

**Potenciál
pozitívnej flexibilizácie práce
a realita jej skutočného uplatňovania v SR**

**Analýza
odpovedí respondentov
pomocou kontingenčných tabuliek**

Samuel Koróny

Ekonomická fakulta
Univerzita Mateja Bela v Banskej Bystrici

Použité metódy 1

- Na analýzu odpovedí respondentov sme použili kontingenčné tabuľky (angl. cross tables).
- Sú základným nástrojom na testovanie vzťahu medzi dvomi kategorickými alebo kvalitatívnymi znakmi (premennými).
- Už niekoľko rokov sú aj súčasťou Microsoft Excel.
- My sme použili štatistický systém SPSS verzia 19 pre všetky tabuľky, testy a grafy.

Použité metódy 2

- V kontingenčných tabuľkách je testovaná vzájomná závislosť (súvis, asociácia) medzi dvojicami znakov - otázok.
- Pri teste sa v políčkach tabuľky porovnávajú skutočné absolútne počty s očakávanými absolútnymi počtami, ktoré by sme mali v prípade nezávislosti otázok.
- Nulová hypotéza: dané dva znaky spolu nesúvisia.
- Alternatívna hypotéza: dané dva znaky spolu súvisia.

Použité metódy 3

- Ak je $p > 0,05$, potom prijímame nulovú hypotézu – „dané dva znaky spolu nesúvisia.“
- Ak je $p \leq 0,05$, potom zamietame nulovú hypotézu a dané dva znaky spolu súvisia.
- Zamietnutie nulovej hypotézy nehovorí nič konkrétne o spôsobe závislosti alebo o príčine významnosti testu.
- Pre častý výskyt malých alebo nulových početností v políčkach kontingenčných tabuliek sme použili presný Fisherov test, nie klasický asymptotický chí kvadrát test.

Použité metódy 4

Na zistenie príčiny významnosti vzťahu dvoch kategorických znakov použijeme tzv. „Adjusted Residual“ (AR).

$$AR_{ij} = \frac{f_{ij} - E_{ij}}{\sqrt{E_{ij} \left(1 - \frac{r_i}{W}\right) \left(1 - \frac{c_j}{W}\right)}}.$$

f_{ij} je skutočná početnosť v i-tom riadku a j-tom stĺpci tabuľky

$c_j = \sum_{i=1}^R f_{ij}$ je súčet za j-ty stĺpec, $r_i = \sum_{j=1}^C f_{ij}$ je súčet za i-ty riadok,

$W = \sum_{i=1}^R r_i = \sum_{j=1}^C c_j$ je celkový súčet tabuľky, $E_{ij} = \frac{r_i c_j}{W}$ je očakávaná početnosť.

Použité metódy 5

- Ak $AR < 0$, potom skutočná početnosť je menšia ako očakávaná (v políčku je menší počet, ako by mal byť pri nezávislosti znakov).
- Ak $AR > 0$, potom skutočná početnosť je väčšia ako očakávaná (v políčku je väčší počet, ako by mal byť).
- Ak $|AR| > 2$, potom rozdiel medzi skutočnou a očakávanou početnosťou je signifikantný na hladine $p < 0,05$.
- Ak $|AR| > 2,6$ potom rozdiel medzi skutočnou a očakávanou početnosťou je signifikantný na hladine $p < 0,01$.
- Ak $|AR| > 3,3$ tak rozdiel medzi skutočnou a očakávanou početnosťou je signifikantný na hladine $p < 0,001$.

Príklad: vzťah medzi vzdelaním respondenta a jeho pracovným časom.

- Nulová hypotéza:
pracovný čas respondenta nezávisí od jeho vzdelania.
- Alternatívna hypotéza:
pracovný čas respondenta závisí od jeho vzdelania.
- Ak je Fisherov test signifikantný ($p < 0,05$), potom hľadáme políčko alebo políčka tabuľky, kde je porušený predpoklad nezávislosti dvoch testovaných odpovedí.

			Pracovný čas				Total
			pevný	pružný počas dňa	pružný počas mesiaca	pružný počas týždňa	
V z d e l a n i e	základné	Count	4	1	0	1	6
		% within Vzdelanie	66,7%	16,7%	0,0%	16,7%	100,0%
		AR	0	0,2	-0,9	0,8	
	učňovské	Count	20	1	2	3	26
		% within Vzdelanie	76,9%	3,8%	7,7%	11,5%	100,0%
		AR	1,2	-1,5	-0,8	0,8	
	stredné	Count	128	15	24	14	181
		% within Vzdelanie	70,7%	8,3%	13,3%	7,7%	100,0%
		AR	1,8	-2,9	0,4	0,1	
	vŠ	Count	116	39	25	13	193
		% within Vzdelanie	60,1%	20,2%	13,0%	6,7%	100,0%
		AR	-2,4	3,6	0,2	-0,6	
Total		Count	268	56	51	31	406
		% within Vzdelanie	66,0%	13,8%	12,6%	7,6%	100,0%

Interpretácia

- Primárnou príčinou významnosti testu je rezíduum s najvyššou absolútnou hodnotou ($AR = 3,6$) vo štvrtom riadku a druhom stĺpci tabuľky. Je kladné, skutočný počet je teda signifikantne vyšší ako očakávaný. Respondenti s vysokoškolským vzdelaním majú vo zvýšenej miere pružný pracovný čas v priebehu dňa.
- Sekundárnou príčinou významnosti testu sú respondenti so stredným vzdelaním. Ich počet s pružným pracovným časom v priebehu dňa je naopak nižší ($AR = -2,9$). Podobne je menej respondentov s vysokoškolským vzdelaním, ktorí majú pevný pracovný čas ($AR = -2,4$). V ostatných políčkach tabuľky sú rozdiely medzi skutočnými a očakávanými početnosťami nevýznamné.

Literatúra

IBM SPSS STATISTICS BASE. 2010. manuál

Ďakujeme
za pozornosť.