Porovnání distribucí u párově závislých dat ve čtvercových kontingenčních tabulkách

JAN ŘEHÁK
Ústav pro filozofii a sociologii ČSAV, Praha
BLANKA ŘEHÁKOVÁ
Ústav pro výzkum veřejného mínění při FSÚ, Praha

Párově závislá data se ve statistické literatuře nazývají také párově spřažená, vzájemně propojená, data z párově závislých výběrů, párovaná data, data závislých dvojic. Jejich zpracování vychází z modelu čtvercové kontingenční tabulky. V sociologických analýzách se vyskytují často a ve značně rozličných situacích. Je proto překvapující, že metodika jejich zpracování není mezi výzkumníky nikterak známa a že nebyla zařazena ani do renomovaných programových systémů. Přitom teorie zpracování čtvercových tabulek je značně bohatá a řeší celou řadu užitečných úloh. (Práce v tomto směru však zdaleka nebyla završena.)

Cílem této stati je shrnout známé metody (McNemarův, Stuartův, Bhapkarův, binomický test) pro jednu z úloh, které lze ve čtvercových tabulkách řešit — pro porovnání marginálních distribucí, a to pro případ nominálních znaků, obohatit metodiku analýzy pro obdobnou situaci s ordinálními znaky a navrhnout některé postupy a míry pro případ neshody obou distribucí.

V první části popíšeme ty analytické situace, které jsou pro sociologickou práci typické a pro něž je uváděná metodika vhodná, druhá část obsahuje statistické postupy pro nominální znaky, čtvrtá pro ordinální znaky a v částech třetí a páté jsou ilustrační příklady z analytické praxe.

Případ dichotomických znaků není samostatně vydělen a rozebírán, neboť byl obsahem statí J. Řeháka—B. Řehákové [1978a], [1978b], v nichž byla ukázána aplikace znaménkového testu i jeho asymptotické normální, resp. chí-kvadrátové varianty.

U vícehodnotových znaků se mění úloha podle jejich typů. Testování nulové hypotézy $\boldsymbol{H_0}$: „*obě distribuce jsou shodné*" provádíme proti různým alternativám:

a) pro nominální znaky
   $\boldsymbol{H_1}$: „*distribuce jsou rozdílné*" (liší se v četnosti alespoň u jedné kategorie),
b) pro ordinální znaky
   $\boldsymbol{H_2}$: „*distribuce se liší ve svém tvaru*"(liší se jejich kumulační distribuční funkce),
   $\boldsymbol{H_3}$: „*distribuce se liší posunutím mediánů*" (rozdílností předem nepředvídanou),
   $\boldsymbol{H_4}$: „*medián distribuce* **A** *je vyšší než medián pro* **B** (obdobně „nižší"),
c) pro kardinální znaky
   $\boldsymbol{H_5}$: „*průměry obou distribucí se liší*"
   $\boldsymbol{H_6}$: „*průměr pro proměnnou* $\boldsymbol{X}$ *je vyšší než pro* $\boldsymbol{Y}$ (obdobně „nižší").

Tyto alternativy jsou obdobné jako alternativy u porovnání dvou nezávislých souborů; porovnání zde však nevychází z tabulky $2 \times K$ (řádky $\equiv$ soubory, sloupce $\equiv$ $\equiv$ hodnoty proměnné), ale z tabulky $K \times K$ (řádky $\equiv$ hodnoty proměnné **A**, sloupec $\equiv$ hodnoty proměnné **B**).[1])

1. Typy párově závislých dat a sociologické úlohy

Údaje párově závislých výběrů nebo párově spřažená data lze charakterizovat dvěma vlastnostmi (a zároveň podmínkami):

— dvě proměnné, $\mathbf{A} = \{a_1, a_2, \ldots, a_K\}$, $\mathbf{B} = \{b_1, b_2, \ldots, b_K)\}$ mají *stejně dlouhé stupnice*, jsou *stejného typu* a *hodnoty vzájemně obsahově korespondují* ($a_k \leftrightarrow b_k$);
— data tvoří *jeden soubor dvoučlenných vektorů*, v nichž první složka odpovídá realizaci proměnné **A** a druhá složka realizaci proměnné **B**, je to tedy matice dat o rozměru $n \times 2$ (n je počet datových dvojic).

Soubor s uvedenými vlastnostmi může vzniknout dvěma způsoby:

1. jde o zjišťování formálně stejných, ale obsahově se lišících proměnných **A**, **B** u souboru n jednotek,

2) jde o zjišťování téže vlastnosti u dvou typů jednotek ve spřažené dvojici (soubor n dvojic), přičemž **A** je informace o jednotce prvního typu a **B** o jednotce druhého typu.

Tabulka 1. Schematická tabulka čtvercového tvaru pro porovnání dvou závislých distribucí

a) Tabulka absolutních četností

		Proměnná B				
		b_1	b_2	b_k	b_K	Součet
Proměnná A	a_1	n_{11}	n_{12}	n_{1k}	n_{1K}	n_{1+}
	a_2	n_{21}	n_{22}	n_{2k}	n_{2K}	n_{2+}
	. . .					
	$a_{k'}$	$n_{k'1}$	$n_{k'2}$	$n_{k'k}$	$n_{k'K}$	$n_{k'+}$
	a_K	n_{K1}	n_{K2}	m_{Kk}	n_{KK}	n_{K+}
	Součet	n_{+1}	n_{+2}	n_{+k}	n_{+K}	n

b) Tabulka teoretických relativních četností (resp. teoretických pravděpodobností)

		Proměnná B						
		b_1	b_2	...	b_k	...	b_K	Součet
Proměnná A	a_1	p_{11}	p_{12}	...	p_{1k}	...	p_{1K}	p_{1+}
	$a_{2\check{r}}$	p_{21}	p_{22}	...	p_{2k}	...	p_{2K}	p_{2+}
	. . .							
	$a_{k'}$	$p_{k'1}$	$p_{k'2}$	...	$p_{k'k}$	...	$p_{k'K}$	$p_{k'+}$
	. . .							
	a_K	p_{K1}	p_{K2}	...	p_{Kk}	...	$p_{KK\,k}$	p_{K+}
	Součet	p_{+1}	p_{+2}	...	p_{+k}	...	p_{+K}	1

Data (v obou případech) lze po vytřídění uspořádat do čtvercové tabulky $\mathbf{A} \times \mathbf{B}$,[2]) která má rozměry $K \times K$ a jejíž řádky a sloupce jsou řazeny stejně podle korespondujících hodnot. Četnosti v tabulce odpovídají buď počtu jednotek (v prvním případě), nebo počtu dvojic (ve druhém případě).

V tabulce 1a značíme:

n_{jk} = počet případů, pro něž $\mathbf{A} = \boldsymbol{a}_j$, $\mathbf{B} = \boldsymbol{b}_k$,

$n_{k+} = \sum_{j=1}^{K} n_{kj}$ = počet případů pro něž $\mathbf{A} = \boldsymbol{a}_k$,

$n_{+k} = \sum_{j=1}^{K} n_{jk}$ = počet případů pro něž $\mathbf{B} = \boldsymbol{b}_k$,

$n = \sum_{j=1}^{K} \sum_{k=1}^{K} n_{jk}$ = počet všech tříděných případů.

(Případem rozumíme buď jednotku, nebo dvojici, a to podle rozlišení uvedeného výše).

V tab. 1b jsou obdobně zavedeny populační relativní četnosti (resp. pravděpodobnosti): p_{jk}, p_{k+}, p_{+k}. Ty nejsou známy, ale odhadujeme je pomocí relativních četností $f_{jk} = n_{jk}/n$, $f_{k+} = n_{k+}/n$, $f_{+k} = n_{+k}/n$, které tvoří obdobnou tabulku.

V sociologické analýze se objevují popsaná data v celé řadě různých situací, z nichž vybíráme typické.

A) *Dva paralelní dotazy témuž respondentovi (nebo obecně dvě souběžně zjišťované skutečnosti o jedné jednotce)*

U dvou proměnných je určena stejná škála odpovědí, otázky jsou však odlišné, ale obsahově příbuzné, takže nás zajímá porovnání výsledků u obou.

Příklad: a) *soubor*: osazenstvo závodu s vysokou fluktuací
otázky: „Jaký vliv bude mít na odchody dělníků opatření?"
A ≡ výstavba vlastní závodní jídelny
B ≡ investice do nových strojů
odpovědi: — lidé budou odcházet stejně jako dřív;
— odchody se značně sníží, ale nejde o jedinou příčinu;
— lidé přestanou odcházet, jde o hlavní příčinu.
úloha pro analýzu dat: Liší se názor na obě příčiny u osazenstva závodu? Která z příčin je považována za závažnější?

b) *soubor*: televizní diváci
otázky: **A** ≡ „Jaký typ pořadu máte v TV nejraději, když přijdete unaven z práce a chcete si odpočinout?"
B ≡ „Jaký typ pořadu máte v TV nejraději o víkendu či o volných dnech, kdy máte klid a dost času?"
odpovědi pro obě otázky:
$\boldsymbol{a}_1 \equiv \boldsymbol{b}_1 \equiv$ sportovní přenos,
$\boldsymbol{a}_2 \equiv \boldsymbol{b}_2 \equiv$ dobrodružný nebo detektivní film nebo inscenace;
$\boldsymbol{a}_3 \equiv \boldsymbol{b}_3 \equiv$ zábavní pořad s hudbou;
$\boldsymbol{a}_4 \equiv \boldsymbol{b}_4 \equiv$ obrazová reportáž z událostí ve světě, ...
úloha pro analýzu dat:
Liší se preference u populace TV diváků v situaci A od preferencí v situaci B? U kterých kategorií?

Případ lze zobecnit na situaci, v níž se vyskytuje podobných otázek více a chceme je srovnávat navzájem. To znamená $M(M-1)/2$ dvojic pro M proměnných a stejně tolik čtvercových tabulek.

Příklad (pokr.) a) Dotaz na účinnost opatření můžeme rozšířit na celou baterii různých příčin
C ≡ investice do odstranění hlučnosti
D ≡ investice do odstranění prašnosti
E ≡ investice do regulace tepla a větrání
F ≡ zvýšení mezd a příplatků
Šest distribucí vede v úplné analýze na $6 \times 5/2 = 15$ tabulek o rozměrech 3×3.
b) Dotazy na TV pořady můžeme rozšířit na:

C ≡ „Jaký typ pořadu máte nejraději vůbec?"
D ≡ „Jaký typ pořadu má nejraději manžel(ka)?"
E ≡ „Jaký typ pořadu mají nejraději diváci?"
F ≡ „Jaký typ pořadu máte nejraději v rozhlase?"
(s obdobnými a srovnatelnými kategoriemi odpovědí)

Všech komparací je 15, v analýze však nás mohou zajímat jen některé, např.: **A** × **B**, **A** × **C** **B** × **C**, **C** × **D**, **C** × **E**, **C** × **F**.

B) *Dva údaje odpovídají měření téže vlastnosti ve dvou časově posunutých situacích*

Nejčastěji se vyskytují situace:

— měření vlastnosti před a po určitém zásahu do zkoumaného systému (např. zavedení *Souboru opatření*, nová technika),
— měření stavu před a po určité samovolné změně v systému, který zkoumáme (změna ve vedení úmrtím, odchodem do důchodu, havárie, výpadek techniky),
— měření vlastnosti před a po určité změně v životě jednotlivce, ve vývoji jednotky (např. vstup do manželství, rozvod, narození dítěte, koupě barevného televizoru, automobilu ap.),
— měření vlastnosti před a po záměrně vyvolaném působení (např. školení, poskytnutí informace),
— panelové studie, tj. měření vlastností těchže jednotek v určeném časovém odstupu za účelem zhodnocení časového vlivu.

Příklad: a) *soubor*: osazenstvo závodu se silnou fluktuací;
otázka položená před (**A**) a po (**B**) úpravě pracovních podmínek a přestavbě závodu: „Jaká je nejsilnější příčina fluktuace v závodě?
odpovědi: 1 ≡ nízké platy, 2 ≡ nízká péče o dělníky, 3 ≡ špatné pracovní podmínky, 4 ≡ těžká práce, 5 ≡ špatné vedení, ...
otázka pro analýzu dat: měly změny v podmínkách vliv na názor dělníků?

b) *soubor*: ekonomicky aktivní občané ve věku 25 až 30 let.
data: informace o oboru, v němž se respondent vyučil, (proměnná **A**) a oboru, v němž v současné době pracuje (proměnná **B**).
otázka pro analýzu dat: nastává při změnách oborů celkové vyrovnání v dané věkové skupině, či některé obory znamenají významný odliv a příliv?

c) *soubor*: diváci TV zahrnutí do panelového souboru
data: údaje o sledování čtyř pořadů *R*, *S*, *T*, *V*, které byly na programu ve dvou zvolených dnech, jsou získány ve tvaru „*sledoval—nesledoval*" a jsou převedeny do nominálních proměnných (**A** ≡ sledování v prvním dnu, **B** ≡ sledování v druhém dnu) s hodnotami: $a_1 \equiv b_1 \equiv$ sledoval všechny pořady *R*, *S*, *T*, *V*; $a_2 \equiv b_2 \equiv$ sledoval *R*, *S*, *T*; $a_3 \equiv b_3 \equiv$ sledoval *R*, *S*, *V* nebo *R*, *S*; $a_4 \equiv b_4 \equiv$ sledoval *R*, *T*, *V* nebo *R*, *T*; $a_5 \equiv b_5 \equiv$ sledoval *R* nebo *R*, *V*; $a_6 \equiv b_6 \equiv$ jiná kombinace; $a_7 \equiv b_7 \equiv$ nesledoval žádný z pořadů (jako rozhodnutí); $a_8 \equiv b_8 \equiv$ nemohl sledovat TV;
otázka pro analýzu dat: Je sledovanost uvedených kombinací stejná v obou dnech? Není-li, ve kterých kategoriích nastává změna?

Případ lze zobecnit na údaje o více časových bodech. Je-li jich M (označme je $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_M$), pak bude nejčastější porovnání vycházet z kombinací $\mathbf{A}_1 \times \mathbf{A}_2$, $\mathbf{A}_2 \times \mathbf{A}_3, \ldots, \mathbf{A}_{M-1} \times \mathbf{A}_M$ (tj $M - 1$ porovnání). V některých případech však budou zajímavá postupná srovnání s výchozím stavem: $\mathbf{A}_1 \times \mathbf{A}_2$, $\mathbf{A}_1 \times \mathbf{A}_3, \ldots$ $\ldots, \mathbf{A}_1 \times \mathbf{A}_M$ (je jich $M - 1$) nebo s konečným stavem: $\mathbf{A}_1 \times \mathbf{A}_M$, $\mathbf{A}_2 \times A_M$, , $\ldots, \mathbf{A}_{M-1} \times \mathbf{A}_M$ (je jich $M - 1$).
Zajímavé též může být srovnání vlivu zpoždění o jeden krok: $\mathbf{A}_1 \times \mathbf{A}_3$, $\mathbf{A}_2 \times \mathbf{A}_4$, ... (je jich $M - 2$) atd. Jinou úlohou je nalézt počet kroků, po nichž se začne projevovat signifikantní rozdíl. Vcelku máme[3]) $M(M - 1)/2$ srovnávacích čtvercových tabulek, většinou nás však zajímají jen vybrané.

C) *Zjišťování téže vlastnosti v různých podmínkách*

má za cíl najít vliv podmínek na zkoumanou vlastnost. Odpovídá situaci experimentu, v němž podmínky jsou záměrně (experimentálně) měněny či jsou analogií řízených

situací. V sociologii je provedení klasického experimentu velmi řídké, neboť kontrola faktorů se většinou vymyká možnostem výzkumníka (používá se však často pro metodologické účely), proto mluvíme pouze o *statistickém uspořádání experimentu* abychom naznačili, že jde o situace, v nichž lze aplikovat modely statistických postupů odvozených pro experimentální situace. Nejde však o experiment v přísném slova smyslu, a tudíž interpretace výsledků tomu musí odpovídat.

Od předchozí situace (B) se případ metodologicky liší tím, že podmínky, za nichž proměnné měříme, nemusí být realizovány ve stejném (pevně daném) pořadí.

Příklady: a) tentýž dotaz na spokojenost s prací v pracovním prostředí a doma;
b) dotaz na kvalitu TV obrazu po vzrušujícím sportovním utkání a po běžném večerním pořadu;
c) výpověď respondenta při řízeném rozhovoru a výpověď při skupinovém rozhovoru;
d) výběr oblíbeného herce (jako vzor, typ) v normální situaci a výběr po shlédnutí filmu, divadelní hry nebo TV inscenace, v nichž některý z populárních herců hraje úspěšně hlavní roli.

Případ lze rozšířit na více (M) různých situací (například v příkladu d) může být dotaz opakován po několika filmech s různými herci). Porovnání v těchto případech závisí na roli jednotlivých proměnných ve struktuře analytického plánu. Nejčastější situace budou dvě:

— porovnání všech $M(M-1)/2$ dvojic situací (tj. zhodnocení všech čtyřpolních tabulek $\mathbf{A}_i \times \mathbf{A}_j$;

— porovnání vybraných situací (proměnné $\mathbf{B}_1, \mathbf{B}_2, \ldots \mathbf{B}_M$) s jednou výchozí (normální, neovlivněnou, standardní, obvyklou, běžnou, pravidelnou ap.) situací (proměnná **A**). Pak $(M+1)$ proměnných dává M porovnání (tabulek): $\mathbf{A} \times \mathbf{B}_1, \mathbf{A} \times \mathbf{B}_2, \ldots$ $\ldots, \mathbf{A} \times \mathbf{B}_M$.

D) *Studie reliability, komparace měřících prostředků*

Spočívá v opakovaném měření dvěma různými instrumenty nebo v opakovaném dotazu s určitým časovým odstupem. Přitom zkoumáme stupeň odchylek i jejich kvalitativní příčiny (náhodnost a systematičnost). Úloha porovnání distribucí však není pro tento případ charakteristická a hlavní. Úlohu zde uvádíme jen jako výrazný a častý případ párovaných dat (měla by být řešena de facto v každém předvýzkumu).

E) *Párované jednotky* (*párově závislé soubory*)

se vyskytují v různých situacích sběru a analýzy dat:

— dvojice, které tvoří strukturovanou, ale přirozeně existující skupinku, v níž jsou jednotky odlišeny svojí podstatou, rolí, funkcí, obsahem;

— dvojice, které sestavujeme pro analýzu tak, že se jejich prvky od sebe liší určitými předem zvolenými vlastnostmi, zatímco jiné (zvolené, kontrolované) vlastnosti mají shodné;

— k sobě přiřazené dvojice v experimentu; jejich prvky mají společné vlastnosti určené modelem experimentu a liší se tím, že na jednu jednotku experimentálně působíme a na druhou ne, popřípadě působíme jinak na první a jinak na druhou.[4])

Smyslem analýzy je sledovat vliv diferenciační vlastnosti u jednotek dvojice s vyloučením vlivu společných vlastností definujících dvojici. V sociologické analýze se nejčastěji vyskytuje případ přirozeně existujících dvojic.

Příklad: a) preference koupě příštího předmětu pro domácnost u muže (**A**) a u ženy (**B**) — společné je tu manželství, byt, životní způsob, dosavadní vybavení, rozdílné je tu hledisko muže a ženy;
b) preference trávení volného času muže (**A**) a ženy (**B**) u manželských dvojic;
c) typ osobnosti ředitele (**A**) a jeho prvního zástupce (**B**) nebo jejich názor;
d) povolání otce (**A**) a syna (**B**) (intergenerační mobilita);
e) typ hlavního pořadu na 1. kanálu TV (**A**) a na 2. kanálu TV (**B**) v určitém intervalu vysílacího času dne (n = počet hodnocených dní; jeden hodnotitel);

Případ párovaných jednotek lze rozšířit též na skupinky o více členech. Tyto skupinky jsou však opět strukturované a údaje **A**, **B**, **C**, ... odpovídají přesně definovaným postavením členů ve skupince (z hlediska řešení úlohy).[5]) Možnosti rozšíření jsou velmi bohaté.
Příklad (pokr.):

V případě a) **C** ≡ názor tchána (otce manžela)

V případě b) **C** ≡ názor staršího dítěte
D ≡ názor mladšího dítěte
(pro rodiny se dvěma dětmi ve věku od 5—15 let)

V případě c) **C** ≡ ředitel pro kádrovou a personální práci

V případě d) **C** ≡ povolání matky
D ≡ povolání dcery

V případě e) **C** ≡ typ hlavního pořadu na rozhlasové stanici Praha (v téže době)
D ≡ typ hlavního pořadu na rozhlasové stanici Vltava (v téže době)
(typy pořadů jsou voleny srovnatelné pro televizi a rozhlas).

Dalším obdobným případem je porovnání subjektivních a objektivních údajů, resp. měřených a skutečných, odhadovaných a faktických. Metodologické i statistické aspekty této úlohy však jsou od předchozích odlišné. Proto lze data uspořádat obdobně (**A** ≡ subjektivní názor, měření, odhad; **B** ≡ objektivní údaj, skutečnost, fakt), úlohu sem neřadíme. Cílem analýzy dat v takovém případě je určit stupeň objektivity či přesnosti jednoho souboru (nepřesných) údajů proti zjištěné (přesné) informaci.

Uvedené příklady měly charakterizovat především párově závislá data v různých analytických situacích. Téměř ve všech uvedených příkladech má analýza shody, resp. neshody, marginálních distribucí zásadní význam, lze však řešit i další úlohy, které už nelze popsat a rozebrat v rámci rozsahu tohoto článku.

2. Statistická metodika porovnání párově závislých distribucí-nominální znaky

Porovnání dvou párově závislých distribucí $\{p_{k+}\}$ a $\{p_{+k}\}$ z tab. 1 (nazýváme je též marginální distribuce) provádíme v několika krocích:

krok 1: *test hypotézy shody distribucí*: v případě ověření shody konstatujeme výsledek, v případě odmítnutí shody postupujeme dále;

krok 2: *měření stupně neshody*;

krok 3: určení kategorií, v nichž se projevuje *významný rozdíl*;

krok 4: sestrojení *znaménkového schématu* pro rozdíly v kategoriích, které ukazuje směr rozdílů a stupeň jejich statistických významností.

Tyto čtyři kroky lze provést statistickou metodikou, která je v úplnosti navržena v této části. Základní, první krok, je založen na Stuartově (resp. Bhapkarově) testu a pro dvojhodnotové znaky (resp. pro jevy) na McNemarově nebo binomickém (znaménkovém) testu. Oba testy pro vícehodnotové znaky vyžadují pro větší K počítač nebo kalkulačku, která poskytuje inverzi matice, proto je též navržena metoda simultánního postupu s využitím snadno aplikovatelných testů McNemarových, pro něž uvádíme tabulku postupných hladin významnosti (tab. A). V této části je též navržena metodika dalších kroků, které aplikujeme po zamítnutí hypotézy shody.

Krok 1. Testování shody závislých distribucí — nominální znaky

U dvou párově závislých znaků A a B předpokládáme obsahovou korespondenci hodnot (a_k, b_k), $k = 1, 2, ..., K$, a tudíž smysluplnost komparace četností jejich populačních výskytů p_{k+} a p_{+k}.

Formulace hypotéz:

$\boldsymbol{H_0}$: $p_{k+} = p_{+k}$ pro všechna $k = 1, 2, ..., K$

$\boldsymbol{H_A}$: alespoň pro jedno k je $p_{k+} \neq p_{+k}$

Nulová hypotéza značí shodu marginálních distribucí, alternativní hypotéza říká, že alespoň u jedné kategorie se populační četnosti liší, není však předem specifikováno, u které (může to být kterákoli).
Data: Absolutní četnosti $\|n_{jk}\|$, uspořádané ve čtvercové tabulce o rozměru $K \times K$, s počtem pozorování n (viz tabulka 1a).
Intuitivní model: Čím blíže budou všechna čísla $f_{k+} - f_{+k}$ $(= (n_{k+} - n_{+k})/n)$ k nule, tím spíše bude platit $\boldsymbol{H}_0$. Čím větší bude zvolená míra nepodobnosti obou distribucí (tj. alespoň některé z rozdílů $f_{k+} - f_{+k}$ budou v absolutní hodnotě velké), tím spíše bude platit $\boldsymbol{H}_A$.

Postupy testování $\boldsymbol{H}_0$ *vs* $\boldsymbol{H}_A$:

A) *Případ dvouhodnotového znaku*[6] $(K = 2)$

A1) *McNemarův test* $(n \geq 30)$

a) zvolíme hladinu významnosti α a k ní zjistíme kritickou hodnotu $\chi^2_{\alpha,1}$ pro test chí-kvadrát s df = 1

b) spočteme statistiku Q:

$$(1) \qquad Q = \frac{(n_{1+} - n_{+1})^2}{n_1 + n_{+1} - 2n_{11}} = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}}$$

c) je-li $Q \geq \chi^2_{\alpha,1}$, zamítáme hypotézu shody, v opačném případě k tomu není důvod.

A2) *Binomický (znaménkový) test*, (pro libovolné n; vždy pro $n < 30$)

a) zvolíme hladinu významnosti;

b) určíme $S = \min(n_{12}, n_{21})$ a tuto hodnotu porovnáme s kritickou hodnotou S_α binomického (znaménkového) dvoustranného testu[7]) s počtem pozorování $\bar{n} = n_{12} + n_{21}$.

Druhá možnost spočívá v tom, že v tabulkách binomického rozložení s $p = 0.5$ nalezneme hodnotu distribuční funkce pro $S = \min(n_{12}, n_{21})$ a tuto hodnotu násobíme dvěma. Je-li výsledné číslo $\alpha^* \leq \alpha$, pak zamítáme hypotézu shody. Pro $\alpha^* > \alpha$ není důvodu (na zvolené hladině) shodu nepřijmout. Pro případ $n_{12} = n_{21}$ automaticky přijímáme $\boldsymbol{H}_0(\alpha^* = 1.0)$.

B) *Vícehodnotový znak* $(K > 2)$

B1) *Stuartův test* $(n \geq 30)$ (A. Stuart [1965])

a) zvolíme hladinu významnosti α a k ní nalezneme kritickou hodnotu $\chi^2_{\alpha,K-1}$ pro test chí-kvadrát s $(K - 1)$ stupni volnosti.

b) určíme hodnoty matice $\mathbf{V} = \|V_{jk}\|$

$$(2) \qquad V_{kk} = n_{k+} + n_{+k} - 2n_{kk} \quad \text{pro} \quad k = 1, 2, \ldots, K-1$$
$$V_{jk} = -(n_{jk} + n_{kj}) \quad \text{pro} \quad j \neq k = 1, 2, \ldots, K-1$$

a vektor $\boldsymbol{d}$ o složkách

$$(3) \qquad d_k = n_{k+} - n_{+k}, \quad k = 1, 2, \ldots, K-1$$

c) zjistíme inverzní matici $\mathbf{V}^{-1}$ a hodnotu

$$(4) \qquad Q_S = \boldsymbol{d}' \, \mathbf{V}^{-1} \, \boldsymbol{d}$$

d) Je-li dosažená významnost α^* pro Q_S vzhledem k rozložení χ^2 s $df = K - 1$ menší nebo rovna α

(5) $\alpha^* \leq \alpha$

nebo (ekvivalentně)

(6) $Q_S \geq \chi_\alpha^2,\ _{K-1}$,
zamítneme $\boldsymbol{H}_0$ ve prospěch $\boldsymbol{H}_A$.
V opačném případě není důvodu (na hladině α) zamítat hypotézu $\boldsymbol{H}_0$.

B2) *Bhapkarův test* ($n \geq 30$) (Bhapkar [1966])

Postup je zcela obdobný k B1) jen místo matice $\mathbf{V}$ vytváříme matici $\mathrm{W} = \|W_{jk}\|$ o rozměrech $(K-1) \times (K-1)$ a prvky:

(7) $W_{kk} = n_{k+} + n_{+k} - 2n_{kk} - n^{-1}(n_{k+} - n_{+k})^2$, $k = 1, 2, \ldots, K-1$
$W_{jk} = -(n_{jk} + n_{kj}) - n^{-1}(n_{k+} - n_{+k})(n_{j+} - n_{+j})$, $j \neq k = 1, 2, \ldots, K-1$

a statistiku Q_B jako

(8) $Q_B = \boldsymbol{d}'\ W^{-1}\ \boldsymbol{d}$.

B3) *Test minimální diskriminační informace* ($n \geq 30$) (Ireland, Ku, Kullback [1969])

a) Určíme kritickou hladinu α a k ní kritickou hodnotu
$\chi^2_{\alpha,K-1}$ (obdobně jako v A).

b) Testová statistika se zjistí jako

(9) $$Z = 2n \sum_{j=1}^{K} \sum_{k=1}^{K} r_{jk} \ln \frac{r_{jk}}{f_{jk}},$$

kde r_{jk} se získá iteračním postupem

(10) $$r_{jk}^{(0)} = f_{jk}, \qquad r_{jk}^{(m+1)} = \left[\frac{r_{+j}^{(m)}\, r_{k+}^{(m)}}{r_{j+}^{(m)}\, r_{+k}^{(m)}}\right]^{\frac{1}{2}} r_{jk}^{(m)}\, C_m,$$

kde

$$C_m = 1 \Big/ \sum_{j=1}^{k} \sum_{k=1}^{k} \left[\frac{r_{+j}^{(m)}\, r_{k+}^{(m)}}{r_{j+}^{(m)}\, r_{+k}^{(m)}}\right]^{\frac{1}{2}} r_{jk}^{(m)},$$

(11) $$r_{+j}^{(m)} = \sum_{k}^{K} r_{kj}^{(m)}, \qquad r_{k+}^{(m)} = \sum_{j}^{K} r_{kj}^{(m)}$$

C) *Simultánní aplikace* $(K-1)$ *testů pro rovnosti četností u jednotlivých kategorií* (*J. Řehák—B. Řeháková [1984]*)

a) určíme kritickou hladinu α a k ní hodnotu $\alpha/(K-1)$

b) zjistíme statistiky McNemarova testu[9]) pro $\overline{n}_k = n_{k+} + n_{+k} - 2n_{kk} \geq 30$

(12) $$Q_k = \frac{(n_{k+} - n_{+k})^2}{n_{k+} + n_{+k} - 2n_{kk}}, \qquad k = 1, 2, \ldots, K$$

Pro $\overline{n}_k < 30$ určíme významnost binomického (znaménkového) testu (též i pro $\overline{n}_k \geq 30$, jsou-li k dispozici tabulky):
určíme $S_k = \min(n_{k+} - n_{kk}, n_{+k} - n_{kk})$ a pro ně zjistíme $\alpha_k^* = 2P_B(\boldsymbol{X} \leq S_k)/p = 0.5, \overline{n}_k)$ podle tabulek distribuční funkce binomického rozložení (pro $S_k = n_k/2$ položíme $\alpha_k^* = 1.0$). Dosažené hladiny ≤ 0.10 pro $n_k \leq 30$ viz tab. B.

c) Je-li alespoň jedno z Q_k větší či rovno hodnotě $\chi^2_{\alpha/(K-1),1}$ pro kategorie u nichž, je $\overline{n}_k \geq 30$ nebo alespoň jedna dosažená významnost binomických testů (popřípadě i McNemarových testů, pokud jsou známy jejich α^*) je menší či rovna hodnotě $\alpha/(K-1)$, pak zamítáme hypotézu shody. Hodnoty $\chi^2_{\alpha/(K-1),1}$ nalezneme v tab. A.

Poznámky k metodám:

1. Každá z metod má své uplatnění

a) Pro $n \geq 30$ a rutinní analýzu dat doporučujeme Stuartův test, který vychází důsledně z platnosti $\boldsymbol{H}_0$.

b) Pro $n < 30$ je vhodné použít simultánní sérii binomických testů (viz C).

c) Metoda simultánního využití McNemarových testů je výhodná pro ruční výpočty, nemáme-li k dispozici počítač. V praxi není nutno počítat všechna Q_k, ale stačí odhadnout, která budou patřit mezi největší. Hodnoty potřebné k testu (tj. kritické hodnota $\chi^2_{\alpha/(k-1),1}$) jsou uvedeny v tab. A.

d) Bhapkarův test byl zaveden jako vylepšení testu Stuartova. Vzhledem k tomu, že však nevychází z platnosti $\boldsymbol{H}_0$ a má tak jiné zázemí, je jeho statistika vhodná pro hledání simultánních konfidenčních intervalů pro složky vektoru rozdílů $(p_{k+} - p_{+k})$. Platí, že

(13) $\quad Q_B = Q_S(1 - Q_S/n)\,, \quad \mathbf{W} = \mathbf{V} - dd'/m$

Konečné slovo o tom, která ze statistik Q_B, Q_S je vhodnější, může dát buď další teoretické zhodnocení, či spíše extenzívní simulační studie.

e) Test minimální diskriminační informace je možno použít jen ve spojení s výpočetní technikou. Jeho význam se projeví ve spojení s dalšími úlohami ve čtvercových tabulkách.

2. Test shody lze založit též na statistice $\sum_j^K (f_{+j} - f_{j+})^2$ (čtverec eukleidovské vzdálenosti distribucí), jejíž významnost lze zjistit aplikací algoritmu *AS* 106 autorů Sheil, O'Muircheartaigh [1977].

3. U McNemarových testů lze též aplikovat korekce pro spojitost, záměnou čitatele v (1) a (12) za

(14) $\quad (|n_{k+} - n_{+k}| - 1/2)^2$

4. Pro $K = 2$ přichází Stuartův test (4) v McNemarův test (1) bez korekcí (14).

Krok 2: *Koeficient neshody*

Stupeň neshody dvou marginálních distribucí čtvercových tabulek můžeme měřit pomocí

$$(15) \quad C_{nom} = \sqrt{\frac{\sum_{k=1}^{K} (f_{k+} - f_{+k})^2}{2}}$$

Vlastnosti koeficientu:

a) je vždy definován

b) $0 \leq C_{nom} \leq 1$

c) $C_{nom} = 0$ právě když empirické četnosti zcela splývají, $f_{k+} = f_{+k}$ pro $k = 1, 2, \ldots, K$

d) $C_{nom} = 1$ právě když obě distribuce jsou si maximálně nepodobné ve smyslu soustředění všech pozorování do rozdílných kategorií, tj. $f_{k+} = 1, f_{+j} = 1, k \neq j$

Směrodatná chyba koeficientu C_{nom} je dána vzorcem:

$$(16) \quad s_c = \frac{1}{\sqrt{n}} \frac{1}{2C_{nom}} \{\sum\sum_{r<s} (f_{rs} + f_{sr})(q_r - q_s)^2 - [\sum\sum_{r<s} (f_{rs} - f_{sr})(q_r - q_s)]^2\}^{\frac{1}{2}},$$

kde $q_r = f_{r+} - f_{+r}$

a *100 (1 − α) %ní interval spolehlivosti* je určen jako

(17) $C_{nom} \pm Z_\alpha s_c$,

kde Z_α je kritická hodnota dvoustranného standardního normálního testu.[10])

Krok 3: Určení kategorií s rozdílnými četnostmi

Úlohu řešíme pomocí simultánního prověření všech K hypotéz o dvojicích četností

$$\boldsymbol{H}_{0k} : p_{k+} = p_{+k} \quad \text{proti} \quad \boldsymbol{H}_{Ak} : p_{k+} \neq p_{+k}, \quad k = 1, 2, \ldots, K.$$

Testování provedeme buď porovnáním statistik Q_k s postupnými kritickými hodnotami,[11] nebo porovnáním dosažených hladin významnosti α_k^* s postupnými hladinami významnosti α/m $(m = 1, 2, \ldots, K)$.[12])

1. *Použití statistik* Q_k (α je předem určená hladina významnosti): do tabulky sestavíme K hodnot Q_k od nejvyšší po nejnižší (značíme $Q_{(k)}$) a pod ně vypíšeme K postupných kritických hladin $\chi^2_{\alpha/m,1}$, $m = K, K-1, \ldots, 2, 1$ podle tabulky A:

$$Q_{(1)}, Q_{(2)}, Q_{(3)}, \ldots, Q_{(K)}$$
$$\chi^2_{\alpha/K,1}, \quad \chi^2_{\alpha/(K-1),1}, \quad \chi^2_{\alpha/(K-2),1}, \ldots, \chi^2_{\alpha,1}$$

Významné budou všechny rozdíly četností, které odpovídají těm $Q_{(1)}, Q_{(2)}, \ldots, Q_{(M)}$, která jsou v *nepřerušené* řadě větší, nebo rovny příslušným kritickým hodnotám. Jakmile je první $Q_{(M+1)}$ menší než k ní přiřazená hodnota, jsou *ona a všechny následující* statistiky $Q_{(M+2)}, \ldots, Q_{(K)}$ nevýznamné; u kategorií, které odpovídají prvním M statistikám, konstatujeme platnost alternativní hypotézy a u zbylých $K-M$ kategorií platnost nulové hypotézy.

2. *Použití dosažených hladin významnosti* α_k^*

Vytvoříme tabulku, v níž seřadíme α_k^* od nejvyšší k nejnižší a pod ně vypíšeme postupné hladiny významnosti:

$$\alpha_{(1)}^*, \alpha_{(2)}^*, \alpha_{(3)}^*, \ldots, \alpha_{(K)}^*$$
$$\alpha/K, \quad \alpha/(K-1), \alpha/(K-2), \ldots, \alpha$$

Pro *nepřerušenou* řadu vztahů $\alpha^*_{(1)} \leq \frac{\alpha}{K}, \alpha^*_{(2)} \leq \frac{\alpha}{(K-1)}, \ldots, \alpha^*_{(M)} \leq \frac{\alpha}{(K-M+1)}$

jsou všechny příslušné rozdíly významné a četnosti jim odpovídající považujeme za různé. Jakmile však první $\alpha^*_{(M+1)}$ je větší než její příslušná kritická hladina, považujeme *k ní* odpovídající četnosti a *všechny zbývající* za stejné.

Krok 4: Znaménkové schéma rozdílů četností

Znaménkové schéma slouží v této úloze jako indikace stupně významnosti rozdílu a rychlá grafická orientace na ty kategorie, u nichž je rozdíl nejprokázanější. Provedeme jej tak, že každému rozdílu $f_{+k} - f_{k+}$, $k = 1, 2, \ldots, K$ přiřadíme znaménko podle pravidla:[13])

0	nebyl prokázán rozdíl
+ nebo −	byl prokázán rozdíl na hladině $\alpha_1 = 0.05$
++ nebo −−	byl prokázán rozdíl na hladině $\alpha_2 = 0.01$
+++ nebo −−−	byl prokázán rozdíl na hladině $\alpha_3 = 0.001$

Kladná znaménka značí vyšší obsazení sloupcové kategorie $f_{+k} - f_{k+} > 0$, a záporná naopak.

Určování znaménkového schématu provádíme trojí aplikací Holmovy metody pro vyhledávání významných rozdílů:

1. na hladině $\alpha_1 = 0.05$ určíme nuly a + nebo −,

2. u těch kategorií, u nichž zůstává významnost na $\alpha_2 = 0.01$, připíšeme jedno znaménko + nebo —,
3. u těch kategorií, které zůstanou významné pro $\alpha_3 = 0.001$ připíšeme ještě jedno znaménko + nebo —.

3. Příklady analýzy shody pro nominální data

V této části uvádíme příklady z výzkumné praxe, u nichž lze metodiku využít. Zároveň ukážeme i numerické postupné řešení pro případnou kontrolu.[14])

Příklad: *Preference kupního záměru*

U $n = 269$ manželských dvojic s nezařízeným bytem byla položena tatáž otázka manželovi i manželce: „*Který z pěti předmětů si chcete koupit jako první?*" Odpovědi: 1 ≡ *televize*, 2 ≡ *neautomatická pračka*, 3 ≡ *ledničku*, 4 ≡ *automatická pračka*, 5 ≡ *kuchyňské vybavení*.
Statistické porovnání dvou obdržených empirických distribucí
A ≡ odpověď muže: (5.3 %, 23.9 %, 38.3 %, 28.4 %, 4.2 %)
B ≡ odpověď ženy: (3.4 %, 9.8 %, 48.9 %. 23.9 %, 14.0 %)
provedeme postupem, který vychází ze čtvercové tabulky **A** × **B** — viz tab. 2.

Tabulka 2. Záměry koupě prvního předmětu vybavení domácnosti: porovnání názoru manžela a manželky u nevybavených domácností (absolutní četnosti a celková procenta)

Názor manžela		Názor manželky: televize	neautomatická pračka	lednička	automatická pračka	kuchyňské vybavení	Součet	Procento
	televize	3 (1.1)	1 (0.4)	6 (2.3)	2 (0.8)	2 (0.8)	14	5.3
	neautomatická pračka	2 (0.8)	11 (4.2)	39 (14.8)	7 (2.7)	4 (1.5)	63	23.9
	lednička	2 (0.8)	12 (4.5)	57 (21.6)	24 (9.1)	6 (2.3)	101	38.3
	automatická pračka	1 (0.4)	2 (0.8)	27 (10.2)	28 (10.6)	17 (6.4)	75	28.4
	kuchyňské vybavení	1 (0.4)	0 (0.0)	0 (0.0)	2 (0.8)	8 (3.0)	11	4.2
	Součet	9	26	129	63	37	264	
	Procento	3.4	9.8	48.9	23.9	14.0		

Přímé porovnání četností je dáno rozdílem četností mezi **B** a **A**: (—1.9; —14.1; 10.6; —4,5; 9.8) a naznačuje rozdíly. Statistické testování hypotézy H_0: „*rozložení názoru mužů ≡ rozložení názoru žen*" provedeme Stuartovým testem.
Stuartův test pro tab. 2 (pro hladinu významnosti $\alpha = 0.05$):
a) z výpočtů vynecháme poslední kategorii;
b) rozdíly četnosti $\mathbf{d} = (5, 37, -28, 12)$ pro první čtyři kategorie (viz (3));
c) matice $\mathbf{V} = \|V_{ij}\|$ typu 4×4 bude mít prvky (podle (2)):

$$V_{11} = 14 + 9 - 2 \times 3 = 17\,, \quad V_{12} = -(2 + 1) = -3\,, \quad \text{atd.}$$

$$\mathbf{V} = \begin{Vmatrix} 17 & -3 & -8 & -3 \\ -3 & 67 & -51 & -9 \\ -8 & -51 & 116 & -51 \\ -3 & -9 & -51 & 82 \end{Vmatrix}$$

Inverzní matice[15]) $\mathbf{V}^{-1}$ (uvádíme její stonásobek):

$$100 \times \mathbf{V}^{-1} = \begin{Vmatrix} 8.1988 & 2.9143 & 2.9168 & 2.4340 \\ 2.9143 & 4.6068 & 3.4348 & 2.7485 \\ 2.9168 & 3.4348 & 3.8345 & 2.8686 \\ 2.4340 & 2.7485 & 2.8686 & 3.3943 \end{Vmatrix}$$

d) výpočet testové statistiky

$$Q_S = \frac{1}{100}\,(5 \times 5 \times 8.1988 + 5 \times 37 \times 2.9168 + \ldots) = 39.5648\,.$$

e) Pro df = 4 je kritická hladina[16]) pro $\alpha = 0.05$ rovna $\chi^2_{0.05,4} = 9.4877$
$Q_S > 9.49$, a tudíž zamítáme hypotézu $\boldsymbol{H_0}$.
obdobně: dosažená hladina významnosti $\alpha^* = .000\,000\,053\,29$, což je menší než .00005 a což zapisujeme počítačem jako $\alpha^* = 0.0000$. Protože $\alpha^* \leqq \alpha$, zamítáme $\boldsymbol{H_0}$.
Obdobný výpočet Bhapkarova testu dává $Q_B = 46.5395$ s dosaženou hladinou významnosti $\alpha^* = 0.0000$ (přesně 1.99×10^{-9}).
Při ručních výpočtech můžeme vyjít z dílčích čtyřpolních tabulek, na něž aplikujeme postupně McNemarovy testy (viz (12)).Bude-li kterýkoli z nich významný na hladině $0.05/4 = 0.0125$, bude prokázána rozdílnost obou distribucí na $\alpha = 0.05$. Postup probíhá podle tab. 3.

Tabulka 3. Čtyřpolní tabulky vzniklé postupnou souběžnou dichotomizací znaků **A** *a* **B** *z tab. 2*
Uvedeny jsou statistiky Q_k a $Z_k = \sqrt{Q_K}$ společně se znaménkem ukazujícím na směr rozdílu („+" = sloupcová kategorie, má vyšší obsazení)

$k = 1$

3	11	14
6	244	250
9	255	264

$\alpha_1^* = 0.3323$ (—)
(binomický test)

$k = 2$

11	52	63
15	186	201
26	238	264

$Q_2 = 20.4328$
$Z_2 = 4.5203$ (—)
$\alpha_2^* < 0.00001$

$k = 3$

57	44	101
72	91	163
129	135	264

$Q_3 = 6.7586$
$Z_3 = 2.5997$ (+)
$\alpha_3^* = 0.0093$

$k = 4$

28	47	75
35	154	189
63	201	264

$Q_4 = 1.7561$
$Z_4 = 1.3252$ (—)
$\alpha_4^* = 0.1868$

$k = 5$

8	3	11
29	224	253
37	227	264

$Q_5 = 21.1250$
$Z_5 = 4.5962$ (+)
$\alpha_5^* < 0.00001$

Poznámky: 1. Q_k jsou počítány bez korekcí na spojitost.
2. Dosažené hladiny významnosti pro $k = 2, 3, 4, 5$ byly nalezeny v tabulkách J. Likeš—J. Laga [1978 : tab 1A]. Viz též J. Řehák—B. Řehákóvá [1984 : tabC]
3. Viz tab. B pro $\alpha_1^* = .3323$.

U $k = 1$ není splněna podmínka pro McNemarův test, proto byla dosažená významnost zjištěna pomocí binomického testu (viz tab. 3).
V tabulce je vidět, že tři z pěti rozdílů mají $\alpha^* \leqq 0.0125$, a proto i tento postup vede k zamítnutí hypotézy shody obou názorových distribucí. (K závěru stačilo spočítat pouze jeden z testů, který je významný na $\alpha/4$).
Podrobnou informaci z tab. 3 využijeme nyní pro určení těch kategorií, u nichž se vyskytne významný rozdíl. K tomu použijeme Holmovu metodu postupných hladin významnosti. Pro $\alpha = 0.05$ lze postup shrnout:

k	5	2	3	4	1
α^*	< .00001	< .00001	.0093	.1868	.3323
postupná hladina	.0100	.00125	.0167	.0250	.0500
významnost a směr	+	—	+	0	0

Dosažená významnost je menší než příslušná hranice α/m, u $l = 5, 2, 3$ ($m = 5, 4, 3$), u nichž považujeme rozdíly za prokázané, zatímco u $l = 4{,}1$ ($m = 2, 1$) prokázány rozdíly nejsou.

Pro získání znaménkového schématu opakujeme postup pro $\alpha = 0.01$ a $\alpha = 0.001$

postupná hladina	.0020	.0025	.0033	.0050	.0100
významnost	+ +	— —	+	0	0
postupná hladina	.00020	,00025	.00033	.00050	.00100
významnost	+ + +	— — —	+	0	0

Výsledek lze sumarizovat přehledně v tabulce 4.

Tabulka 4. Rozdílnost v názorech na preferenční vybavení domácností u mladých manželů (procenta rozložení a statistická významnost rozdílů)

	Záměr koupě prvního předmětu v %				
	televize	neautomatická pračka	lednička	automatická pračka	kuchyňské vybavení
názory žen	3.4	9.8	48.9	23.9	14.0
názory mužů	5.3	23.9	38.3	28.4	4.2
rozdíl	—1.9	—14.1	10.6	—4.5	10.0
významnost	0	— — —	+	0	+ + +

Poznámky: a) „+" znamená vyšší požadavek u žen
„—" znamená vyšší požadavek u mužů
b) „+" značí významnost na 0.05
„+ + +", „— — —" značí významnost na 0.001

Koeficient neshody C_{nom} (viz (15)) spočteme jako:[17]

$$C_{nom} = \sqrt{\tfrac{1}{2}((.034 - .053)^2 + (.098 - .239)^2 + \ldots + (.140 - .042)^2)} = 0.1468\,.$$

Výsledkem statistické analýzy je tedy konstatování, že názory mužů a žen nejsou shodné. Muži více preferují neautomatickou pračku, ženy ledničku a kuchyňské vybavení (soupravu pro vaření a stolování, příbory, kuchyňský robot).[18] Při další analýze je možno hledat specifické faktory neshody názorů pomocí podobné analýzy pro dílčí soubory (třídění třetího a vyšších stupňů) určené např. hodnotovými orientacemi mladé dvojice, typologií vzájemných vztahů a typologií životního způsobu rodin, z nichž mladí manželé přicházejí, vzdělanostní nebo profesní strukturou. Statistická příčina by se projevila vymizením neshody v dílčích souborech,[19] určených jejími hodnotami.

Příklad: Preferenční postoje k opatřením

Ve velkém podniku bylo dotazováno 53 řídících pracovníků různých úrovní (náhodný výběr) na preference v opatřeních vzhledem k situaci v podniku a v řízeném kolektivu. Dvě otázky zněly: **A** ≡ *„Dáváte přednost výstavbě nové kapacity, nebo modernizaci techniky?"* **B** ≡ *„Dáváte přednost výstavbě nové kapacity, nebo investicím do sociálních opatření?"* Výzkumná otázka: lze obě preferenční distribuce považovat za shodné, nebo se projevuje jejich statisticky významný rozdíl? Statistický model: $\boldsymbol{H_0}$ ≡ *„obě distribuce jsou shodné"* vs $\boldsymbol{H_A}$ ≡ *distribuce se liší alespoň u jedné ze tří kategorií škály"*, testujeme hypotézu $\boldsymbol{H_0}$ proti hypotéze $\boldsymbol{H_A}$.

Data jsou uvedena v tab. 5a, test provedeme pomocí Stuartovy statistiky Q_S. Výsledky postupu viz tab. 5b.

Tabulka 5. Porovnání preferenčních názorů na investice (soubor 53 řídících pracovníků)
a) Třídění 2. stupně pro dva názory

		preference			
		nová kapacita	nemá preferenci	sociální opatření	
preference	nová kapacita	13	3	6	22
	nemá preferenci	2	8	8	18
	nová technika	2	2	9	13
		17	13	23	53

b) Komparace distribucí

Preferenční rozložení četností	—	0	+
nová kapacita — sociální opatření	32.1 %	24.5 %	43.3 %
nová kapacita — nová technika	41.5 %	34.0 %	24.5 %
Rozdíl v procentech	—9.4 %	—9.5 %	18.9 %

Stuartova statistika $Q_S = 5.5882$, df $= 2$, $\alpha^* = 0.0612$, $C_{nom} = 0.1634$.

Dosažená významnost $\alpha^* = 0.06$ je na hranici běžného rizika $\alpha = 0.05$. Přesto, že formální statistické rozhodnutí ($\alpha^* > \alpha$) vede k přijetí hypotézy shody, provedeme explorační posouzení rozdílů kategorií, neboť smysluplná věcná interpretace, která je navíc popřípadě podpořena jinými věcnými výsledky, se přesto může stát součástí závěrů.

Další výzkumná otázka tedy zní: naznačují data hypotézu rozdílu četnosti v některé z kategorií (a tím závěr o tendenci v názorech)?

Četnosti v tabulce jsou nízké, proto použijeme binomický test.[20]) Postup:

kategorie	$n_{k+} - n_{kk}$	$n_{+k} - n_{kk}$	min	$\bar{n}$	tabulková hodnota	dosažená významnost $\alpha_k{}^*$
1	9	4	4	13	.1334	.2668
2	10	5	5	15	.1509	.3018
3	14	4	4	18	.0154	.0309

I když je $\alpha_3{}^* = .0309 < 0.05$, nevypovídá tento vztah o rozdílnosti celé distribuce,[21]) neboť podle Holmovy metody postupných hladin musí být pro odmítnutí $\boldsymbol{H_0}$: min $(\alpha_1, \alpha_2, \alpha_3) \leqq \leqq 0.05/2 = 0.025$. I zde však je výsledek na hranici. Ve skutečnosti také byla přijata hypotéza $\boldsymbol{H_A}$ pro její věcný význam. Příklad zároveň ukazuje výhodu výpočtů dosažených hladin významnosti, které umožní daleko lépe posoudit, do jaké míry jsou $\alpha_k{}^*$ rozdílné od α (které je vždy voleno libovolně), a tím posoudit rizika spojená s vytvářením věcných závěrů.

V daném příkladě se ovšem také nabízí otázka metodologická, zda testování hypotéz o minimálních distribucích neztrácí na informaci, neboť u ordinálních znaků může být důležitější alternativa rozdílnosti tvarů distribucí nebo posunu mediánů.

Při rozboru příkladů jsme prováděli zdlouhavé výpočty. V rutinní analýze dat však z počítače budeme požadovat pouze stručný výsledek ve tvaru tabulky 4 nebo 5b.

4. Statistická metodika porovnání párově závislých distribucí - ordinální znaky

U ordinálních znaků můžeme formulovat dvě úlohy, které se svým významem liší, ale obě vycházejí z toho, že kategorie škály jsou uspořádány. Je to porovnání tvarů

distribucí, testování jejich shody a posouzení posunu mediánů (a tím i posunu rozložení četností) u jednoho ze znaků proti druhému. Účelem této části je uvést nové postupy pro porovnání distribučních funkcí a adaptaci Wilcoxonova testu pro závislé výběry na situaci kategorizovaných dat.

Testy pro porovnání distribučních funkcí (tj. tvarů rozložení četností) zavádíme jako analogie k testům Stuartovu a Bhapkarovu z části 2. Vycházíme přitom z výsledků práce J. Řehák, B. Řeháková [1979], z níž plyne tvar distanční funkce stojící v základu testů. Pro postup si zavedeme další značení, které doplňuje symboliku části 2:

$$F_{+k} = (\sum_{s=1}^{k} n_{+s})/n\,, \qquad F_{jk} = (\sum_{r=1}^{j}\sum_{s=1}^{k} n_{rs})/n$$

$$F_{k+} = (\sum_{k=1}^{r} n_{r+})/n\,,$$

jsou empirické distribuční funkce; jejich populační hodnoty jsou

$$P_{k+}, P_{+k}, P_{jk}\,.$$

Kroky analýzy jsou obdobné jako v části 2 o nominálních datech; testování rozdílnosti kumulačních funkcí v polích (krok 3) má však velmi specifický význam a totéž platí o znaménkovém schématu, které získáme (obdobně jako u nominálních znaků) trojnásobným opakováním kroku 3.

Krok 1: *Testy shody závislých distribučních funkcí* (J. Řehák, B. Řeháková [1984])

Formulace hypotéz: $\boldsymbol{H_0} \equiv P_{k+} = P_{+k}$ pro $k = 1, 2, \ldots, K-1$

$\boldsymbol{H_A}$ = alespoň pro jedno k je $P_{k+} \neq P_{+k}$

Data: Čtvercová tabulka druhého stupně třídění obou znaků **A** a **B** (viz tab 1a)

Intuitivní model: Neplatnost $\boldsymbol{H_0}$ bude indikována velkými rozdíly některých výběrových hodnot $F_{k+} - F_{+k}$, $k = 1, 2, \ldots, K-1$

Hladina významnosti: α (určena předem).

Postupy testování H_A *vs* H_0

A) *Přímý test shody* (analogie k testu Stuartovu), $n \geq 30$

a) určíme vektor $\boldsymbol{D}$ o složkách $D_k = n(F_{k+} - F_{+k})$, $k = 1, 2, \ldots, K-1$,

b) určíme prvky matice $\mathbf{V} = \|V_{jk}\|$ o rozměru $(K-1) \times (K-1)$:

(18) $V_{kk} = n(F_{k+} + F_{+k} - 2F_{kk})$

$V_{jk} = n(F_{j+} + F_{+j} - F_{jk} - F_{kj})$ pro $j < k$

$V_{jk} = V_{kj}$ pro $j > k$

k $\mathbf{V}$ určíme inverzní matici $\mathbf{V}^{-1}$,

c) určíme testovou statistiku

(19) $Q_H = \boldsymbol{D}'\,\mathbf{V}^{-1}\,\boldsymbol{D}$,

d) je-li $Q_H \geq \chi^2_{\alpha, K-1}$ nebo α^* pro rozložení chí-kvadrát s $df = K-1$ je $\alpha^* \leq \alpha$, zamítáme $\boldsymbol{H_0}$ ve prospěch $\boldsymbol{H_A}$ (obě nerovnosti jsou ekvivalentní). V opačném případě není důvod $\boldsymbol{H_0}$ zamítat.

B) *Přímý test shody* (analogie k Bhapkarovu testu), $n \geq 30$

Postupujeme obdobně jako u A); místo Q_H však spočítáme

(20) $Q_P = \boldsymbol{D}'\,\mathbf{W}^{-1}\,\boldsymbol{D}$,

kde prvky matice $\mathbf{W} = \|W_{jk}\|$ jsou dány jako:

$$W_{kk} = n[F_{k+} + F_{+k} - 2F_{kk} - (F_{k+} - F_{+k})^2]]$$

$$(21) \quad W_{jk} = n[F_{j+} + F_{+j} - F_{jk} - F_{kj} - (F_{j+} - F_{+j})(F_{k+} - F_{+k})]$$

pro $j < k$

$$W_{jk} = W_{kj} \quad \text{pro} \quad j > k$$

C) *Simultánní aplikace* $(K - 1)$ *testů pro rovnosti jednotlivých kumulovaných četností v kategoriích*

Test se provádí obdobně jako v postupu E) u metodiky pro nominální znaky (viz část 2). Místo (12) však zjišťujeme statistiky:

$$(22) \quad Q_k = n \frac{(F_{k+} - F_{+k})^2}{F_{k+} + F_{+k} - 2F_{kk}}, \qquad k = 1, 2, \ldots, K - 1$$

Pro $\bar{n}_k = n(F_{k+} + F_{+k} - 2F_{kk}) \leq 30$ aplikujeme (též zcela obdobně) binomické testy pro dvojice četnosti

$$(23) \quad n(F_{k+} - F_{kk}), \quad n(F_{+k} - F_{kk}).$$

Poznámky k metodě:

1. a) Pro použití rutinních výpočtů počítačem doporučujeme aplikaci statistiky Q_H, která vychází z nulové hypotézy ($\boldsymbol{H}_0$), za jejichž předpokladů je odvozena. Pro ruční výpočty je tato metoda vhodná jen pro malá K (jednoduchý je postup pro $K = 3$; pro větší K je možno využít programovatelné kalkulačky); je však možno též využít i postup C.

b) Obdobně k (13) platí:

$$(24) \quad Q_P = Q_H/(1 - Q_H/n), \qquad \mathbf{W} = \mathbf{V} - D\,D'/n$$

2. Test shody lze založit též na statistice $\sum_{k}^{K-1} (F_{k+} - F_{+k})^2$, jejíž významnost lze zjistit aplikací algoritmu *AS* 106 (viz Sheil, O'Muircheartaigh [1977]).

3. a) U McNemarových testů lze též aplikovat korekce pro spojitost, které znamenají záměnu čitatele $n^2(F_{k+} - F_{+k})^2$ v upraveném vzorci (22) za

$$(15) \quad (n|F_{k+} - F_{+k}| - \tfrac{1}{2})^2$$

b) Lze též využít normální test $Z_k = \sqrt{Q_k}$, viz poznámka 11.

c) Pro $K = 3$ je postup ekvivalentní k aplikaci dvou statistik Q_1 a Q_3 podle (12).

4. Pro $K = 2$ jsou situace pro ordinální a nominální znak totožné.

5. Pro výpočty je výhodné použít vzorce s dosazením empirických absolutních kumulovaných četností: $M_{k+} = nF_{k+}$, $M_{+k} = nF_{+k}$, $M_{jk} = nF_{jk}$.

Krok 2: Koeficient neshody pro ordinální závislá data

$$(26) \quad C_{ord} = \sqrt{\frac{\Sigma(F_{k+} - F_{+k})^2}{K - 1}}$$

Má vlastnosti:

a) je vždy definován.

b) $0 \leq C_{ord} \leq 1$,

c) $C_{ord} = 0$, právě když obě empirické distribuce splynou, $F_{k+} = F_{+k}$ pro všechna k, což je ekvivalentní k $f_{k+} = f_{+k}$ pro všechna k. ($C_{ord} = 0$ právě když $C_{nom} = 0$).
d) $C_{ord} = 1$ právě když $f_{1+} = 1$ a $f_{+K} = 1$ nebo $f_{+1} = 1$ a $f_{K+} = 1$.
Koeficient má směrodatnou chybu

$$(27)\quad s_c = \frac{1}{\sqrt{n}} \frac{1}{(K-1)\,C_{ord}} \left\{ \sum\sum_{r<s} (f_{rs} + f_{sr}) \left[\sum_{j=r}^{s-1} (F_{j+} - F_{+j}) \right]^2 - \left[\sum\sum_{r<s} (f_{rs} - f_{sr}) \sum_{j=r}^{s-1} (F_{j+} - F_{+j}) \right]^2 \right\}^{\frac{1}{2}}$$

a 100 $(1-\alpha)$%ní interval spolehlivosti pro populační koeficient je

$$(28)\quad C_{ord} \pm Z_\alpha\, s_c\,,$$

kde Z_α je kritická hodnota dvoustranného standardního normálu testu (viz poznámka 10).

Krok 3: Určení kategorií, u nichž se projevil významný rozdíl u kumulovaných četností a znaménkové schéma určující stupeň významnosti rozdílů

Postup se provádí postupnou Holmovou metodou za použití McNemarových nebo binomických testů (viz postup *C* této části). Na rozdíl od obdobné situace v části dvě zde posuzujeme pouze $(K-1)$ četnosti ($F_{K+} = F_{+K} = 1$), a proto postupné hladiny začínají od $\alpha/(K-1), \ldots, \alpha$. Aplikace však, vzhledem k významu rozdílů distribučních funkcí, bude požadována pouze ve speciálních situacích.

U ordinálních dat je, kromě porovnání tvaru distribucí, zajímavá hypotéza o posunutí mediánu. Tuto hypotézu prověříme modifikovaným Wilcoxonovým testem pro kategorizovaná závislá data. Úprava vychází z prací J. W. Pratt [1959] a E. Cureton [1967].

Test shody dvou mediánů u závislých marginálních rozložení ve čtvercové tabulce

Formulace hypotéz: $H_0 \equiv \tilde{X}_A = \tilde{X}_B$ (mediány jsou shodné)
$H_1 \equiv \tilde{X}_A \neq \tilde{X}_B$ (mediány jsou rozdílné)
$H_2 \equiv \tilde{X}_A < \tilde{X}_B$ (medián řádkového znaku je menší)
$H_3 \equiv \tilde{X}_A > \tilde{X}_B$ (medián řádkového znaku je větší)

Pro test vybereme předem jednu z alternativ a testujeme tak H_0 vs H_1 nebo H_0 vs H_2 nebo H_0 vs H_3.
Data: Čtvercová tabulka $\mathbf{A} \times \mathbf{B}$, $\|n_{jk}\|$, n = počet pozorování v tabulce (viz tabulka 1a).

Intuitivní model: Posunutí distribucí se projeví v soustředění dat do horní nebo do dolní části tabulky (nad hlavní nebo pod hlavní diagonálu). Čím větší posun mediánů, tím větší soustředění dat do jedné z polovin očekáváme.
Hladina významnosti: α (předem určeno).

A) *Modifikovaný Wilcoxonův test pro závislá kategorizovaná data (jeden výběr)*

a) značení a předběžné výpočty:
součty pozorování na diagonálách $= d_q = \sum_{j-i=q} n_{ij}$, $q = -(K-1), \ldots, -1, 0, 1, \ldots, K-1$

$$D_q = d_q + d_{-q}\,, \quad q = 1, 2, \ldots, K-1$$

(29) $R_1 = \frac{D_1 + 1}{2}$, $R_2 = \frac{2D_1 + D_2 + 1}{2}$, $R_3 = \frac{2D_1 + 2D_2 + D_3 + 1}{2}$,

$$R_q = \frac{2(D_1 + D_2 + \ldots + D_{q-1}) + D_q + 1}{2} \qquad q = 1, 2, \ldots, K-1$$

$$R = \sum_{q=1}^{K-1} d_q R_q + d_o C\,, \qquad C = \sum_{q=1}^{K-1} d_q\,, \qquad T = \sum_{q=1}^{K-1} (D_q^3 - D_q)$$

b) testová statistika je

(30) $$W = \frac{2R - [n(n+1) - d_o(d_o+1)]/2 \pm 1}{\sqrt{[n(n+1)(2n+1) - d_o(d_o+1)(2d_o+1) - \frac{T}{2}]/6}}$$

kde $\pm$ znamená redukci absolutní hodnoty čitatele o jedničku (tj. -1 pro kladný a $+1$ pro záporný výraz).

Závěr plyne z porovnání W a kritické hodnoty standardního normálního testu: Z_α (pro dvoustranný test, viz pozn. 10), Z_α' (pro jednostranný test[22]) nebo ze zjištěné hodnoty $\alpha' = 1 - \Phi(|W|)$, kde Φ je distribuční funkce standardního normálního rozložení.

Testovací[23]) úloha	H_0 se nezamítá	Přijímá se alternativní hypotéza
H_0 vs H_1 $(\tilde{X}_A \neq \tilde{X}_B)$	$\lvert W\rvert < Z_\alpha$ nebo $\alpha^* = 2\alpha' > \alpha$	$\lvert W\rvert \geq Z_\alpha$ $\alpha^* = 2\alpha' \leq \alpha$
H_0 vs H_2 $(\tilde{X}_A < \tilde{X}_B)$	$W < Z_\alpha'$ $\alpha^* = \alpha' > \alpha$	$W \geq Z_\alpha'$ $\alpha^* = \alpha' \leq \alpha$
H_0 vs H_3 $(\tilde{X}_A > \tilde{X}_B)$	$W > -Z_\alpha'$ $\alpha^* = \alpha' > \alpha$	$W \leq -Z_\alpha'$ $\alpha^* = \alpha' \leq \alpha$

B) *Modifikace binomického (znaménkového) testu*

a) zjistíme počet pozorování nad hlavní diagonálou

(31) $$C = \sum_i \sum_j n_{ij} = \sum_{q=1}^{K-1} d_q$$

a pod hlavní diagonálou

(32) $$D = \sum_{i<j} \sum n_{ij} = \sum_{q=1}^{K-1} d_{-q}$$

(33) $$\bar{n} = C + D\,,$$

b) četnosti C a D odpovídají četnostem „+" a „−" u znaménkového testu s $\bar{n} = C + D$, jehož variantu volíme podle volby alternativní hypotézy.[24]) Dosaženou významnost získáme z tabulky B.

Pro $n \geq 30$ lze použít a proximaci standardním normálním testem:

(34) $$Z = \frac{2C - \bar{n}}{\sqrt{\bar{n}}}\,.$$

přičemž rozhodnutí o hypotézách provádíme shodně s tabulkou Wilcoxonova testu (viz A).

Pro $\boldsymbol{H}_0$ vs $\boldsymbol{H}_1$ a $\bar{n} \geq 30$ lze použít McNemarovu statistiku:

$$(35) \quad X^2 = \frac{(C-D)^2}{C+D} \text{ popřípadě } X^2 = \frac{(|C-D| - \frac{1}{2})^2}{C+D},$$

kterou porovnáme s $\chi^2_{\alpha,1}$.

Poznámky k metodě:

1. Použitelnost Wilcoxonova testu se spojeními nebyla vzhledem k n v literatuře hodnocena; doporučujeme jej používat pro $n - d_0 \geq 30$.
2. Binomický test nevyužívá jemnou informaci o rozložení dat na diagonálách v tabulce, hodí se však pro rychlou informaci a ruční analýzu. Ve většině případů očekáváme stejné výsledky při jeho aplikaci jako při aplikaci Wilcoxonova testu, zvláště u malých rozměrů tabulek.

5. Příklady analýzy shody pro ordinální závislá data

U ordinálních znaků můžeme aplikovat všechny metody zavedené pro nominální případ (část 2). Můžeme však aplikovat i metody speciální, které vycházejí z vlastnosti uspořádanosti ordinální stupnice. Odstupňovaně se měnící škála umožňuje navíc zkoumat rozdílnost tvarů rozložení a posunutí mediánů.

Příklad: *Preferenční postoje k opatřením* (viz tabulka 5a)
Nominální analýza ukázala, že významnost rozdílu obou distribucí se pohybuje těsně u hranice přijaté hladiny. Dalším krokem bude proto nyní aplikace statistiky Q_H pro test hypotézy $\boldsymbol{H}_0 \equiv (P_{k+} = P_{+k}$, pro všechna $k)$ vs $\boldsymbol{H}_A \equiv$ (alespoň pro jedno k je $P_{k+} \neq P_{+k})$.
Výpočet statistiky Q_H:
a) $\mathbf{D} = \|5, 10\|$
b) Prvky matice $\mathbf{V}$ zjistíme z kumulativních četností, které dostaneme z tab. 6a ($M = nF$):

13	16	22	22
15	26	40	40
17	30	53	53
17	30	53	53

Prvky matice $\mathbf{V}$ a $\mathbf{V}^{-1}$ dostaneme:[25])

$$V_{11} = 22 + 17 - 2 \times 13 = 13, \qquad V_{22} = 40 + 30 - 2 \times 26 = 18$$
$$V_{12} = 22 + 17 - 16 - 15 = 8, \qquad V_{21} = V_{12} = 8$$

$$\mathbf{V} = \begin{Vmatrix} 13 & 8 \\ 8 & 18 \end{Vmatrix}, \qquad \mathbf{V}^{-1} = \begin{Vmatrix} .105\,882\,4 & -.047\,058\,8 \\ -.047\,058\,8 & .076\,470\,6 \end{Vmatrix}$$

Výpočet statistiky:

$$Q_H = 5 \times 5 \times 0.105\,882\,4 - 2 \times 5 \times 10 \times 0.047\,058\,8 + 10 \times 10 \times 0.076\,470\,6 = 5.588 < 5.991 \quad (=\chi^2_{0,05,2})$$

Podobně jako u nominálního přístupu ani zde nebyl prokázán signifikantní odklon od nulové hypotézy o shodě distribučních funkcí na hladině $\alpha = 0.05$.

Simultánní testování provedeme pomocí dvou binomických testů (4.9), (4.14), které mají významnosti 0.2668 a 0.0309, které jsou obě vyšší než 0.05/2 = 0.0250, a tudíž rozdíl není prokázán. (Použité testy se vyskytly i v nominálním případě).

Posun mediánů na stupnici znaků zde má zvláštní význam, neboť jde o specifickou alternativu, odpovídající rozdílnosti ve stupni názorové preference (aniž nás zajímal tvar rozložení). K tomu zjistíme C a D podle (31) a (32) a použijeme tab. B pro vyhledání binomického testu: významnosti binomického testu: $C = 3 + 6 + 8 = 17$, $D = 2 + 2 + 2 = 6$, $\alpha^* = 0.0347$ ($\leqq \alpha = 0.05$). Tak je posunutí prokázáno na hladině 0.05 (s významností 0.03): *„sociální opatření"* dostávají u řídících pracovníků vyšší stupeň názorové preference proti *„investicím do nových kapacit"* než *„investice do nové techniky"*.

Příklad: Porovnání názorů na řešení problému ve vyústění televizní inscenace

U manželských dvojic byly zjišťovány názory na vyústění televizní hry, která se zabývala problémy manželského soužití. Výzkumná otázka zněla: *„Liší se názory mužů od názorů žen?"* Tabulka 6 shrnuje data.

Tabulka 6. Názory manželských dvojic na řešení manželského problému v televizní hře (soubor 435 manželských dvojic, v tabulce jsou uvedeny absolutní četnosti a celkové procento)

		názor ženy			
		1	2	3	Celkem
názor muže	1	66 (15.2)	21 (4.8)	4 (0.9)	91 (20.9)
	2	33 (7.6)	150 (34.5)	8 (1.8)	191 (43.9)
	3	5 (1.1)	68 (15.6)	80 (18.4)	153 (35.2)
	Celkem	104 (23.9)	239 (54.9)	92 (21.1)	435 (100)

$Q_H = 44.47$, df = 2, $\alpha^* = 0.0000$
Kategorie odpovědí: 1 = nesouhlasím, řešení mělo být opačné
2 = řešení bylo správné jen zčásti
3 = souhlasím, řešení bylo správné

Četnosti v tabulce ukazují, že existuje značně vysoká souvislost mezi názory mužů a žen v rodinách a že tedy diferencující faktory působí na názory souběžně, že zde působí značně silný homogenizační faktor či že už vzájemný výběr partnerů je apriori podmíněn podobnými názory (Pearsonův lineární korelační koeficient $r = 0.65$ ukazuje na $100r^2 = 42$ % společné variance). Jsou však distribuce shodné? Nedochází k posunu tím, že zde působí specifické faktory pro muže a ženy či že souběžné faktory působí s různou intenzitou u různých pohlaví a tak dostaneme rozdíl ve stupni souhlasu s vyústěním děje? Stupeň souhlasu pro soubor můžeme charakterizovat mediánem; pro muže je to 2.16, pro ženy 1.97. Je však tento rozdíl vysvětlitelný náhodnými vlivy nebo vlivy systematickými? To prověříme testem hypotézy o shodě mediánů proti hypotéze o jejich rozdílnosti (dvoustrannou hypotézu volíme proto, že jsme předem nevěděli, který z mediánů by měl být větší.)

Znaménkový test:

$$C = 33\,, \quad D = 106\,, \quad \tilde{n} = 139\,, \quad Z = (66 - 139)/11.790 = -6.192$$

To ukazuje na vysokou významnost, přijímáme tedy hypotézu o rozdílnosti mediánů:

$$6.192 > 1.96 = Z_{0,05}$$

Modifikovaný Wilcoxonův test:

$$d_2 = 4, \quad d_1 = 29\,, \quad d_0 = 296\,, \quad d_{-1} = 101\,, \quad d_{-2} = 5$$
$$D_1 = 130\,, \quad D_2 = 9\,, \quad R_1 = 65.5\,, \quad R_2 = 135$$
$$T = 2197590\,, \quad R = 2439.5 + 9768 = 12207.5$$

$$W = \frac{2 \times 12207.5 - 50874 + 1}{\sqrt{(113062044 - 1098795)/6}} = -\frac{26458}{4319.78} = -6.1248$$

Závěr je zcela stejný jako u znaménkového testu, neboť $6.1248 > 1.96$. Prokázaný statistický rozdíl ukazuje, že u mužů je názor posunut mnohem blíže k jednoznačnému souhlasu než u žen.

Stejně i test Q_H ukazuje na vysokou statistickou významnost rozdílů ve tvarech distribucí ($Q_H = 44.47$, $df = 2$, $\alpha^* = 0.0000$). Koeficient neshody $C_{ord} = 0.1014$.

Závěr

Shrnutí technik pro posouzení shody marginálních distribucí ve čtvercových tabulkách, resp. shody dvou párově závislých rozložení četností, může být už jen stručně doplněno známým vzorcem pro porovnání závislých průměrů, tj. o specifickou situaci kardinálního znaku.

Postup testování hypotéz $\boldsymbol{H}_0$ vs $\boldsymbol{H}_1$, $\boldsymbol{H}_0$ vs $\boldsymbol{H}_2$, $\boldsymbol{H}_0$ vs $\boldsymbol{H}_3$, kde hypotézy se týkají populačních průměrů obou vstupujících proměnných μ_A, μ_B: $\boldsymbol{H}_0 \equiv \mu_A = \mu_B$, $\boldsymbol{H}_1 \equiv \mu_A \neq \mu_B$, $\boldsymbol{H}_2 \equiv \mu_A < \mu_B$, $\boldsymbol{H}_3 \equiv \mu_A > \mu_B$ se provádí pomocí z-testu ($n \geqq 30$)

$$(36) \quad Z = \sqrt{n}\,\frac{\bar{X}_B - \bar{X}_A}{\sqrt{s_A^2 + s_B^2 - 2\,\mathrm{cov}\,(A, B)}}$$

Ve vzorci je

$$\bar{X}_A = \sum_{k=1}^{K} n_{k+} x_k\,, \quad \bar{X}_B = \sum_{k=1}^{K} n_{+k} x_k\,,$$

$$s_A^2 = \frac{1}{n-1}\left(\sum_{k=1}^{K} n_{k+} x_k^2 - n\bar{X}_A^2\right), \qquad s_B^2 = \frac{1}{n-1}\left(\sum_{k=1}^{K} n_{+k} x_k^2 - n\bar{X}_B^2\right).$$

Kovariance se spočte jako

$$\mathrm{cov}\,(A, B) = \frac{1}{n-1}\left(\sum_{j=1}^{K}\sum_{k=1}^{K} n_{jk} x_j x_k - n\bar{X}_A\bar{X}_B\right);$$

$(x_1, x_2, \ldots, x_K)$ je společná kvantifikace obou proměnných **A** a **B**. Rozhodnutí se provádí zcela analogicky k postupu, jenž je vyjádřen v tabulce Wilcoxonova testu.

Připomínkou vzorce (36) je završena metodika porovnávání závislých distribucí pro standardní typy znaků. Pro zobecněné typy znaků (viz J. Řehák, B. Řeháková [1979]) je možno odvodit podobné testy jako Q_S, Q_H (resp. Q_B, Q_P).

Metodika testování shody může být zobecněna pomocí postupu simultánní statistické inference na porovnání více dvojic současně a tím na vytváření souhrnných závěrů o více dvojicích distribucí, tj. závěrů komplexního charakteru.

Ve čtvercových tabulkách lze řešit i jiné úlohy, jako je symetrie rozložení celé tabulky, quasisymetrie tabulky, problémy propojenosti obou marginálních distribucí (jejich strukturní shoda) atp.

Poznámky

[1]) Soubor statistických technik pro párově závislá data se v terminologii neparametrických postupů nazývá *„techniky pro jeden výběr"*. U pořadových dat se např. rozlišuje Wilcoxonův test pro jeden výběr (odpovídá párově spřaženým datům) a Wilcoxonův test pro dva výběry (odpovídá porovnání dvou nezávislých souborů).

[2]) Symbolem **A** × **B** budeme značit tabulku četnosti 2. stupně třídění, v níž je **A** řádková a **B** sloupcová proměnná. Vyšší stupně třídění lze značit obdobně **A** × **B** × **C** atd.

[3]) Při M-násobném zjišťování vlastností v časové posloupnosti se ovšem vyskytuje řada praktických a technických problémů, které spočívají v tom, že respondenti nejsou k dispozici po celou dobu, tj. že o nich není známa informace od začátku do konce, či že uprostřed od nich chybí údaje. To vede ke složité metodologii plánování takových časových sledování, která ovšem má vliv i na zpracování dat. Podrobný rozbor je mimo možnosti a cíle tohoto článku, proto jen upozorňujeme na problém.

[4]) Sem patří i situace, v níž na některé jednotky necháme působit experimentální podmínky a na některé jednotky jej necháme působit jen zdánlivě (tzv. „placebo"), přičemž oba soubory párujeme podle společných vlastností, jejichž vlivy chceme vyloučit.

[5]) Popsaná situace se liší od *analýzy skupinkové homogenity*, při níž nerozlišujeme jednotky věcně či podle jejich role ve skupině (v takové analýze také nemusí být skupinky stejně velké). Úloha skupinkové homogenity nebo heterogenity má svoji speciální statistickou metodiku, jež není v této stati popisována.

[6]) Případ shody dvouhodnotového znaku je totožný s testováním shody četností dvou (libovolných) jevů **a** a **b**, jejichž četnosti jsou odhadovány z těchže dat. Vztah obou jevů lze vyjádřit čtyřpolní tabulkou, v níž $\mathbf{A} = (a, \bar{a})$, $\mathbf{B} = (b, \bar{b})$, kde $\bar{a}$ je nevýskyt a, $\bar{b}$ je nevýskyt b. Tato úloha je převoditelná na značení používané v textu: $n_{11} = n_{ab} \equiv$ četnost společného výskytu obou jevů $n_{1+} = n_a$, $n_{+1} = n_b$.

[7]) Test aplikujeme tak, že dvojici čísel (n_{12}, n_{21}) považujeme za výsledek binomického rozložení s $p = \frac{1}{2}$, respektive pro značení znaménkového testu za dvojici (n_+, n_-).

[8]) Kritické hodnoty lze nalézt v běžných statistických tabulkách nebo v příručkách neparametrických technik. Viz též J. Řehák, B. Řeháková [1978a, s.85, tab. B], kde je případ aplikace znaménkového testu podrobně rozebrán.

[9]) Každá z K statistik, Q_k, odpovídá čtyřpolní tabulce vzniklé souběžnou dichotomizací obou proměnných **A**, **B**: $(a_k, \bar{a}_k)$, $(b_k, \bar{b}_k)$, $k = 1, 2, \ldots, K$. Test je založen na Bonferroniho nerovnosti. Ve skutečnosti by stačilo pracovat s kterýmikoli $(K - 1)$ statistikami.

[10]) Pro běžnou rutinu se používá: pro 95%ní spolehlivost je $Z_{0,05} = 1.96$, pro 99%ní spolehlivost $Z_{0,01} = 2.58$.

[11]) V praxi je někdy výhodnější a jednodušší pracovat se statistikami $Z_k = \sqrt{Q_k}$, jež hodnotíme podle standardního rozložení — dosažená hladina významnosti pro Z_k je $\alpha_k^* = 2(1 - \Phi(Z_k))$, tabulky distribuční funkce $\Phi(x)$ jsou snadno dostupné. Porovnání v metodě 1 provádíme pak $Z_{(1)} \gtrless Z_{\alpha/K}, Z_{(2)} \lesseqgtr Z_{\alpha/(K-1)}, \ldots$

Hodnoty $Z_{\alpha/M}$ jsou uvedeny v tabulce A.

[12]) Metodika simultánního testování je aplikací Holmova postupu sekvenčního zamítání hypotéz (Holm [1977], [1979]).

[13]) Posloupnost hladin významnosti $\alpha_1, \alpha_2, \alpha_3$ můžeme volit libovolně. V textu je popsána nejobvyklejší varianta, která se v praxi plně osvědčila. Proti tradičnímu postupu konstrukce znaménkových schémat, v nichž jsou jednotlivé statistiky porovnávány s pevnou hladinou významnosti, vede aplikace Holmovy sekvenční metody k závěrům, jež mají platnost na hladině významnosti α. (Obdobný způsob lze aplikovat i v kontingenčních tabulkách $R \times S$ asociačního či komparačního typu.)

[14]) Za výpočty Stuartova a Bhapkarova testu děkují autoři Dr. J. Trundovi, který pro ně využil svůj konverzační programový systém ISS.

[15]) Inverzi získáme většinou výpočtem na programovatelném kalkulátoru či kalkulačce nebo na počítači. Existují jednoduché metody inverze, které se bez počítače obejdou, jsou však pro větší K zdlouhavé.

[16]) Získáváme ji buď výpočtem na počítači nebo vyhledáním v tabulkách distribuční funkce rozložení chí-kvadrát.

[17]) Koeficient neshody má význam především při porovnávání různých dvojic znaků pro tentýž soubor (i když každý jednotlivý koeficient lze posuzovat vzhledem ke krajním hodnotám jeho definičního intervalu).

[18]) Zdánlivě nelogický výsledek se vyskytuje u porovnání rozdílů 10.6 % a 10.0 %, z nichž druhý (dokonce menší) je podstatně statisticky významnější. Rozdíly však odpovídají velmi různým poměrům: u „kuchyňského vybavení" je to 29 : 3 = 9,67 zatímco u kategorie „lednička" je poměr četností 72 : 44 = 1.64.

[19]) Role třídění vyšších stupňů pro hledání příčin diferenciačních jevů je obdobná postupům v analýze asociací a parciálních asociací. Podrobnější metodika však přesahuje rámec tohoto článku.

[20]) K tomu využijeme tabulku B, která obsahuje dosažené významnosti binomického dvoustranného testu.

[21]) Hodnota $\alpha_3^* = .0308$ by vypovídala o hypotéze $\boldsymbol{H}_0 : p_{3+} = p_{+3}$ proti $\boldsymbol{H}_A : p_{3+} \neq p_{+3}$ v případě, že by byla formulována před pohledem na data jako samostatná hypotéza, která nás zajímá bez ohledu na strukturu ostatních rozdílností.

[22]) Pro jednostranný test jsou obvyklé kritické hodnoty: $Z'_{0,05} = 1.64$, $Z'_{0,01} = 2.33$.

[23]) $\boldsymbol{H}_0$ vs $\boldsymbol{H}_2$ vyjadřuje ve skutečnosti test $(\tilde{X}_A \geqq \tilde{X}_B)$ vs $(\tilde{X}_A < \tilde{X}_B)$, $\boldsymbol{H}_0$ vs $\boldsymbol{H}_3$ vyjadřuje ve skutečnosti test $(\tilde{X}_A \leqq \tilde{X}_B)$ vs $(\tilde{X}_A > \tilde{X}_B)$.

[24]) Vysoká hodnota C $(> D)$ indikuje $\tilde{X}_A < \tilde{X}_B$, vysoká hodnota D $(> C)$ naopak $\tilde{X}_A > \tilde{X}_B$; kritické hodnoty znaménkového testu viz např. J. Řehák, B. Řeháková [1978a].

[25]) Pro matici 2×2 zjistíme inverzi $\mathbf{V}^{-1}$ snadno podle vzorečků: $V_{11}^{(-1)} = V_{11}/|V|$, $V_{22}^{(-1)} = V_{22}/|V|$, $V_{12}^{(-1)} = -V_{12}/|V|$, $V_{21}^{(-1)} = -V_{12}/|V|$, kde $|V| = V_{11}V_{22} - V_{12}V_{21}$. U symetrických matic platí navíc $V_{12}^{(-1)} = V_{21}^{(-1)}$, takže stačí zjistit jen tři prvky. V příkladě: $|V| = 13 \times 18 - 8 \times 8 = 170$, $V_{11}^{(-1)} = 18/170 = 0.1058824$, $V_{22}^{(-1)} = 13/170 = 0.0764706$, $V_{12}^{(-1)} = V_{21}^{(-1)} = -8/170 = 0.0470588$.

Literatura

Anděl, J.: *Matematická statistika.* Praha, SNTL 1978.

Bhapkar, V. P.: *A Note on the Equivalence of Two Criteria for Hypotheses in Ca-*

tegorial Data. Journal of the American Statistical Association 61, 1966, s. 228—235.
Cureton, E.: *The Normal Approximation to the Signed — Rank Sampling Distribution When Zero Differences Are Present.* Journal of the American Statistical Association 62, 1967, s. 1068—1069.
Holm, S.: *Sequentially Rejective Multiple Test Procedures.* Statistical Research Report 1977 — 1, University of Umeå, Sweden 1977.
Holm, S.: *A Simple Sequentially Rejective Multiple Test Procedure.* Scandinavian Journal of Statistics 6, 1979, s. 65—70.
Ireland, C. T. — Ku, H. H. — Kullback, S.: *Symmetry and Marginal Homogeneity of an r x r Contingency Table.* Journal of the American Statistical Association 64, 1969, s. 1323—1341.
Pratt, J. W.: *Remarks on Zeros and Ties in the Wilcoxon Signed Rank Procedures.* Journal of the American Statistical Association 54, 1959, s. 655—667.
Řehák, J. — Řeháková, B.: *Základní statistické testy pro rozložení diskrétních znaménkových dat.* Sociologický časopis XIV, 1978a, č. 1, s. 78—97
Řehák, J. — Řeháková B.: *Bayesovský znaménkový test.* Sociologický časopis XIV, 1978b, č. 2, s. 173—192.
Řehák, J. — Řeháková, B.: *Základní charakteristiky proměnných s konečným počtem hodnot a distanční analýza jejich rozložení.* Sociologický časopis XV, 1979, č. 2, s. 214—233.
Řehák, J. — Řeháková, B.: *Analýza kategorizovaných dat v sociologii,* Praha, Academia 1984 (v tisku).
Sheil, J. — O'Muircheartaigh, I.: *Algorithm AS 106; The Distribution of Nonnegative Quadratic Forms in Normal Variables.* Applied Statistics 26, 1977, s. 92—98.
Stuart, A.: *A test for Homogeneity of the Marginal Distributions in a Two-Way Classification.* Biometrika 42, 1955, s. 412—416.

Tabulková příloha

Tabulka A. Postupné kritické hodnoty testu chí-kvadrát (df = 1) a standardního normálního testu pro aplikaci sekvenčního Holmova postupu simultánní inference

Počet dílčích hypotéz	Celkový závěr na hladině								
	$\alpha = 0.05$			$\alpha = 0.01$			$\alpha = 0.001$		
K	α/K	$\chi^2_{\alpha/K,1}$	α/K	α/K	$\chi^2_{\alpha/K,1}$	$Z_{\alpha/K}$	α/K	$\chi^2_{\alpha/K,1}$	$Z_{\alpha/K}$
1	.05000	3.84	1.960	.01000	6.64	2.576	.00100	10,83	3.291
2	.02500	5.02	2.241	.00500	7.88	2.807	.00050	12.12	3.481
3	.16667	5.73	2.394	.00333	8.62	2.935	.00033	12.87	3.588
4	.01250	6.24	2.498	.00250	9.14	3.023	.00025	13.41	3.662
5	.01000	6.64	2.576	.00200	9.55	3.090	.00020	13.83	3.719
6	.00833	6.96	2.638	.00167	9.89	3.144	.00017	14.17	3.765
7	.00714	7.24	2.690	.00143	10.17	3.189	.00014	14.46	3.803
8	.00625	7.48	2.734	.00125	10.42	3.227	.00013	14.72	3.836
9	.00556	7.69	2.773	.00111	10.63	3.261	.00011	14.94	3.865
10	.00500	7.88	2.807	.00100	10.83	3.291	.00010	15.14	3.891
11	.00455	8.05	2.838	.00091	11.00	3.317	.00009	15.32	3.914
12	.00417	8.21	2.865	.00083	11.17	3.341	.00008	15.48	3.935
13	.00385	8.36	2.891	.00077	11.31	3.364	.00008	15.63	3.954
14	.00357	8.49	2.914	.00071	11.45	3.384	.00007	15.77	3.971
15	.00333	8.62	2.935	.00067	11.58	3.403	.00007	15.90	3.988
16	.00313	8.73	2.955	.00063	11.70	3.421	.00006	16.02	4.003
17	.00294	8.84	2.974	.00059	11.81	3.437	.00006	16.14	4.017
18	.00278	8.95	2.991	.00056	11.92	3.452	.00006	16.25	4.031
19	.00263	9.05	3.008	.00053	12.02	3.467	.00005	16.35	4.044
20	.00250	9.14	3.023	.00050	12.12	3.481	.00005	16.45	4.056

Poznámka: Postupné hladiny α/k a postupné kritické hodnoty pro K hypotéz nalezneme tak, že první hodnota je v K-tém řádku a ostatní jsou postupně nad ní.

Tabulka B. Dosažená hladina významnosti α^* pro dvoustranný binomický test. Sloupce odpovídají hodnotě S = menší z obou četností testu. Řádky odpovídají hodnotě $\bar{n} - S$ = větší z obou četností testu. Tabulka obsahuje desetinnou část hodnoty α^*

$\bar{n} \leq 30$

$\bar{n} - S$	$S = 0$	$S = 1$	$S = 2$	$S = 3$	$S = 4$	$S = 5$	$S = 6$	$S = 7$	$S = 8$	$S = 9$	$S = 10$
5	0625										
6	0313	0703									
7	0156	0391	0654								
8	0078	0215	0386	0923							
9	0039	0117	0225	0574	0768						
10	0020	0063	0129	0352	0490	0963					
11	0010	0034	0074	0213	0309	0636	0784				
12	0005	0018	0042	0127	0192	0414	0525	0931			
13	0002	0010	0023	0075	0118	0266	0347	0639	0755		
14	0001	0005	0013	0044	0072	0169	0227	0433	0522	0872	
15	0001	0003	0007	0026	0043	0106	0146	0290	0357	0614	0987
16	0000	0001	0004	0015	0026	0066	0094	0192	0241	0428	
17	0000	0001	0002	0009	0015	0041	0059	0125	0161		
18	0000	0000	0001	0005	0009	0025	0037	0081			
19	0000	0000	0000	0002	0005	0015	0023	0052			
20	0000	0000	0000	0001	0003	0009	0014				
21	0000	0000	0000	0001	0002	0005					
22	0000	0000	0000	0000	0001	0003					
23	0000	0000	0000	0000	0001						
24	0000	0000	0000	0000							
25	0000	0000	0000	0000							
26	0000	0000	0000	0000							
27	0000	0000	0000								
28	0000	0000	0000								
29	0000	0000									
30	0000										

Poznámky: a) zápis .0000 značí hodnotu menší než .00005.
b) Uvedeny jsou jen významnosti menší než .1000.
c) $\alpha^* = 2P_B\,(X \leq S|p = 0.5, \bar{n})$; tabulka vznikla adaptací z tab. 24 v J. Likeš — J. Laga: *Základní statistické tabulky*. Praha, SNTL 1978.

Резюме

Я. Ржегак — Б. Ржегакова: Сопоставление распределений у парнозависимых данных в квадратных таблицах сопряжения

Авторы статьи занимаются важной аналитической задачей, нередко встречающейся в социологической работе: сопоставлением двух парнозависимых распределений в квадратных контингенциальных таблицах. Методика обработки таких распределений, к сожалению, игнорируется даже широко известными и распространенными в мире программными системами (например, SPSSX). В вышеназванной статье предлагается решение проблемы во всей ее целостности и полноте. В первой части, названной Типы парнозависимых данных и социологические задачи, характеризуются свойства данных, из-за которых используется вышеназванная техника, и описываются основные ситуации, ведущие к возникновению парнозависимых (сопряженных) данных:
— два параллельных, но разных по содержанию вопроса, обращенных к тому же анкетируемому (или два параллельно выясняемых факта об одной и той же единице);
— двое данных об измерении того же свойства в двух неодинаковых по времени ситуациях;

— обнаружение того же свойства в разных условиях;
— изучение надежности, сопоставимости средств измерения;
— спаренные единицы (два парнозависимых набора, или комплекта).

Отдельные случаи подтверждаются примерами, взятыми из исследовательской практики, и у них дается обобщение на большее число переменных.

Вторая часть — Статистическая методика сопоставления парнозависимых распределений: номинальные переменные — содержит описание задачи сопоставления на четыре аналитических шага (первый шаг — критерий согласия; второй шаг — измерение степени несовпадения; третий шаг — измерение отличающихся по значению категорий; четвертый шаг — схема знаков).

Первый шаг решается с помощью известных критерий: биноминального, Мак-Немары, Стюарта и Ъхапкара и с помощью минимальной дискриминационной информации. Эти методы подытожены и описаны. Авторы предлагают также метод одновременных статистических выводов, удобный для подсчетов вручную (он не требует инверсии матриц). Далее предлагается коэффициент несовпадения и указывается его асимптотическая стандартая ошибка для построения интервалов достоверности. Предлагается методика для третьего и четвертого шагов.

Третья часть — Примеры анализа совпадения для номинальных данных — иллюстрирует методику предшествующей части на практических примерах, взятых из обычного социологического анализа.

В четвертой части — Статистическая методика сопоставления парнозависимых распределений: порядковые переменные — указывается критерий согласия кумулятивных функций распределения. Вводятся критерии аналогичные методам второй части. Более того, приводится адаптация одновыборочного критерия Уилкоксона и критерия знаков к случаю квадратной таблицы сопряжения.

Пятая часть — Примеры анализа согласия для порядковозависимых данных — играет роль иллюстрации методики четвертой части на конкретных социологических примерах.

В заключение для полноты приводится также парный тест Z для сопоставления зависимых средних величин в таблице сопряжения.

В статье приводятся также практические таблицы постепенных критических значений стандартного нормального критерия и хи-квадрат критерия с одной ступенью свободы, а также таблица достигнутых уровней значимости двустороннего биноминального (знакового) критерия. Они позволяют быстро осуществить одновременную оценку данных и в тех случаях, когда вычислительная машина не обрабатывает критерии согласия многозначных переменных.

Summary

J. Řehák — B. Řeháková: Comparison of Matched Data Distributions in Square Contingency Tables

The paper is concerned with an important analytical task frequently faced in sociological work: the comparison of two matched distributions (paired data) in square contingency tables. The statistical methodology for analysis of this kind of data is unfortunately ignored even by well-known program systems all over the world (e. g. SPSSH). The paper offers a complete solution of the problem in the full range of tasks. The first part *Types of matched data and sociological tasks*, characterizes the properties of data for which the mentioned techniques may be used, and describes the basic situations of the origin of matched data:
— two parallel questions, differing in content, given to the same respondent (or two parallel facts about one unit)
— the data on measuring the property under different conditions,
— measurement of the same property under different conditions,
— study of reliability, comparison of the means of measurement,
— paired units (two matched sets).

Individual cases are illustrated by examples from research practice. The generalization for more than two variables is also described.

The second part, *Statistical methodology of comparing distributions of matched data — nominal variables*, describes four analytical steps (1. test of homogenity, 2. measurement of the degree of heterogenity, 3. determination of significantly different categories, 4. sign scheme).

The first step is solved by well-known methods: binomial, Mc Nemar's, Stuart's and Bhaphar's tests ,and the test of minimal discriminative information. These are described. Moreover, the procedure of simultaneous statistical inference is proposed that is

suitable for manual calculations and does not require the inversion of matrices. Also the coefficient of dissimilarity is suggested together with its asymptotic standard error for the construction of confidence intervals. The methodology for exploratory steps 3 and 4 is designed.

The third part, *Examples of homogenity analysis for nominal data,* illustrates the methodology of the preceding part giving practical examples from routine sociological analysis.

The fourth part, *Statistical methodology of comparing distribution of matched data — ordinal variables,* introduces new procedures for testing homogeneity of cumulative distribution functions which correspond to ordinal variables. Tests analogical to the methods discussed in Part 2 are introduced. Moreover, an adaptation of Wilcoxon's one-sample test and sign test for the square contingency table is presented.

The fifth part, *Examples of homogeneity analysis for ordinal dependent data,* illustrates the methodology of Part 4 by practical sociological examples.

In conclusion, the z-test for comparing dependent means of cardinal variables in the contingency table is mentioned for the sake of completeness.

The paper also contains statistical tables of sequential critical values of the normal test and the x^2-test with one degree of freedom and a table of the attained levels of significance of the two-sided binomial (sign) test. It facilitates a rapid simultaneous testing in the data even in cases when the computer output does not supply homogeneity tests of multivalue variables.