VIII. Kontingenční tabulky

Test dobré shody
Test nezávislosti
Test homogenity

Vytvořil Institut biostatistiky a analýz, Masarykova univerzita
J. Jarkovský, L. Dušek

Opakovanie

Jak vznikají informace ?
– různé typy dat znamenají různou informaci

<u>Data poměrová</u> — Kolikrát ?

<u>Data intervalová</u> — O kolik ?

Spojitá data

<u>Data ordinální</u> — Větší, menší ? — Kategoriální otázky

<u>Data nominální</u> — Rovná se ? — Otázky „Ano/Ne“

Diskrétní data

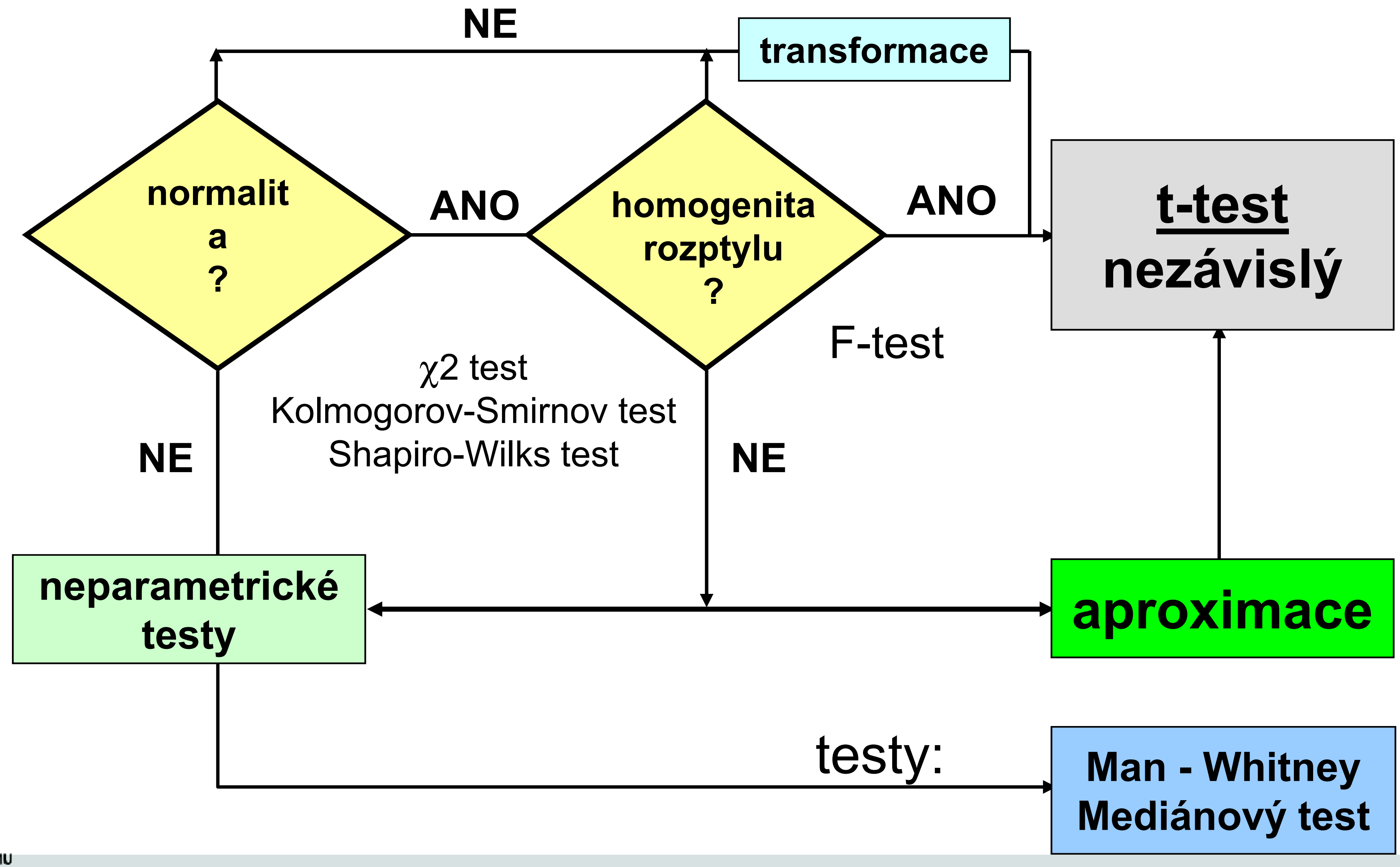
Podíl hodnot větší/menší než specifikovaná hodnota ?

Procenta odvozené hodnoty

Samotná znalost typu dat ale na dosažení informace nestačí ....

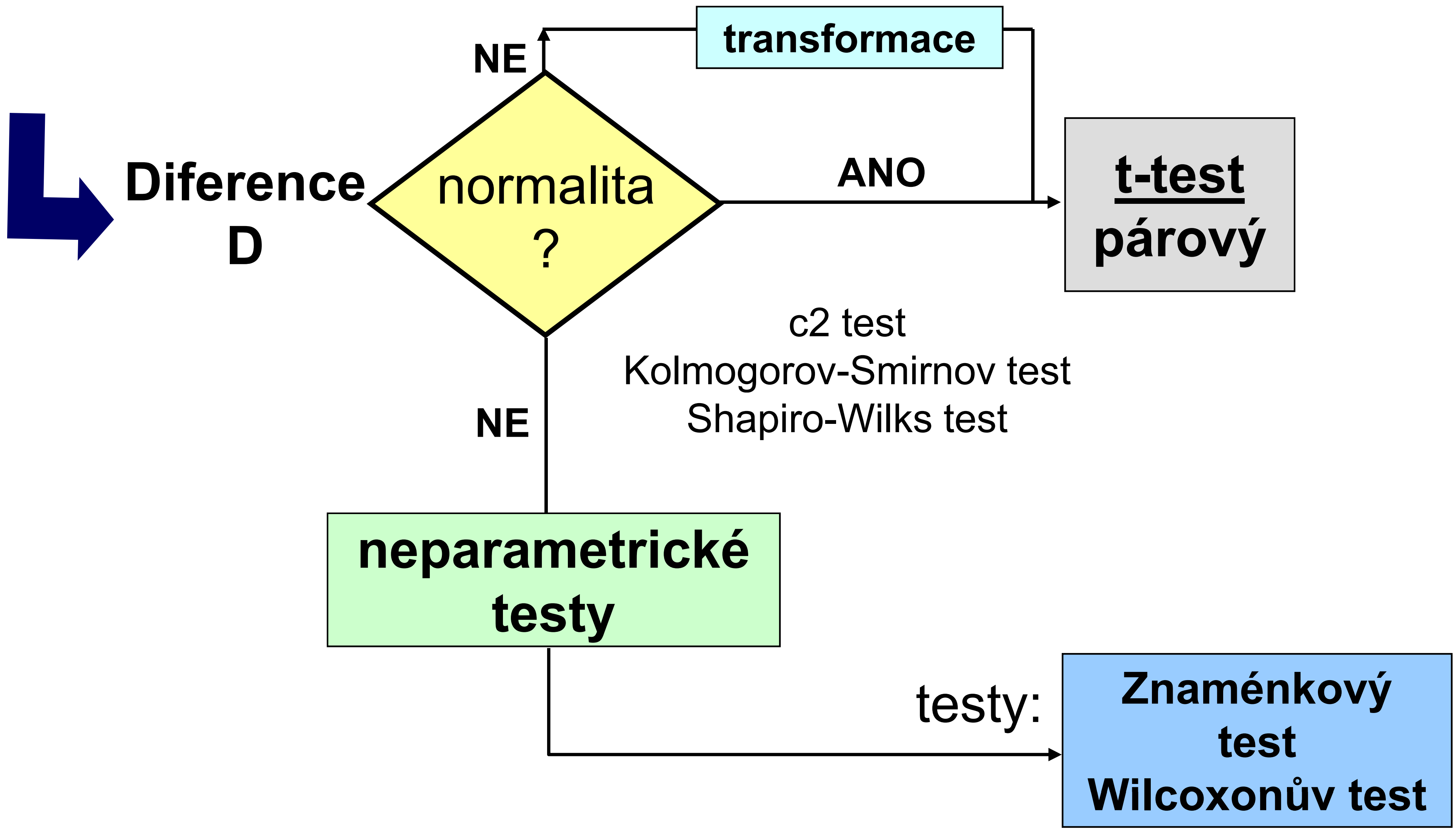
Dvouvýběrové testy: schéma analýzy

Nezávislé uspořádání



Dvouvýběrové testy: schéma analýzy

Párové uspořádání



Opakovanie

- Jednovýberový t-test:
  - Je priemerná hodnota parametru (prem 4), popisujúceho predispoziciu na istú chorobu rovná 100?
- Dojvýberový t-test nepárový:
  - Je priemerná hodnota tohto parametru u mužov rovnaká ako u žien? Alebo: Má jedno z pohlaví vyššiu predispozíciu na danú chorobu ako druhé?
  - Je rýchlosť sedimentácie krvných teliesok pred podaním lieku (prem before) u mužov rovnaká ako u žien?
- Dvojvýberový t-test párový:
  - Premenné marker pred a marker po udávajú hodnotu pred liečbou a po liečbe. Ovplyvňuje liečba hodnotu markeru?
  - Hodnota before popisuje rýchlosť sedimentácie krvných teliesok pred podaním a after rýchlosť po podaní lieku. Ovplyvňuje liek rýchlosť sedimentácie?

Diskrétne dáta a kontingenčná tabuľka

Vytvořil Institut biostatistiky a analýz, Masarykova univerzita
J. Jarkovský, L. Dušek

Motivácia

- Pri spracovaní dát sa často stretávame s úlohou zistiť, či dva znaky nominálneho alebo ordinárneho typu sú stochasticky nezávislé.

- Príklad:
  - Závisí farba vlasov na farbe očí
  - Je poradie žiakov v dvoch rôznych predmetoch nezávislé
  - Závisia známky z matematiky na známkach z biológie

Kontingenčná tabuľka

- Majme dva znaky X s počtom variant r a Y s počtom variant s

	y	**$y_{[1]}$**	**…**	**$y_{[s]}$**	**$n_{j.}$**
x	n_{jk}				
$x_{[1]}$		n_{11}	…	n_{1s}	$n_{1.}$
…		…	…	…	…
$x_{[r]}$		n_{r1}	…	n_{rs}	$n_{r.}$
$n_{.k}$		$n_{.1}$	…	$n_{.s}$	n

- $n_{j.}$ – marginálna absolútna četnosť
- n_{jk} – simultánna absolútna četnosť
- n_{jk}/n – pravdepodobnosti (mariginálne alebo simultánne)

Anotace

- Analýza kontingenčních tabulek umožňuje analyzovat vazbu mezi dvěma kategoriálními proměnnými. Základním způsobem testování je tzv. chi-square test, který srovnává pozorované četnosti kombinací kategorií oproti očekávaným četnostem, které vychází z teoretické situace, kdy je vztah mezi proměnnými náhodný.
- Test dobré shody je využíván také pro srovnání pozorovaných četností proti očekávaným četnostem daným určitým pravidlem (typickým příkladem je Hardy-Weinbergova rovnováha v genetice)
- Specifickým typem výstupů odvozených z kontingenčních tabulek jsou tzv. odds ratia a relativní rizika, využívaná často v medicíně pro identifikaci a popis rizikových skupin pacientů.

Test dobrej zhody – multinomické rozdelenie

- Môže nastať len určitý počet situácií (nejaké kategórie, z ktorých vyberáme).
- Vždy musí nastať nejaká situácia (musíme vybrať jednu možnosť).
- Nemôžu nastať dve situácie zároveň (vyberáme vždy len jednu možnosť).
- Napr. Poranenie- ľahké, stredné, ťažké.
- Jedno poranenie nemôže byť ľahké a ťažké zároveň a poranenie musí patriť do jednej kategórie.
- Chceme testovať, či teoretická pravdepodobnosť je rovnaká ako v nazbieraných dátach.

Test dobré shody - základní teorie

$$\chi^2_{(s.v.)} = \sum \frac{\left[\text{pozorovaná četnost} - \text{očekávaná četnost}\right]^2}{\text{očekávaná četnost}}$$

Porovnávame s tabuľkovou hodnotou a zamietame, ak je vyrátaná hodnota väčšia ako tabuľková

$$\chi^2_{(s.v.)} = \underbrace{\frac{\left[\text{pozorovaná četnost} - \text{očekávaná četnost}\right]^2}{\text{očekávaná četnost}}}_{\text{1. jev}} + \underbrace{\frac{\left[\text{pozorovaná četnost} - \text{očekávaná četnost}\right]^2}{\text{očekávaná četnost}}}_{\text{2. jev}} + \ldots$$

Test dobré shody: příklad I

Ověřte na datech z pokusu se 100 květinkami určitého druhu, že barva květů se geneticky štěpí v poměru žlutá : červená = 3 : 1.

H_0: Pozorovaná frekvence pro jednotlivé barvy květů jsou vzorkem populace mající poměr mezi žlutými a červenými květy 3 :1.
Součet frekvencí u obou barev květů (f_i) se rovná 100 a pozorované frekvence u kategorií barvy budou srovnány s očekávanými frekvencemi (uvedeny v závorkách):

	Kategorie barvy		
	Žlutá	Červená	n
$f_{poz.}$	84	16	100
$f_{oček.}$	75	25	

$$\chi^2 = \sum \frac{(f_{poz.} - f_{oč.})^2}{f_{oč.}} = \frac{(84-75)^2}{75} + \frac{(16-25)^2}{25} = 4{,}320$$

St. volnosti = n = k - 1 = 1

Zamítáme hypotézu shody srovnávaných četností

Při testování H_0 jsme použili matematický zápis (0,025 < P < 0,05). Z tabulek $\chi2$ rozložení vidíme, že pravděpodobnost překročení hranice 2,706 je 0,1 (10 %), což může být stručně zapsáno jako
P ($\chi2 \geq 2{,}706$) = 0,10.
Dále lze zjistit pro P ($\chi2 \geq 3{,}841$) = 0,05. V řešené úloze jsme dospěli k hodnotě testové statistiky $\chi2$ = 4,320. Pro tento případ lze tedy psát 0,025 < P ($\chi2 \geq 4{,}320$) < 0,05; a jednodušeji 0,025 < P < 0,05. Jde v podstatě o přibližné určení hranic chyby 1. druhu.

Test dobré shody: příklad II

Tento příklad je rozšířením problému z příkladu 1 na srovnání pozorovaných a očekávaných frekvencí pro více kategorií sledovaného znaku:

✓ Celkem bylo zkoumáno 250 semen určitého druhu rostliny a roztříděno do následujících kategorií: žluté/hladké; žluté/vrásčité; zelené/hladké; zelené/vrásčité. Předpokládaný poměr výskytu těchto kategorií v populaci je 9 : 3 : 3 : 1. Následující tabulka obsahuje původní data z pozorování a dále postup při testování H_0.

	žluté/hladké	žluté/vrásčité	zelené/hladké	zelené/vrásčité	n
$f_{poz.}$	152	39	53	6	250
$f_{oček.}$	140,6250	46,8750	46,8750	15,6250	

$\nu = k - 1 = 3$

$$\chi^2 = \frac{11{,}3750^2}{140{,}6250} + \frac{7{,}8750^2}{46{,}8750} + \frac{6{,}1250^2}{46{,}8750} + \frac{9{,}6250^2}{15{,}6250} = 8{,}972$$

Zamítáme hypotézu shody pozorovaných četností s očekávanými

Test dobré shody: příklad III

Složitější příklady řešené srovnáváním frekvencí je možné rozdělit na testování dílčích hypotéz:

✓ Předpokládejme, že chceme pro data z předchozí úlohy testovat hypotézu existence štěpného poměru 9 : 3 : 3 pro první tři kategorie semen:

	žluté/hladké	žluté/vrásčité	zelené/hladké	n
$f_{poz.}$	152	39	53	244
$f_{oček.}$	146,400	48,800	48,800	

n = k - 1 = 2

$$\chi^2 = \frac{5,600^2}{146,40} + \frac{9,800^2}{48,80} + \frac{4,200^2}{48,80} = 2,544$$

Nezamítáme hypotézu shody pozorovaných četností s očekávanými.

✓ Nyní otestujeme hypotézu štěpného poměru kategorií zelené/vrásčité:ostatní typy = 1:15

	zelené/vrásčité	ostatní	n
$f_{poz.}$	6	244	250
$f_{oček}$	15,625	234,375	

n = k - 1 = 1

$$\chi^2 = \frac{9,625^2}{15,625} + \frac{9,625^2}{234,375} = 6,324$$

Zamítáme hypotézu shody pozorovaných četností s očekávanými.

Test dobré shody: příklad IV - využití aditivity testu

✓ U 193 párů dvojčat byly zjištěny následující poměry pohlaví: **56 Ch - Ch**
72 Ch – H
65 H - H

? ***Za předpokladu, že narození chlapečka má stejnou pravděpodobnost jako narození holčičky, lze očekávat poměry pro výše uvedené skupiny = 0,25 : 0,5 : 0,25. Ověřte tento předpoklad na uvedeném vzorku populace.***

Σ 193 párů 1/4 : 1/2 : 1/4
očekávané četnosti = 48,25 : 96,50 : 48,25

$$\chi^2_{(2)} = 13{,}28$$

Proč lze v předchozím případě očekávat zamítnutí H_0?

Testujte následující hypotézy:
1) Jsou relativní počty párů se shodným pohlavím ve shodě s očekávanými četnostmi? (ignorujte Ch – H páry)
2) Je relativní četnost kombinace Ch - Ch a H - H párů oproti párům s rozdílným pohlavím ve shodě s očekávanými četnostmi?

Σ 121 párů 1 : 1
očekávané četnosti = 60,5 : 60,5

$$\chi^2_{(1)} = 0{,}669 \qquad \frac{H - H}{Ch - Ch}$$

Σ 193 párů 1 : 1
očekávané četnosti = 96,5 : 96,5

$$\chi^2_{(1)} = 12{,}44$$

Test dobré shody: příklad V

Města - zatížení exhalacemi - třídy (A > B > C > D)

Svět: A : B : C : D = 2 : 3 : 6 : 4

Konkrétní země (n = 184 měst): A : B : C : D = 32 : 151 : 182 : 116

H_0: shoda f_i a F_i a = 0,05

F_A: 64,13 F_C: 192,39

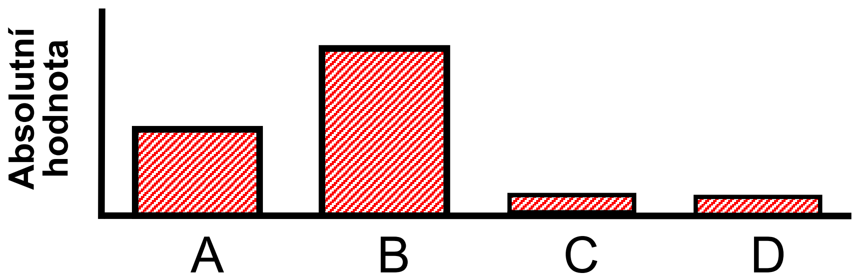
F_B: 96,19 F_D: 128,27

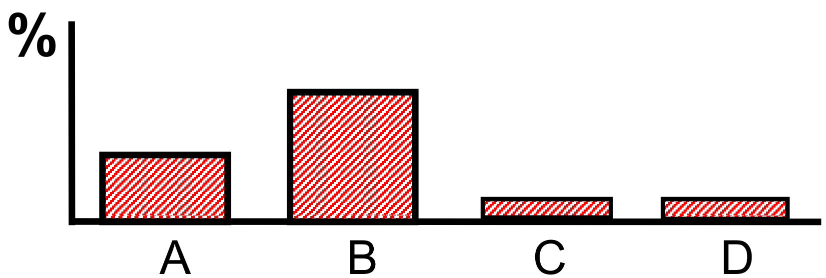
$$\chi^2_{(3)} = \frac{(32 - 64,13)^2}{64,13} + \ldots\ldots\ldots + \frac{(116 - 128,27)^2}{128,27} = \underline{\underline{49,06}}$$

$$\textit{Tabulky}: \chi^{2\,(\nu)}_{1-\alpha} = \chi^{2\,(3)}_{0,95} = 7,81$$

Zamítáme hypotézu shody pozorovaných četností s očekávanými.

Příspěvek kategorií A, B, C, D k celkové hodnotě $\chi 2$

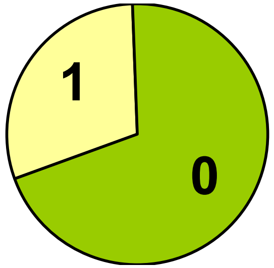




Test dobré shody – binomické data

Binomické jevy (1/0)

$$\chi^2_{(1)} = \underbrace{\frac{\left[\text{pozorovaná četnost} - \text{očekávaná četnost}\right]^2}{\text{očekávaná četnost}}}_{\text{I. jev 1}} + \underbrace{\frac{\left[\text{pozorovaná četnost} - \text{očekávaná četnost}\right]^2}{\text{očekávaná četnost}}}_{\text{II. jev 2}}$$



Příklad

✓ **10 000 lidí hází mincí** → **rub: 4 000 případů (R)**; **líc: 6 000 případů (L)**

? **Lze výsledek považovat za statisticky významně odlišný (nebo neodlišný) od očekávaného poměru R : L = 1 : 1 ?**

$$\chi^2_{(1)} = \frac{(4000 - 5000)^2}{5000} + \frac{(6000 - 5000)^2}{5000} = 400$$

Tabulková hodnota: $\chi^2_{(0,95)}(\nu = 1) = \underline{\underline{3,84}}$ $\quad (0,95 = 1 - \alpha)$

Rozdíl je vysoce statisticky významný (p << 0,001)

Test nezávislosti

- Sledujeme dva znaky.
- Tieto znaky nadobúdajú len konečné množstvo hodnôt.
  - Napríklad: farba vlasov: - svetlá, gaštanová, čierna, hrdzavá
  - Napríklad: farba očí: modrá, šedá-zelená, hnedá

- Chceme testovať, či sú tieto znaky nezávislé
- H0: znak 1 a znak 2 sú nezávislé proti H1: sú na sebe závislé
  - H0: farba vlasov a farba očí sú na sebe nezávislé
  - H1: farba vlasov a farba očí sú na sebe závislé
- H0 zamietame, ak je vyrátaná hodnota väčšia ako príslušná tabuľková alebo porovnaním p-hodnoty a hladiny významonosti

Test nezávislosti
H0 :Nezávislost dvou jevů A a B

B ↓ \ A →	+	-	Podíl (+)
+	a	b	$\frac{a}{(a+b)}$ p_1
-	c	d	$\frac{c}{(c+d)}$ p_2
Podíl (+)	$\frac{a}{(a+c)}$	$\frac{b}{(b+d)}$	

Kontingenční tabulka 2 x 2

N = a + b + c + d

$$P(B^{+}) = \frac{(a+b)}{N}$$

$$P(B^{-}) = \frac{(c+d)}{N}$$

Očekávané četnosti:

$$F_{(A)} = \frac{(a+b)(a+c)}{N} \qquad F_{(C)} = \frac{(a+c)(d+c)}{N}$$

$$F_{(B)} = \frac{(a+b)(b+d)}{N} \qquad F_{(D)} = \frac{(b+d)(c+d)}{N}$$

$$\chi^2_{\nu=1} = \sum_{i=1}^{4} \frac{(f_i - F_i)^2}{F_i}$$

$$\nu = 1 = (r-1)*(c-1)$$

$$P_{(A)};\ P_{(B)}$$

$$\chi^2_c = \sum \sum \frac{\left(\left|f_{ij} - F_{ij}\right| - 0{,}5\right)^2}{F_{ij}}$$

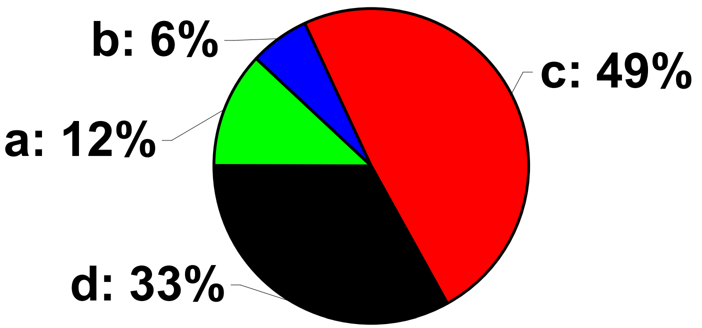
Kontingenční tabulky: příklad

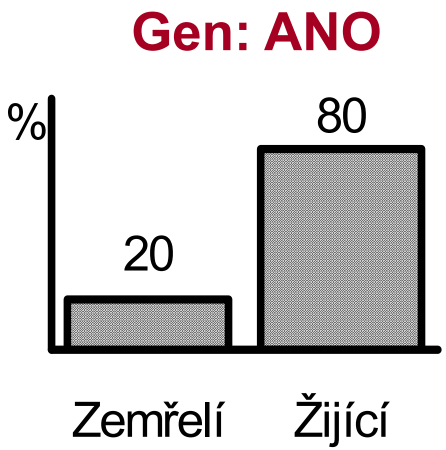
gen \ †	Ano	Ne	Σ
Ano	20	82	102
Ne	10	54	64
Σ	30	136	166

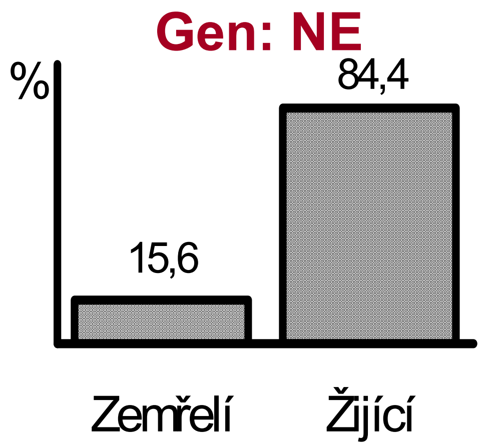
$F_A = 102 * 30 / 166 = 18,43$
$F_B = 102 * 136 / 166 = 83,57$
$F_C = 11,57$
$F_D = 52,43$

$$\chi^2_{(1)} = \frac{(20-18{,}43)^2}{18{,}43} + \frac{(82-83{,}57)^2}{83{,}57} + \frac{(10-11{,}57)^2}{11{,}57} + \frac{(54-52{,}43)^2}{52{,}43} = 0{,}423 \qquad 0{,}423 < \chi^{2\ (1)}_{0,95} = 3{,}84$$

Kontingenční tabulka v obrázku







R x C kontingenční tabulka

Výběr: N lidí ze sociologického průzkumu (delikventi)
Jev **A**: Původ z rozvrácených rodin
Jev **B**: Stupeň zločinnosti I < II < III < IV

A \ B	I.	II.	III.	IV.	Σ
ANO	a	b	c	d	číslo 1
NE	e	f	g	h	
Σ	číslo2				

Stupně volnosti:
(R-1) * (C-1) = 1 * 3 = 3

$$F_a = \frac{\check{c}\acute{\imath}slo\,1 \cdot \check{c}\acute{\imath}slo\,2}{N}$$

Tabulky: $\chi^{2}_{(1-\alpha)}{}^{(\nu)}$

Očekávané četnosti:

$$p_a = \frac{a}{a+e} \qquad p_b = \frac{b}{b+f} \qquad p_c = \frac{c}{c+g} \qquad p_d = \frac{d}{d+h}$$

Test homogenity

- Pravdepodobnosť výskytu znaku v stĺpcoch nezávisí na riadkoch
- Stĺpce napr.: krvné skupiny (0, A, B, AB)
- Riadky napr.: kraje
- H0: Zastúpenie jednotlivých krvných skupín je v každom kraji rovnaký

Test homogenity: příklad

Pomocí χ2 rozložení lze rovněž posuzovat homogenitu většího množství nezávislých pokusů testujících tutéž hypotézu.

Bylo provedeno 6 nezávislých výběrů z populace mladých mužů, kteří v dětství onemocněli těžkým zánětem mozkových blan.
H_0: V této populaci se vyskytují praváci a leváci v poměru 1 : 1.

Nalezněte v literatuře příslušné vztahy pro testování homogenity všech šesti výběrových populací a na základě výsledků tohoto testu rozhodněte o dalším postupu.

Následující tabulka obsahuje původní data a výsledek testování (v závorkách jsou uvedeny očekávané četnosti):

Vzorek	Praváci	Leváci	n	χ2	St. volnosti
1	3 (7)	11 (7)	14	4,5714	1
2	4 (8)	12 (8)	16	4,000	1
3	15 (10)	5 (10)	20	5,000	1
4	14 (9)	14 (9)	18	5,5556	1
5	13 (8,5)	4 (8,5)	17	4,7647	1
6	17 (11)	5 (11)	22	6,5455	1

$$\chi^2_{heterogenita} = 30{,}2$$

$$\nu = s - 1 = 5$$

$$P < 0{,}001$$

Jednoduchým testováním lze zjistit, že všechny testy pro jednotlivé výběry jsou významné, což znamená, že ani v jednom případě nebyla potvrzena shoda očekávaných a pozorovaných četností. Test homogenity štěpného poměru v zkoumaných populacích rovněž vedl k zamítnutí možnosti sloučit jednotlivé výběry a posuzovat je jako celek (kromě testovaného poměru 1 : 1 neexistuje tedy v datech žádný jiný jednotný štěpný poměr mezi oběma vlastnostmi.

V případě, že by tento test neprokázal odchylky mezi jednotlivými výběrovými populacemi, bylo by možné jednotlivé odběry sloučit a posuzovat jako homogenní vzorek.

Test homogenity binomických rozložení

✓ **Jev: Úmrtnost na leukemii**
Předpoklad: $\Pi = 0{,}6$
Absolutní četnost jevu označena r_i

$$\bar{p} = \frac{\sum p_i}{S}$$

Sledovalo s autorů z s zemí:

Autor	n_i	r_i	p_i
1			
2			
.			
.			
.	$\sum n_i = N$		
s			

Test homogenity binomických rozložení

$$\chi^2_{S-1} = \frac{\left(\sum r_i\, p_i - \bar{p} \sum r_i\right)}{\bar{p}\left(1 - \bar{p}\right)}$$

Po možném sloučení s výběrů

Test shody reálného r $\left(\sum r_i\right)$ a $n \cdot \Pi$

$$\chi^2_{(1)} = \frac{\left(\left|\sum r_i - N \cdot \Pi\right| - \frac{1}{2}\right)^2}{N \cdot \Pi \cdot (1 - \Pi)}$$

$\chi 2$ test - příklad složitější kontingenční tabulky I

Caffeine consumption and marital status in antenatal patiens (from Martin and Bracken, 1987)

	Caffeine consumption (mg/day)				
Marital status	**0**	**1 - 150**	**151 - 300**	**> 300**	**Total**
Married	652	1537	598	242	3029
Divorced, separed or widowed	36	46	38	21	141
Single	218	327	106	67	718
Total	906	1910	742	330	3888

Caffeine consumption and marital status data

	Caffeine consumption (mg/day)				
Marital status	**0**	**1 - 150**	**151 - 300**	**> 300**	**Total**
Married	22 %	51 %	20 %	8 %	3029 (100 %)
Divorced, separed or widowed	26 %	33 %	27 %	15 %	141 (100 %)
Single	30 %	46 %	15 %	9 %	718 (100 %)
Total	23 %	49 %	19 %	8 %	3888 (100 %)

$\chi 2$ test - příklad složitější kontingenční tabulky II

Expected frequencies

	Caffeine consumption (mg/day)				
Marital status	**0**	**1 - 150**	**151 - 300**	**> 300**	**Total**
Married	705,8	1488	578,1	257,1	3029
Divorced, separed or widowed	32,9	69,3	26,9	12,0	141
Single	167,3	352,7	137	60,9	718
Total	906	1910	742	330	3888

Contributions of each cell

	Caffeine consumption (mg/day)				
Marital status	**0**	**1 - 150**	**151 - 300**	**> 300**	**Total**
Married	4,11	1,61	0,69	0,89	7,30
Divorced, separed or widowed	0,30	7,82	4,57	6,82	19,51
Single	15,36	1,88	7,02	0,60	24,86
Total	19,77	11,31	12,28	8,31	51,66

$\chi2$ test - příklad frakcionace složitější kontingenční tabulky I

Cílem rozsáhlejšího průzkumu populace bylo prozkoumat vztah mezi dvěma typy chorob a krevními skupinami u lidí. Konkrétní data jsou uvedena v tabulce:

Krevní skupina	Žaludeční vředy	Rakovina žaludku	Kontrola	Celkem
0	983	383	2892	4258
A	679	416	2625	3720
B	134	84	570	788
Celkem	1796	883	6087	8766

Vypočítejte testovou charakteristiku pro tuto kontingenční tabulku a otestujte nulovou hypotézu nezávislosti jevů ($\chi2$ = 40,54; 4 st. volnosti)

$\chi 2$ test - příklad frakcionace složitější kontingenční tabulky II

K podrobnějšímu průzkumu složitějších tabulek výrazně napomáhá přepis původní tabulky do podoby procentického zastoupení kategorií:

Krevní skupina	Žaludeční vředy	Rakovina žaludku	Kontrola
0	983	383	2892
A	679	416	2625
B	134	84	570
Celkem	1796	883	6087

Z této tabulky je patrné:

1. **Jsou jenom malé rozdíly v distribuci krevních skupin u kontroly a u skupiny nemocných rakovinou žaludku.**
2. **Pacienti s vředy mají mnohem častěji krevní skupinu 0.**

Na základě těchto poznatků je možné sestrojit menší kontingenční tabulku, která otestuje hypotézu o shodné distribuci krevních skupin pro nemocné rakovinou a pro zdravé lidi.

Sestavte tuto tabulku a otestujte nulovou hypotézu.
($\chi 2$ = 5,64 (2 st. v.), P je přibližně rovna 0,06)

$\chi 2$ test - příklad frakcionace složitější kontingenční tabulky III

- Z tohoto dílčího testu vyplývá možnost sloučení skupiny nemocných rakovinou a zdravých lidí neboť se vzhledem k distribuci krevních skupin chovají jako homogenní populace. Dalším logickým krokem v podrobné analýze je testování shody relativních četností výskytu krevních skupin A a B mezi kombinovaným vzorkem (sloučená skupina s rakovinou a kontrola) a mezi vzorkem lidí nemocných žaludečními vředy - tzn. nyní neuvažujeme krevní skupinu 0. Výsledkem tohoto testu je $\chi 2 = 0{,}68$ (1 st. vol.); $P > 0{,}7$. Vzorky pro krevní skupiny A a B lze tedy sloučit do směsného vzorku A + B.

- Nyní otestujeme shodu relativních četností výskytu skupiny 0 oproti A + B, a to mezi kombinovanou populací (kontrola + nemocní rakovinou) a mezi vzorkem nemocných vředařů ($\chi 2 = 34{,}29$; 1 st. vol.). Lze tedy shrnout, že vysoká hodnota původního $\chi 2$ se 4 st. volnosti byla způsobena zvýšenou četností lidí s krevní skupinou 0 mezi nemocnými žaludečními vředy.

χ2 test - příklad frakcionace složitější kontingenční tabulky IV

Průběh hodnocení lze shrnout do tabulky:

Srovnání	St. volnosti	χ2
0, A, B skupina u pacientů s rakovinou (**r**) x kontrola (**k**)	2	5,64
A, B skupina u pacientů s vředy x kombinovaný vzorek (**r + k**)	1	0,68
0, A, B skupina u pacientů s **s** vředy x kombinovaný vzorek (**r + k**)	1	34,29
Celkem	**4**	**40,61**

Celkový součet testových statistik χ2 (40,61) odpovídá přibližně původní hodnotě χ2 (40,54). Což platí i o stupních volnosti (4). Tato skutečnost potvrzuje, že jsme detailním rozborem vyčerpali informační obsah původní kontingenční tabulky a kromě popsané závislosti (zvýšený výskyt krevní skupiny 0 u lidí s žaludečními vředy) jsou jednotlivé kategorie zkoumaných jevů zcela nezávislé.