Neparametrické testy

- parametrické a neparametrické testy
- pořadové neparametrické testy
- test Chí-kvadrát
  - test nezávislosti proměnných
  - test dobré shody

Parametrické testy

- t-testy a analýza rozptylu jsou tzv. parametrické testy
- parametr = charakteristika populace (průměr, rozptyl)
- parametrické testy používají při výpočtech charakteristiky populace (parametry)

Parametrické testy

- parametrické testy pracují s předpoklady o charakteristikách populace
- např. u t-testu předpokládáme, že směrodatné odchylky výběrů mohou posloužit jako odhad pro směrodatnou odchylku populace
- podobně počítají s normálním rozdělením měřeného znaku

Parametrické testy

- pokud nejsou tyto předpoklady splněny, můžeme dojít k nepřesným výsledkům

Neparametrické testy

- neparametrické testy nezávisí na charakteristikách populace ani o nich nečiní žádné závěry
- není vyžadováno normální rozdělení znaku
- proto jsou tyto testy označovány také jako „distribution-free“ testy

Neparametrické testy

- proč potom vůbec používat parametrické testy?
  - mnoho parametrických testů je poměrně „odolných“ (tzv. robustních) vůči narušení předpokladů testu (např. menší odchylky od normálního rozdělení výsledky nezkreslí)
  - parametrické testy mají větší statistickou sílu než neparametrické (větší pravděpodobnost zjištění rozdílu, pokud skutečně existuje)
  - pro některé typy analýz neparametrické metody nejsou (např. neexistuje obecně přijímaná neparametrická faktoriální ANOVA)

Neparametrické testy

- hlavní **výhody** neparametrických testů
  - nejsou omezeny předpokladem normálního rozdělení
  - jsou často založeny na pořadí, dají se použít i pro ordinální data (kde můžeme spočítat pouze medián, nikoli průměr) i pro nominální (test Chí-kvadrát)
  - nejsou citlivé na extrémní hodnoty (jsou většinou založeny na mediánu)

Neparametrické testy

- hlavní **nevýhody** neparametrických testů
  - menší statistická síla
  - pro složitější analýzy často není neparametrická varianta metody k dispozici

Neparametrické testy

- přehled neparametrických ekvivalentů parametrických testů
  - t-test pro nezávislé výběry – Mann-Whitney U test
  - t-test pro závislé výběry – Wilcoxon test
  - analýza rozptylu – Kruskall-Wallis test
  - opakovaná měření (ANOVA) – Friedman Rank Test

Mann-Whitney U test - příklad

- chceme zjistit, zda se levoruké a pravoruké osoby liší v prostorových schopnostech
- náhodně vybereme 10 leváků a 10 praváků (podobného věku, stejný počet mužů a žen) a zadáme jim test prostorových schopností

Mann-Whitney U test - příklad

- jaká bude naše hypotéza?

Mann-Whitney U test - příklad

- jaká bude naše hypotéza?
  - skóry v testu prostorových schopností se **liší** u leváků a praváků

Mann-Whitney U test - příklad

- jaká bude nulová hypotéza?

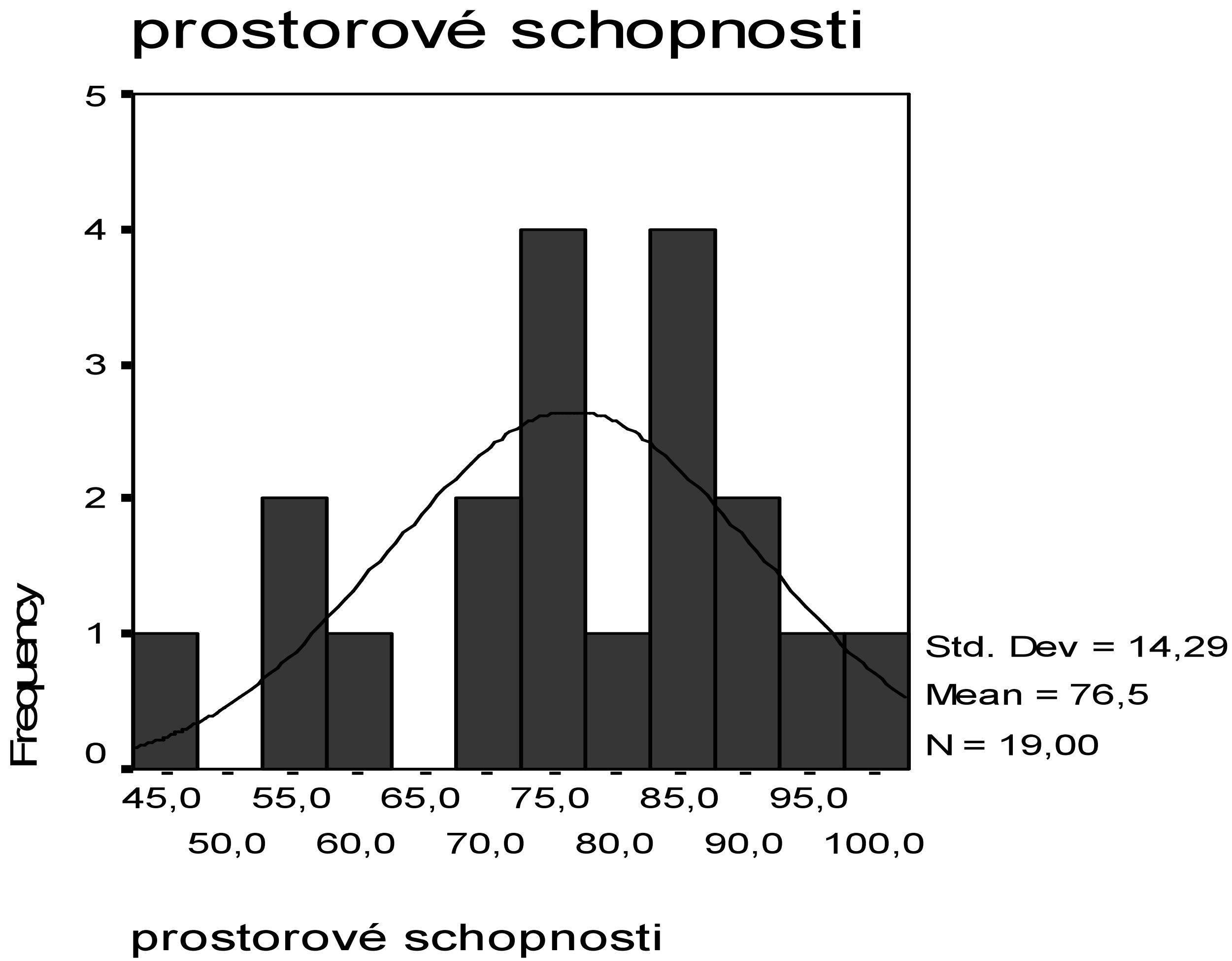
Mann-Whitney U test - příklad

- jaká bude nulová hypotéza?
  - skóry v testu prostorových schopností se u leváků a praváků **neliší**

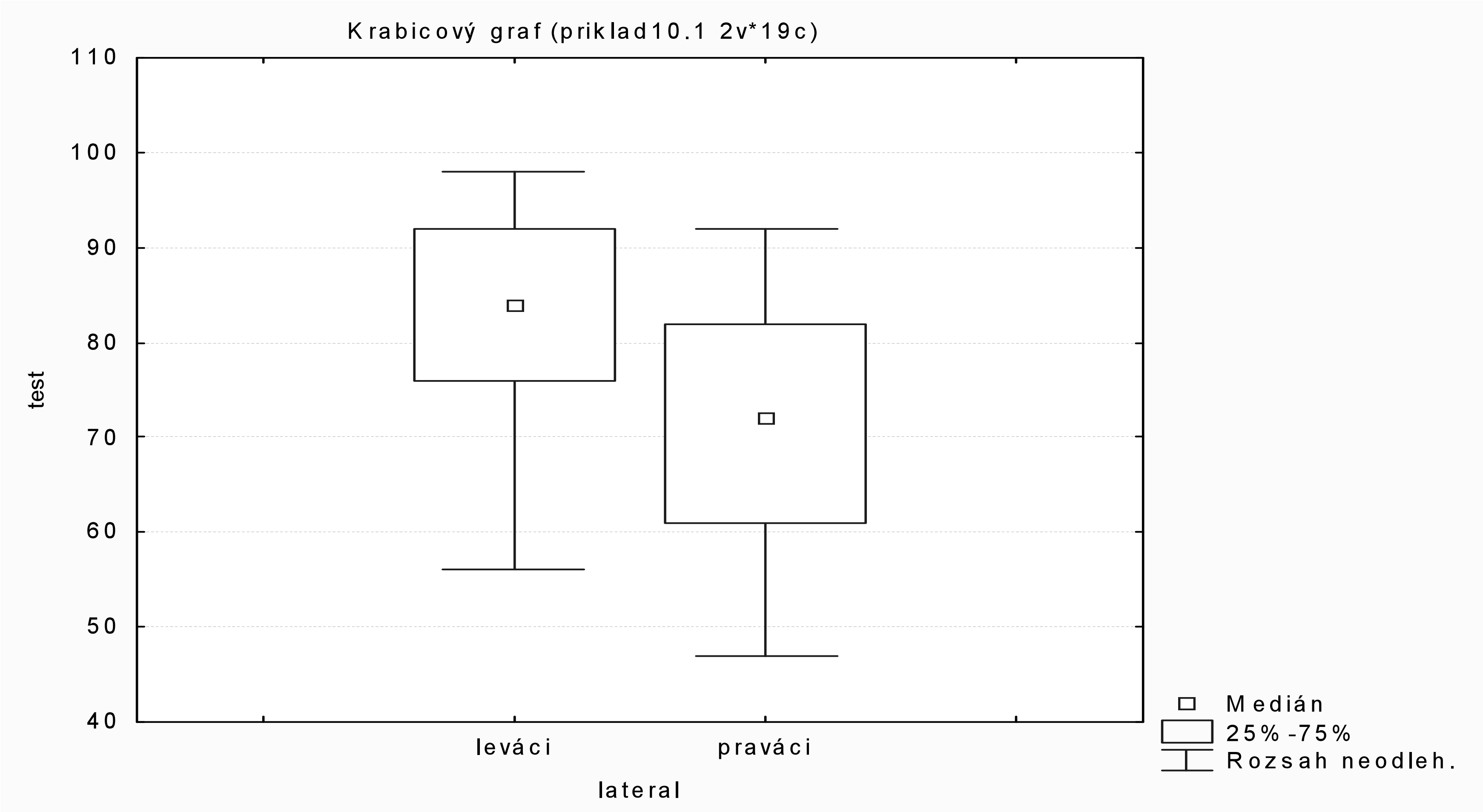
- testujeme nulovou hypotézu (začneme s předpokladem, že platí a ptáme se: jaká je pravděpodobnost pozorovaných rozdílů, pokud H0 platí?)

leváci	praváci
87	47
94	68
56	92
74	73
98	71
83	82
92	55
84	61
76	75
- (nedostavil se)	85

Mann-Whitney U test - příklad



Mann-Whitney U test - příklad



Mann-Whitney U test - příklad

- na základě takto malého vzorku nemůžeme rozhodnout, zda je rozdělení skorů z testu prostorových schopností normální
- počty osob ve skupinách jsou příliš malé (9 a 10)
- vhodnější než t-test bude proto neparametrický test

Mann-Whitney U test - příklad

- **1. krok**
  - seřadit skóry podle velikosti - bez ohledu na skupinu
  - a přidělit jim pořadí (rank)

leváci		**praváci**	
skór	**pořadí**	**skór**	**pořadí**
87	15	47	1
94	18	68	5
56	3	92	16,5
74	8	73	7
98	19	71	6
83	12	82	11
92	16,5	55	2
84	13	61	4
76	10	75	9
-	-	85	14
	$\Sigma R_1 = 114{,}5$		$\Sigma R_2 = 75{,}5$

Mann-Whitney U test - příklad

- **2. krok**
  - sečíst pořadí v obou skupinách

$$\Sigma R_1 = 114{,}5$$

$$\Sigma R_2 = 75{,}5$$

(pokud se leváci a praváci neliší, průměrné pořadí skórů by mělo být u obou skupin podobné)

Mann-Whitney U test - příklad

- **3. krok**
  - vypočítat **U** pro obě skupiny
  - podle vzorce

    $$\mathbf{U_1 = (n_1)(n_2) + n_1 (n_1+1)/2 - \Sigma R_1}$$

    $$\mathbf{U_2 = (n_1)(n_2) + n_2 (n_2+1)/2 - \Sigma R_2}$$

Mann-Whitney U test - příklad

- **výpočet U**

$$U_1 = (n_1)(n_2) + n_1\,(n_1{+}1)/2 - \Sigma\, R_1$$

$$U_1 = (9)(10) + 9\,(9{+}1)/2 - 114{,}5$$

$$\mathbf{U_1 = 20{,}5}$$

$$U_2 = (n_1)(n_2) + n_2\,(n_2{+}1)/2 - \Sigma\, R_2$$

$$U_2 = (9)(10) + 10(10{+}1)/2 - 75{,}5$$

$$\mathbf{U_2 = 69{,}5}$$

Mann-Whitney U test - příklad

- **4. krok**
  - vybrat menší z vypočítaných U
  - v našem příkladu je to U_1 (=20,5)
- **5. krok**
  - najít v tabulce kritickou hodnotu U pro zvolenou hladinu významnosti
  - pro α = .05, při $n_1 = 9$ a $n_2 = 10$
    $\boldsymbol{U_{\text{krit.}} = 20}$

Mann-Whitney U test - příklad

- **6. krok**
  - porovnat vypočítanou hodnotu U a kritickou hodnotu U
  - u tohoto testu je rozdíl **statisticky významný,** pokud je vypočítaná hodnota <u>menší než kritická</u> hodnota U
  - 20,5 není menší než 20 → **nemůžeme zamítnout nulovou hypotézu**

Mann-Whitney U test - příklad

- **závěr**: rozdíl mezi leváky a praváky v testu prostorových schopností není statisticky významný
- neznamená to nutně, že kdybychom prozkoumali celou populaci leváků a praváků, nebyl by mezi nimi rozdíl – pouze se nám tento rozdíl nepodařilo prokázat (hlavně díky malému N)

Test Chí-kvadrát

- chí-kvadrát může být použit
  - pro testování rozdělení jedné proměnné (test dobré shody)
  - testování nezávislosti dvou proměnných

Test Chí-kvadrát

- chí-kvadrát pro testování nezávislosti proměnných se používá pro nominální nebo ordinální proměnné
- data jsou uspořádána do tzv. kontingenční tabulky (viz příklad)

Příklad

- zajímá nás, jak souvisí model manželství s jeho vydařeností
  - model manželství má kategorie: dominance žena, dominance muž, kooperace
  - vydařenost má 3 kategorie – vydařené, průměrné, nevydařené
- pozn.: jde o manželství rodičů respondentů, tak jak je posuzují oni (zdroj dat – výzkum doc. Plaňavy)

Příklad

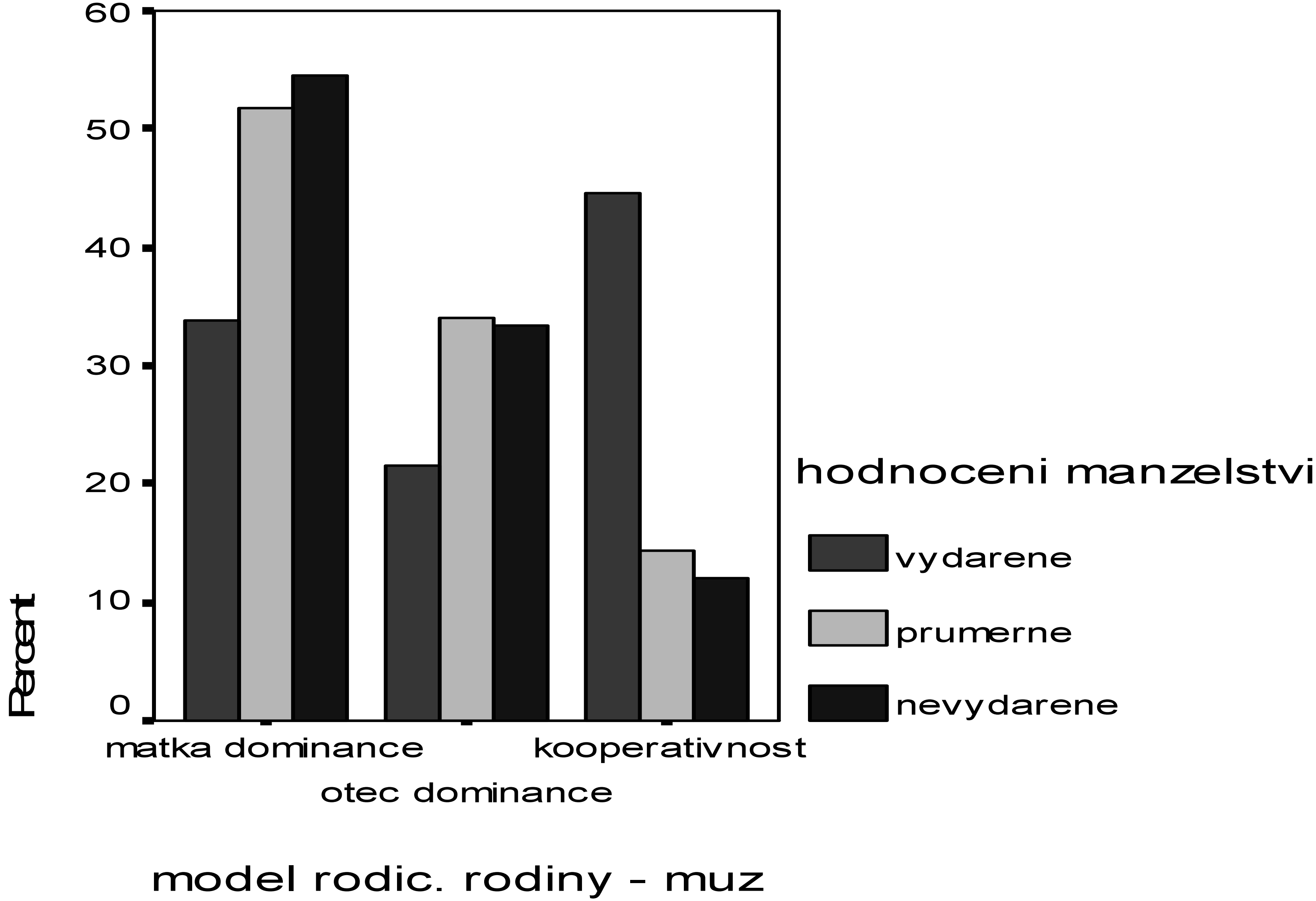
- otázka zní: liší se podíl vydařených, průměrných a nevydařených manželství u rodin, kde dominovala matka, rodin, kde dominoval otec a u rodin, kde nedominoval ani jeden z nich?

Kontingenční tabulka

model rodic. rodiny - muz * hodnoceni manzelstvi rodicu - muz Crosstabulation

Count

		hodnoceni manzelstvi rodicu - muz			Total
		vydarene	prumerne	nevydarene	
model rodic. rodiny - muz	matka dominance	22	29	18	69
	otec dominance	14	19	11	44
	kooperativnost	29	8	4	41
Total		65	56	33	154



60
50
40
30
20
10
0
Percent
matka dominance
otec dominance
kooperativnost
hodnoceni manzelstvi
vydarene
prumerne
nevydarene
model rodic. rodiny - muz

Test Chí-kvadrát

- chí-kvadrát porovnává očekávané a pozorované četnosti
- očekávané jsou četnosti za předpokladu, že proměnné jsou nezávislé

model rodic. rodiny - muz * hodnoceni manzelstvi rodicu - muz Crosstabulation

			hodnoceni manzelstvi rodicu - muz			
			vydarene	prumerne	nevydarene	Total
model rodic. rodiny - muz	matka dominance	Count	22	29	18	69
		% within model rodic. rodiny - mu	31,9%	42,0%	26,1%	100,0%
	otec dominance	Count	14	19	11	44
		% within model rodic. rodiny - mu	31,8%	43,2%	25,0%	100,0%
	kooperativnost	Count	29	8	4	41
		% within model rodic. rodiny - mu	70,7%	19,5%	9,8%	100,0%
Total		Count	65	56	33	154
		% within model rodic. rodiny - mu	42,2%	36,4%	21,4%	100,0%

Příklad

- v našem příkladu bylo 42,2% vydařených manželství
- pokud by proměnné (model a vydařenost manželství) byly vzájemně nezávislé, poměr vydařených manželství v jednotlivých modelech manželství by měl být přibližně stejný (a odrážet celkový podíl) – 42%
- podobně ostatní kategorie…

Test Chí-kvadrát

- očekávané četnosti – výpočet:

$$O_{ij} = (ř_i \; s_j)/ N$$

(pro každé políčko tabulky se vynásobí celkové četnosti z příslušného řádku se sloupcovými četnostmi a vydělí celkovým počtem osob)

Příklad

rodic. rodiny - muz * hodnoceni manzelstvi rodicu - muz Crosstab

Count

		dnoceni manzelstvi rodicu - mu			Total
		vydarene	prumerne	nevydarene	
model rodic rodiny - mu:	matka dominar	22	29	18	69
	otec dominance	14	19	11	44
	kooperativnost	29	8	4	41
Total		65	56	33	154

Příklad

- pro první políčko tabulky (vydařená manželství s dominantní matkou) je očekávaná četnost

  $O_{ij} = (ř_i \ s_j)/ N$

  $O_{11} = (ř_1 \ s_1)/ N$

  $O_{11} = (69*65)/ 154$

  $\underline{O_{11} = 29{,}12}$

Očekávané četnosti

model rodic. rodiny - muz * hodnoceni manzelstvi rodicu - muz Crosstabulation

			hodnoceni manzelstvi rodicu - muz			Total
			vydarene	prumerne	nevydarene	
model rodic. rodiny - muz	matka dominance	Count	22	29	18	69
		Expected Count	29,1	25,1	14,8	69,0
	otec dominance	Count	14	19	11	44
		Expected Count	18,6	16,0	9,4	44,0
	kooperativnost	Count	29	8	4	41
		Expected Count	17,3	14,9	8,8	41,0
Total		Count	65	56	33	154
		Expected Count	65,0	56,0	33,0	154,0

Test Chí-kvadrát

- chí-kvadrát porovná očekávané četnosti s pozorovanými

$$\chi^2 = \Sigma\ [(\text{pozor. četnosti} - \text{oček.})^2/\text{oček.}]$$

Příklad

$\chi^2 = \Sigma$ [(pozor. četnosti – oček.)^2/oček.]

χ^2 = (-7,1)^2/29,1 +3,9^2/25,1 + 3,2^2/14,8 + (-4.6)^2/18,6 + 3^2/16 + 1,6^2/9,4 + 11,7^2/17,3 + (-6,9)^2/14,9 + (-4,8)^2/8,8

χ^2 = 18, 71

Test Chí-kvadrát

- pro vyhledání kritické hodnoty χ^2 v tabulce musíme vypočítat ještě počet stupňů volnosti (df)
- **df = (ř-1) (s-1)**

(tj. počet řádků -1 krát počet sloupců -1)

Příklad

- df = (ř-1) (s-1)

  df = (3-1) * (3-1)

  <u>df = 4</u>
- v tabulkách vyhledáme kritickou hodnotu χ^2 pro df = 4 a 5% hladinu významnosti
- **χ^2_{krit} = 9,49**

Příklad

- $\chi^2_{\text{krit}} = 9{,}49$
- $\chi^2 = 18{,}71$
- **závěr:** vypočítaná hodnota je větší než kritická hodnota - <u>očekávané a pozorované četnosti se liší na 5% hladině významnosti</u> (tj. je malá pravděpodobnost, že proměnné jsou nezávislé)

Test Chí-kvadrát ve Statistice

- Pearsonův chí-kv. : 18,7117, sv=4, p=,000896

Chí-kvadrát pro 1 proměnnou

- tzv. test dobré shody (goodness-of-fit test)
- opět porovnává očekávané a pozorované četnosti
- předpokladem očekávaných četností není tentokrát nezávislost proměnných (máme jen 1)

Test dobré shody

- jak určíme očekávané četnosti?
- 2 způsoby:
  - předpoklad vyplývá z teorie (např. u genetických dat – poměr osob s projevem dominantní a recesivní alely)
  - nebo můžeme předpokládat náhodné rozdělení do kategorií

Příklad

- je počet sebevražd stejný každý den v týdnu?
- zjistíme data pro rok 2000 (ČR)

Příklad

pondělí	255
úterý	247
středa	240
čtvrtek	206
pátek	236
sobota	192
neděle	226

Příklad

- **očekávané četnosti**
  - stejný počet sebevražd pro každý den v týdnu
  - celkem 1602 sebevražd
  - očekávaná četnost pro každý den je 228,9

Příklad

Pozorované vs. očekávané četnosti (Tabulka5)
Chí-kvadrát =13,44444 sv =6 p < ,036499

	pozorov.	ocekav.	P - O	(P-O)^2/O
C: 1	255,000	228,857	26,1429	2,98636
C: 2	247,000	228,857	18,1429	1,43829
C: 3	240,000	228,857	11,1429	0,54254
C: 4	206,000	228,857	-22,8571	2,28286
C: 5	236,000	228,857	7,1429	0,22294
C: 6	192,000	228,857	-36,8571	5,93579
C: 7	226,000	228,857	-2,8571	0,03567
Sčt	1602,000	1602,000	-0,0000	13,44444

Příklad

- vzorec pro výpočet je stejný
- $\chi^2 = 13,44$
- df = k -1 (počet kategorií -1)
- df = 6
- pro df =6 a 5% hladinu významnosti je $\chi^2_{krit} = 12,59$
- **rozdíl je statisticky významný**

Příklad

Pozorované vs. očekávané četnosti (Tabulka5) Chí-kvadrát =13,44444 sv =6 p < ,036499				
	pozorov.	**ocekav.**	**P - O**	**(P-O)^2/O**
C: 1	255,000	228,857	26,1429	2,98636
C: 2	247,000	228,857	18,1429	1,43829
C: 3	240,000	228,857	11,1429	0,54254
C: 4	206,000	228,857	-22,8571	2,28286
C: 5	236,000	228,857	7,1429	0,22294
C: 6	192,000	228,857	-36,8571	5,93579
C: 7	226,000	228,857	-2,8571	0,03567
Sčt	1602,000	1602,000	-0,0000	13,44444

Omezení Chí-kvadrátu

- 2 potenciální problémy:
  - malý počet osob – pokud má velké % políček tabulky očekávanou četnost menší než 5 (v ideálním případě by všechna měla mít oček. četnost nejméně 5 osob)
  - příliš velký počet osob – čím vyšší N, tím vyšší χ^2 (vyjdou významné i malé rozdíly)

Kontrolní otázky

- hlavní rozdíl mezi parametrickými a neparametrickými testy
- výhody a nevýhody neparametrických testů
- kdy je možno využít chí–kvadrát jako test nezávislosti proměnných? (pro jaké typy proměnných?)
- kdy se chí–kvadrát využívá jako test dobré shody?