.

Řešení:

a) Volili jsme Kendallovo tau_b, protože všechny škály mají stejný rozsah a proto všechny kontingenční tabulky budou čtvercové.

Correlations

			Silný vůdce v politickém systému	Experti rozhodují	Zemi vládně armáda	Demokratický systém
Kendall's tau_b	Silný vůdce v politickém systému	Correlation Coefficient	1,000	,163**	,264**	-,210**
		Sig. (2-tailed)	,	,000	,000	,000
		N	1758	1692	1733	1731
		Correlation Coefficient	,163**	1,000	,104**	-,033
	Experti rozhodují	Sig. (2-tailed)	,000	,	,000	,125
		N	1692	1748	1720	1721
		Correlation Coefficient	,264**	,104**	1,000	-,180**
		Sig. (2-tailed)	,000	,000	,	,000
	Zemi vládně armáda	N	1733	1720	1811	1780
		Correlation Coefficient	-,210**	-,033	-,180**	1,000
		Sig. (2-tailed)	,000	,125	,000	,
		N	1731	1721	1780	1809

** . Correlation is significant at the .01 level (2-tailed).

b) Korelace mezi proměnnými byla spíše slabší, jen:

- mezi preferencí vlády armády a preferencí silného vůdce byla středně silná asociace (0,264) - pozitivní, což znamená, že preference jednoho byla spojena s preferencí druhého)
- mezi preferencí demokracie a preferencí silného vůdce byla nalezena středně silná asociace (-0,264) - negativní, což znamená, že preference jednoho byla spojena s odmítáním druhého)

LEKCE 9: MĚŘENÍ (SÍLY) ASOCIACE MEZI DVĚMA SPOJITÝMI PROMĚNNÝMI: KORELAČNÍ KOEFICIENTY A GRAFY

C9.4: Podívejte se na souvislosti (zvolte dobře koeficient) mezi znepokojením nad životními podmínkami osob podle toho jak jsou nám blízcí či vzdálení (q79). Dovedli byste zde odhalit nějakou tendenci (napovím, že se zde rýsují přinejmenším dva, možná tři typy respondentů - víte co mám na mysli)?

Řešení:

Correlations

			Nejbližší příbuzní	Sousedé	Lidé z regionu	Obyvatelé ČR	Evropané	Lidstvo
Kendall's tau_b	Nejbližší příbuzní	Coefficient	1,000	,599**	,293**	,167**	,169**	,018
		Sig. (2-tailed)	.	,000	,000	,000	,000	,359
		N	1937	1904	1895	1894	1850	1859
	Sousedé	Coefficient	,599**	1,000	,544**	,289**	,270**	,114**
		Sig. (2-tailed)	,000	.	,000	,000	,000	,000
		N	,599	1908	1887	1877	1834	1838
	Lidé z regionu	Coefficient	,293**	,544**	1,000	,516**	,365**	,265**
		Sig. (2-tailed)	,000	,000	.	,000	,000	,000
		N	1895	1887	1900	1879	1836	1839
	Obyvatelé ČR	Coefficient	,167**	,289**	,516**	1,000	,512**	,442**
		Sig. (2-tailed)	,000	,000	,000	.	,000	,000
		N	1894	1877	1879	1901	1840	1844
	Evropané	Coefficient	,169**	,270**	,365**	,512**	1,000	,496**
		Sig. (2-tailed)	,000	,000	,000	,000	.	,000
		N	1850	1834	1836	1840	1853	1831
	Lidstvo	Coefficient	,018	,114**	,265**	,442**	,496**	1,000
		Sig. (2-tailed)	,359	,000	,000	,000	,000	.
		N	1859	1838	1839	1844	1831	1865

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

Zdá se, že se vytvářejí dva typy respondentů.

- Jedni jsou orientováni, co se týče obav o životní podmínky ostatních, spíše na své nejbližší - příbuzné a sousedy (žlutá barva).
- Druzí jsou orientováni regionálně (zelená barva).
- Třetí jsou orientováni globálně (modrá barva).

Poznámka:

Obecně je výklad následující: Čím častěji je člověk znepokojen osudem svých příbuzných, tím častěji je současně znepokojen i osudem svých sousedů (0,599). Slabší je vazba mezi znepokojením s osudem příbuzných a s osudem lidí z regionu (0,293), ještě slabší co se týče znepokojení osudem obyvatel ČR (0,167) a Evropy (0,169) a téměř žádný vztah nenacházíme mezi znepokojením s osudem lidstva (0,018, vztah se blíží 0 a tak by se dalo takřka říci, že z těch, kdo jsou znepokojeni osudem příbuzných je osudem lidstva polovina znepokojena a polovina ne).

LEKCE 9: MĚŘENÍ (SÍLY) ASOCIACE MEZI DVĚMA SPOJITÝMI PROMĚNNÝMI: KORELAČNÍ KOEFICIENTY A GRAFY

C9.5: V minulém cvičení jste testovali nulovou hypotézu, že mezi hrdostí na to být občanem ČR (q71) a preferencí místa, k němuž respondenti pocítují největší příslušnost (q67) není žádná souvislost. Pokuste se nyní o vztahu rozhodnout na základě příslušného koeficientu asociace. Jaký koeficient zvolíte a proč? O co je nyní vaše informace bohatší?

Řešení:

a) Pokud budeme vztah považovat za symetrický (jde o ordinální proměnné), je možné zvolit např. Kendalovo tau_c (tabulka je 5x4) nebo Spearmana.

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal	Kendall's tau-c	,015	,018	,829	,407
	Spearman Correlation	,020	,024	,860	,390 ^c
Interval by Interval	Pearson's R	,052	,025	2,208	,027 ^c
N of Valid Cases		1832			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

c. Based on normal approximation.

Korelace je ale nízká (v případě Kendalova tau 0,015 a v případě Spearmana je 0,020), navíc nesignifikantní (v případě Kendalova tau je sign. 0,407 a v případě Spearmana je 0,390). Nízkou korelaci dokládá i rozložení dat v tabulce, viz první řádek dat v tabulce příslušného cvičení k lekci 8 (ať je člověk hrdý více nebo méně, pocítuje nejsilnější příslušnost k obci, v níž žije). Oproti informaci vyvozené z rozložení či chí-kvadrátu nám dává koeficient asociace informaci i o (velmi malé) síle asociace.

Poznámka:

Signifikance varují, že odmítneme-li nulovou hypotézu (o neexistenci vztahu v populaci), činíme to s vysokým rizikem chyby. V případě Kendalova tau 41% riziko, v případě Spearmanova koeficientu 39% riziko.

b) Pokud bychom považovali vztah za asymetrický, museli bychom použít např. Somersovo D:

Directional Measures

			Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal	Somers' d	Symmetric	,018	,021	,829	,407
		Nejsilnější pocit příslušnosti Dependent	,019	,023	,829	,407
		Hrdost na občanství ČR Dependent	,017	,020	,829	,407

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Ani zde však nenacházíme téměř žádnou asociaci, ať již vybereme jako nezávislou jednu či druhou proměnnou. A i zde signifikance varují, že odmítneme-li nulovou hypotézu (o neexistenci vztahu v populaci), činíme to s vysokým rizikem chyby..