Cvičení 11: Hodnocení kontingenčních tabulek

Úkol 1.: Testování hypotézy o nezávislosti, měření síly závislosti
V roce 1950 zkoumali Yule a Kendall barvu očí a vlasů u 6800 mužů.

Barva očí	Barva vlasů			
	světlá	kaštanová	černá	rezavá
modrá	1768	807	180	47
šedá nebo zelená	946	1387	746	53
hnědá	115	438	288	16

Na asymptotické hladině významnosti 0,05 testujte hypotézu o nezávislosti barvy očí a barvy vlasů. Vypočtěte Cramérův koeficient. Simultánní četnosti znázorněte graficky.

Postup ve STATISTICE:
Otevřeme datový soubor oci_vlasy.sta o 12 případech a třech proměnných (OCI, VLASY, CETNOST).
Před provedením testu je zapotřebí ověřit podmínky dobré aproximace:
Statistiky – Základní statistiky/tabulky – Kontingenční tabulky - Specif. tabulky – List 1 OCI, List 2 VLASY, OK, Váhy - CETNOST, Stav zapnuto, OK – na záložce Možnosti zaškrtneme Očekávané četnosti – Výpočet.

Souhrnná tab.: Očekávané četnosti (oci_vlasy.sta)
Četnost označených buněk > 10
Pearsonův chí-kv. : 1088,15, sv=6, p=0,00000

OCI	VLASY světlá	VLASY kaštanová	VLASY černá	VLASY rezavá	Řádk. součty
modrá	1167,259	1085,976	500,902	47,8622	2802,000
šedá nebo zelená	1304,731	1213,875	559,895	53,4990	3132,000
hnědá	357,010	332,149	153,202	14,6388	857,000
Vš.skup.	2829,000	2632,000	1214,000	116,0000	6791,000

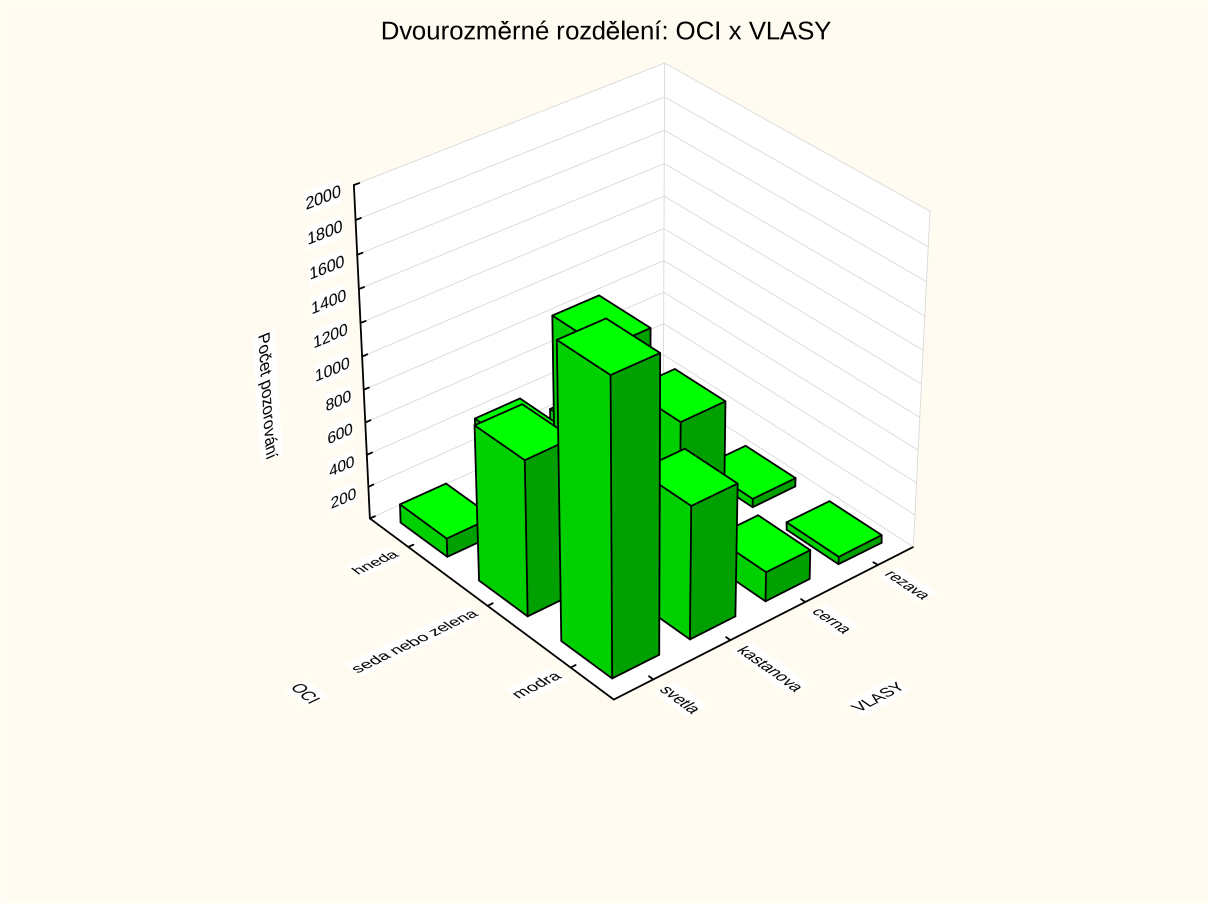
Podmínky dobré aproximace jsou splněny. Všechny teoretické četnosti jsou větší než 5. Nyní budeme testovat hypotézu o nezávislosti proměnných OCI, VLASY.

Návrat do Výsledky; kontingenční tabulky – na záložce Detaily zaškrtneme Pearsonův & M-L Chi - kvadrát, Phi & Cramerovo V – Detailní výsledky – Detailní 2 rozm. tabulky.

Statist.	Chí-kvadr.	sv	p
Pearsonův chí-kv.	1088,149	df=6	p=0,0000
M-V chí-kvadr.	1155,669	df=6	p=0,0000
Fí	,4002923		
Kontingenční koeficient	,3716246		
Cramér. V	,2830494		

Ve výstupní tabulce najdeme mj. hodnotu testové statistiky (Pearsonův chí-kv = 1088,149) s počtem stupňů volnosti (sv = 6) a odpovídající p-hodnotou (p = 0,0000), dále Cramérův koeficient (V = 0,283). Protože p-hodnota je mnohem menší než 0,05, nulovou hypotézu o nezávislosti barvy očí a barvy vlasů zamítáme na asymptotické hladině významnosti 0,05. Cramérův koeficient svědčí o slabé závislosti barvy očí a vlasů.

Pro grafické znázornění četností se vrátíme do Výsledky; kontingenční tabulky – Detailní výsledky – 3D histogramy.



Úkol k samostatnému řešení: Na hladině významnosti 0,05 testujte hypotézu o nezávislosti pedagogické hodnosti a pohlaví a vypočtěte Cramérův koeficient vyjadřující intenzitu závislosti pedagogické hodnosti na pohlaví, jsou-li k dispozici následující údaje:

pohlaví	pedagogická hodnost		
	odb. asistent	docent	profesor
muž	32	15	8
žena	34	8	3

Výsledek: Podmínky dobré aproximace jsou splněny, pouze jediná teoretická četnost klesne pod 5. Testová statistika K nabývá hodnoty 3,5, p = 0,1739, tedy na asymptotické hladině významnosti 0,05 nezamítáme hypotézu o nezávislosti pedagogické hodnosti a pohlaví. Cramérův koeficient: V = 0,187.

Úkol 2.: Fisherův faktoriálový test

100 náhodně vybraných mužů a žen bylo dotázáno, zda dávají přednost nealkoholickému nápoji A či B. Údaje jsou uvedeny ve čtyřpolní kontingenční tabulce.

preferovaný nápoj	pohlaví	
	muž	žena
A	20	30
B	30	20

Na hladině významnosti 0,05 testujte pomocí Fisherova faktoriálového testu hypotézu, že preferovaný typ nápoje nezáleží na pohlaví respondenta.

Návod: Otevřeme datový soubor napoje AB.sta
Statistiky – Základní statistiky/tabulky – Kontingenční tabulky - Specif. tabulky – List 1 NAPOJ, List 2 POHLAVI, OK, Váhy - CETNOST, Stav zapnuto, OK – na záložce Možnosti zaškrtneme Fisher exakt, Yates, McNemar (2x2) – Detailní výsledky – Detailní 2-rozm. tabulky.

	Statist. : NAPOJ(2) x POHLAVI(2) (napoje AB.sta)		
Statist.	Chí-kvadr.	sv	p
Yatesův chí-kv.	3,240000	df=1	p=,07186
Fisherův přesný, 1-str.			p=,03567
Fisherův přesný, 2-str.			p=,07134
McNemarův chí-kv. (A/D)	,0250000	df=1	p=,87437
McNemarův chí-kv. (B/C)	,0166667	df=1	p=,89728

Ve výstupní tabulce je mimo jiné uvedena p-hodnota pro oboustranný a jednostranný test. V našem případě se jedná o oboustranný test (nevíme, zda muži více preferují nápoj A či nápoj B než ženy), zajímáme se tedy o Fisherův přesný, 2-str. Ta je 0,07134. Protože p-hodnota je větší než 0,05, nezamítáme na hladině významnosti 0,05 hypotézu, že preferovaný typ nápoje nezáleží na pohlaví respondenta.

Úkol 3.: Podíl šancí
Pro údaje z úkolu 2 vypočtěte podíl šancí a sestrojte 95% asymptotický interval spolehlivosti pro podíl šancí. Pomocí tohoto intervalu spolehlivosti testujte na asymptotické hladině významnosti 0,05 hypotézu, že preferovaný typ nápoje nezáleží na pohlaví respondenta.

Návod: Nejprve zopakujme teorii:

Ve čtyřpolních tabulkách používáme charakteristiku $OR = \frac{ad}{bc}$, která se nazývá podíl šancí (odds ratio). Můžeme si představit, že pokus se provádí za dvojích různých okolností a může skončit buď úspěchem nebo neúspěchem.

Výsledek pokusu	okolnosti		$n_{j.}$
	I	II	
úspěch	a	b	a+b
neúspěch	c	d	c+d
$n_{.k}$	a+c	b+d	n

Poměr počtu úspěchů k počtu neúspěchů (tzv. šance) za 1. okolností je $\frac{a}{c}$, za druhých okolností je $\frac{b}{d}$. Podíl šancí je $OR = \frac{ad}{bc}$. Považujeme ho za odhad skutečného podílu šancí $o\rho$. Pomocí 100(1-α)% asymptotického intervalu spolehlivosti pro logaritmus skutečného podílu šancí $\ln o\rho$ lze na asymptotické hladině významnosti α testovat hypotézu o nezávislosti nominálních veličin X a Y.

Upozornění: Musí být splněny podmínky dobré aproximace.

Asymptotický 100(1-α)% interval spolehlivosti pro přirozený logaritmus skutečného podílu šancí má meze:

$\ln OR \pm \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}\, u_{1-\alpha/2}$. Jestliže interval spolehlivosti nezahrne 0, pak hypotézu o nezávislosti zamítneme na asymptotické hladině významnosti α.

V našem případě podíl šancí vypočteme ručně: $OR = \frac{ad}{bc} = \frac{20 \cdot 20}{30 \cdot 30} = \frac{4}{9} = 0,\bar{4}$.

Dolní a horní mez intervalu spolehlivosti pro ln ορ zjistíme pomocí STATISTIKY.
Ověříme splnění podmínek dobré aproximace a zjistíme, že všechny teoretické četnosti jsou 25.
Vytvoříme datový soubor o dvou proměnných DM a HM a dvou případech. Do Dlouhého jména proměnné DM napíšeme vzorec pro dolní mez:
=log(4/9)-sqrt(1/20+1/30+1/30+1/20)*VNormal(0,975;0;1)
a analogicky do Do Dlouhého jména proměnné HM napíšeme vzorec pro horní mez:
=log(4/9)+sqrt(1/20+1/30+1/30+1/20)*VNormal(0,975;0;1)

	1 DM	2 HM
1	-1,61108	-0,01078

Výsledek: -1,61108 < ln ορ < -0,01078 s pravděpodobností přibližně 0,95. Protože tento interval spolehlivosti neobsahuje 0**,** na asymptotické hladině významnosti 0,05 zamítáme hypotézu, že preferovaný typ nápoje nezáleží na pohlaví respondenta. Ke stejnému výsledku dospějeme, pokud použijeme Pearsonův chí-kvadrát test o nezávislosti.
Tento výsledek je v rozporu s výsledkem, ke kterému dospěl Fisherův přesný test. Je to způsobeno tím, že test pomocí asymptotického intervalu spolehlivosti a Pearsonův chí-kvadrát test jsou pouze přibližné.

Úkol k samostatnému řešení: 36 mužů onemocnělo určitou chorobou. Někteří z nich se léčili, jiní ne. Někteří se uzdravili, jiní zemřeli. Údaje jsou uvedeny ve čtyřpolní kontingenční tabulce.

přežití	léčení	
	ano	ne
ano	10	6
ne	12	8

Vypočtěte a interpretujte podíl šancí. Pomocí intervalu spolehlivosti pro podíl šancí testujte na asymptotické hladině významnosti 0,05 hypotézu, že přežití nezávisí na léčení proti tvrzení, že léčení zvyšuje šance na přežití.

Výsledek: $OR = 1,\bar{1}$, nulovou hypotézu nezamítáme asymptotické hladině významnosti 0,05, protože levostranný 95% asymptotický interval spolehlivosti pro logaritmus podílu šancí je (-1,03; ∞).

Úkol 4.: Testování hypotézy o symetrii ve čtyřpolní tabulce (McNemarův test)
Máme náhodný výběr 18 pacientů, kteří byli léčeni dvěma různými antihypertenzívy A a B. Každý pacient dostával po dobu jednoho měsíce lék A a po odeznění jeho případných účinků dostával po dobu jednoho měsíce lék B. Výsledek byl klasifikován jako úspěch nebo neúspěch. Za úspěch byl pokládán pokles krevního tlaku alespoň o 15 mm Hg. Každý jiný výsledek byl považován za neúspěch. Lék A byl úspěšný u 4 pacientů, přičemž u jednoho z nich byl úspěšný i lék B. Lék B byl úspěšný u 10 pacientů. Na asymptotické hladině významnosti 0,05 testuje hypotézu, že pravděpodobnost úspěchu je stejná pro oba léky.

Návod:
Údaje z textu úlohy uspořádáme do čtyřpolní kontingenční tabulky.

Lék A	Lék B	
	Úspěch ano	Úspěch ne
Úspěch ano	1	3
Úspěch ne	9	5

Podmínky dobré aproximace McNemarova testu jsou splněny, b + c = 12 > 8.

Testová statistika: $K = \frac{(b-c)^2}{b+c} = \frac{(3-9)^2}{3+9} = 3$.

Kritický obor: $W = \langle \chi^2_{1-\alpha}(1), \infty) = \langle \chi^2_{0,95}(1), \infty) = \langle 3{,}841, \infty)$.

Protože $K \in W$, H_0 nezamítáme na asymptotické hladině významnosti 0,05. Nepodařilo se tedy prokázat, že pravděpodobnost úspěchu se pro léky A a B liší.

Postup v systému STATISTICA:
Otevřeme datový soubor leky AB.sta
Statistiky – Základní statistiky/tabulky – Kontingenční tabulky - Specif. tabulky – List 1 LEK, List 2 USPECH, OK, Váhy - CETNOST, Stav zapnuto, OK – na záložce Možnosti zaškrtneme Fisher exakt, Yates, McNemar (2x2) – Detailní výsledky – Detailní 2-rozm. tabulky.
Dostaneme výstupní tabulku:

	Statist. : LEK(2) x USPECH(2) (leky AB.sta)		
Statist.	Chí-kvadr.	sv	p
Yatesův chí-kv.	,6790178	df=1	p=,40993
Fisherův přesný, 1-str.			p=,20588
Fisherův přesný, 2-str.			p=,27451
McNemarův chí-kv. (A/D)	1,500000	df=1	p=,22067
McNemarův chí-kv. (B/C)	2,083333	df=1	p=,14891

Zajímá nás poslední řádek označený McNemarův chí-kv. (B/C). Testová statistika nabývá hodnoty 2,0833 (systém STATISTICA používá při výpočtu testové statistiky tzv. opravu na nespojitost: $K = \frac{(|n_{12} - n_{21}| - 1)^2}{n_{12} + n_{21}}$). Odpovídající p-hodnota je 0,1489, tedy na asymptotické hladině významnosti 0,05 nezamítáme hypotézu, že pravděpodobnost úspěchu je stejná pro oba léky.

Úkol k samostatnému řešení: Při výuce angličtiny byla použita speciální metoda pro zrychlení konverzace. Tato metoda byla aplikována u 100 studentů s následujícími výsledky: z 52 studentů, kteří byli před aplikací metody rychlí v konverzaci, se jich zlepšilo 24 a ze 48 studentů, kteří byli pomalí v konverzaci, se jich zlepšilo 12. Na hladině významnosti 0,05 testujte hypotézu, že speciální metoda nemá vliv na rychlost reakce v konverzaci.
Výsledek: Testová statistika 6,4 je větší než kritická hodnota 3,84, nulovou hypotézu zamítáme na asymptotické hladině významnosti 0,05.

Další úkoly k samostatnému řešení:

1. Následující kontingenční tabulka ukazuje výsledky velmi slavného lékařského experimentu ze čtyřicátých let, který se zabýval účinkem streptomycinu při léčbě plicní tuberkulózy. Údaje z radiologického hodnocení po 6 měsících byly porovnány s tím, zda pacient patřil do léčené, nebo kontrolní skupiny. Lze na hladině významnosti 0,05 prokázat vztah mezi léčbou a výsledkem?

Radiologické hodnocení	Streptomycin	Kontrolní	Celkem
Významné zlepšení	28	4	32
Střední / malé zlepšení	10	13	23
Beze změn	2	3	5
Střední / malé zhoršení	5	12	17
Významné zhoršení	6	6	12
Smrt	4	14	18
Celkem	55	52	107

Výsledek: Testová statistika 26,96 je větší než kritická hodnota 15,09, nulovou hypotézu zamítáme na asymptotické hladině významnosti 0,05.

2. V průzkumu o kuřáctví bylo dotázáno 92 osob. Z 64 mužů jich kouří 19 a z 28 žen jich kouří 6.
a) Na hladině významnosti 0,05 testujte hypotézu, že kouření se vyskytuje stejně často u mužů a žen. Použijte Pearsonův chí-kvadrát test i Fisherův přesný test.
b) Vypočtěte a interpretujte podíl šancí a stanovte meze 95% intervalu spolehlivosti pro podíl šancí.
Výsledek:
Ad a) Před provedením Pearsonova chí-kvadrát testu je zapotřebí ověřit splnění podmínek dobré aproximace. Je to v pořádku, všechny čtyři teoretické četnosti jsou větší než 5. Testová statistika Pearsonova chí-kvadrát testu je 0,6714, což je menší než kritická hodnota 3,84, nulovou hypotézu nezamítáme na asymptotické hladině významnosti 0,05.
Pro Fisherův přesný test vyjde p-hodnota 0,4576, což je větší než hladina významnosti 0,05, nulovou hypotézu nezamítáme na hladině významnosti 0,05.
Ab b) Podíl šancí je 1,55, což znamená, že u mužů je šance na kouření 1,55 x vyšší než u žen.
$0,5418 < o\rho < 4,4239$ s pravděpodobností aspoň 0,95.