LEKCE09
MĚŘENÍ (SÍLY) ASOCIACE MEZI DVĚMA SPOJITÝMI PROMĚNNÝMI: KORELAČNÍ KOEFICIENTY A GRAFY

SÍLA ASOCIACE

- Je dána rozdíly mezi jednotlivými variantami proměnné.
- Měříme ji speciálními statistickými koeficienty asociace.
- Jednotlivé úrovně měření (nominální, ordinální, kardinální) a některé typy znaků (znaménkové znaky ap.) mají své specifické koeficienty.
- Použití koeficientů je ovlivněno i velikostí kontingenční tabulky (své koeficienty mají čtyřpolní tabulky vyjadřující vztah dvou dichotomických znaků).

K měření síly statistické závislosti MEZI PROMĚNNÝMI (vazby, souvislosti, asociace) jsou určeny sumarizační statistiky nazývané ASOCIAČNÍ ČI KORELAČNÍ KOEFICIENTY (koeficienty, závislosti, explanační síly ap.). Tyto koeficienty podávají souhrnnou informaci o existenci vztahů mezi proměnnými a o jejich síle.

KOEFICIENTY ASOCIACE

K jejich základním charakteristikám KOEFICIENTŮ ASOCIACE patří, že:

- Hodnoty koeficientů se většinou pohybují v intervalech:
  - <0;1>

  Příklad: Mezi volbou politické strany a subjektivní třídou existuje silný vztah. Koeficient vypovídá o síle tohoto vztahu, nikoliv o jeho směru (u nominálního znaku jako je politická strana nemá směr smysl).
  - <-1;+1>

  Příklad: „S růstem vzdělání roste výše platu (pozitivní vztah) nebo naopak (negativní vztah). Koeficient vypovídá o síle tohoto vztahu, znaménko o jeho směru.
- Čím vyšší je hodnota koeficientů (v absolutní hodnotě), tím silnější je vztah.
- Znaménko určuje směr vztahu (koeficienty pro ordinální a kardinální proměnné).

  Záporné koeficienty znamenají negativní asociaci a kladné koeficienty pozitivní asociaci. Znaménko neříká nic o síle vztahu (o té vypovídá absolutní hodnota koeficientu).

  - Nula má obvykle význam neexistence vztahu (někdy ovšem, jak jsme již viděli, je však jen výrazem toho, že vztah sice existuje, ale je nelineární).
  - Hodnota 1,00 má význam existence perfektního vztahu.

Příklad:

ŽÁDNÁ ASOCIACE		STŘEDNÍ ASOCIACE		PERFEKTNÍ ASOCIACE	
65%	65%	30%	75%	0%	100%
35%	35%	70%	25%	100%	0%
korelace 0,000		korelace 0,500		korelace 1,000	

PAMATUJME SI!

- Pro každou úroveň měření (nominální, ordinální a kardinální) jsou určeny zvláštní koeficienty.
  - Máme-li proměnné různého charakteru (například nominální a kardinální), musíme volit vždy koeficient pro proměnnou nižší úrovně (v tomto případě nominální).
  - V některých případech jsou pro takový případ vyvinuty speciální koeficienty (pro zmíněný případ je to například koeficient eta).
- Některé z koeficientů lze použít jen při lineárním vztahu, jiné i pro vztahy nelineární.
- Některé koeficienty rozlišují, která z proměnných je závisle a která nezávisle proměnná (asymetrické), jiné to nerozlišují (symetrické).
- Některé dokonce rozlišují i velikost kontingenční tabulky.
- Prokázání asociace není důkazem kauzality vztahu.

KOEFICIENTY MÍRY ASOCIACE PRO DVĚ RŮZNÉ ÚROVNĚ MĚŘENÍ

Jestliže jedna proměnná je nominální a druhá ordinální nebo kardinální, nebo je-li jedna proměnná ordinální a druhá kardinální, existují celkem tři možnosti jak vybrat pro měření síly asociace mezi nimi vhodný koeficient.

- Použijeme KOEFICIENTU PRO NIŽŠÍ ÚROVEŇ MĚŘENÍ (proměnnou vyšší úrovně měření lze vždy transformovat v proměnnou nižší úrovně měření, nikoliv však naopak). Musíme si být vědomi toho, že tím ztrácíme část informací.
- Jestliže jedna z proměnných má jen 2 varianty (DICHOTOMICKÉ PROMĚNNÉ), můžeme ignorovat její úroveň měření a volbu koeficientu určí druhá (nedichotomická) proměnná.
- Použijeme SPECIÁLNĚ PRO TENTO PŘÍPAD VYVINUTÝCH KOEFICIENTŮ. Příkladem je ETA KOEFICIENT, který může být použit, když závisle proměnná je měřena na intervalové nebo dlouhé ordinální škále a nezávisle proměnná na nominální škále.

PŘEHLED KOEFICIENTŮ MÍRY ASOCIACE

ÚROVEŇ MĚŘENÍ A VELIKOST TABULEK		VHODNÁ METODA	VHODNÝ KOEFICIENT	INFERENČNÍ STATISTIKA
NOMINÁLNÍ NOMINÁLNÍ	2x2	kontingenční tabulka	Phí, Yules Q Goodmanoovo a Kruskalovo tau	chí-kvadrát
NOMINÁLNÍ NOMINÁLNÍ	3 a více x 2 a více	kontingenční tabulka	Lambda, Goodmanoovo a Kruskalovo tau, Cramerovo V	chí-kvadrát
NOMINÁLNÍ ORDINÁLNÍ	nominální s 3 a více	kontingenční tabulka	Theta, Goodmanoovo a Kruskalovo tau, Cramerovo V	Mann-Whitney U test (dichotomická nominální nezávislá); K-sample median test Kruskal-Wallis
NOMINÁLNÍ KARDINÁLNÍ	nominální nezávislá	kontingenční tabulka* porovnání průměrů*	Eta (korelační poměr); lze i Goodmanovo a Kruskallovo tau, Cramerovo V, Eta	F-test (ONEWAY) chí-kvadrát F-test (ONEWAY)
ORDINÁLNÍ ORDINÁLNÍ	obě s málo variantami	kontingenční tabulka	Gamma, Kendallovo tau_b** , Kendallovo tau_c	test významnosti pro Gamma test významnosti pro tau
ORDINÁLNÍ ORDINÁLNÍ	jedna s mnoha variantami	pořadová korelace	Kendallovo tau	test významnosti pro tau
ORDINÁLNÍ ORDINÁLNÍ	obě s mnoha variantami	pořadová korelace	Kendallovo tau Spearmanovo ró	testy významnosti pro tau a pro ró
ORDINÁLNÍ KARDINÁLNÍ	obě s málo variantami	kontingenční tabulka***	Eta, Gamma Kendallovo tau	F-test
ORDINÁLNÍ KARDINÁLNÍ	ordinální s málo variantami	porovnání průměrů pořadová korelace	Eta Kendallovo tau	F-test test významnosti pro tau
ORDINÁLNÍ KARDINÁLNÍ	obě s mnoha variantami	pořadová korelace	Kendallovo tau, Spearmanovo ró	testy významnosti pro tau a pro ró
KARDINÁLNÍ KARDINÁLNÍ	obě s málo variantami	kontingenční tabulka	Pearsonovo R	test významnosti pro R
KARDINÁLNÍ KARDINÁLNÍ	nejméně jedna s mnoha variantami	scattegram regrese	Pearsonovo R, regresení koeficienty	

*podle počtu variant kardinální proměnné ** tau_b pro čtvercovou tabulku
*** pokud je kardinální závislá
Vaus, D. A. de: Surveys in Social Research. Unwin Hyman, London 1990, p.182.

KORELAČNÍ MATICE