

Příklady k procvičování

Příklad 10.1. Testování hypotézy o nezávislosti, měření síly závislosti

Otevřete si soubor ped.hodnost.txt. Na hladině významnosti $\alpha = 0.05$ testujte hypotézu o nezávislosti pedagogické hodnosti a pohlaví. Dále vypočtete Cramérův koeficient vyjadřující intenzitu závislosti pedagogické hodnosti na pohlaví. Data v souboru mají následující tvar:

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

Řešení příkladu 10.1

- H_0 : Znaky pohlaví a pedagogická hodnost stochasticky nezávislé.
- H_1 : Znaky pohlaví a pedagogická hodnost stochasticky nezávislé.
- Hladina významnosti $\alpha =$

Podmínka dobré aproximace

	odb. asistent	docent	profesor	
muz	36.3	12.65	6.05	1
zena	29.7	10.35	4.95	2
				3

Podmínky dobré aproximace splněny. Všechny teoretické četnosti až na jednu jsou než 5.

Pearsonův χ^2 test

Pearson's Chi-squared test	4
	5
	6
data: data	7
X-squared = 3.4988, df = 2, p-value = 0.1739	8

q	9
1 5.991465	10

a) Test kritickým oborem

Hodnota testovací statistiky K je Kritický obor má tvar
Protože, H_0 na hladině významnosti $\alpha =$

b) Test p -hodnotou P -hodnota vyšla Protože p -hodnota, H_0
na hladině významnosti $\alpha =$

Pro zjištění míry závislosti v kontingenční tabulce použijeme koeficient.

Crameruv.koeficient	11
1 0.1870496	12

Hodnota Cramérova koeficientu je

Interpretace výsledků: Znaky pohlaví a pedagogická hodnost jsou / nejsou stochasticky nezávislé. Mezi pohlavím a pedagogickou hodností existuje stupeň závislosti.

Příklad 10.2. Podíl šancí

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

Vypočtete a interpretujte podíl šancí. Pomocí intervalu spolehlivosti pro logaritmus podílu šancí testujte na asymptotické hladině významnosti $\alpha = 0.05$ hypotézu, že přežití nezávisí na léčení, proti tvrzení, že léčení zvyšuje šance na přežití.

Řešení příkladu 10.2

- H_0 : \rightarrow
- H_1 : \rightarrow
- Hladina významnosti $\alpha =$

Podmínka dobré aproximace

Podmínky dobré aproximace splněny. Všechny teoretické četnosti jsou než 5.

	ano	ne	
ano	9.777778	6.222222	13
ne	12.222222	7.777778	14
			15

Výpočet (logaritmu) podílu šancí

	OR	lnOR	
1	1.111111	0.1053605	16
			17

Podíl šancí $OR =$ Logaritmus podílu šancí $\ln(OR) =$

a) Test kritickým oborem

	t_0	q	
1	0.1528731	-1.644854	18
			19

Hodnota testovací statistiky t_0 je Kritický obor má tvar
Protože, H_0 na hladině významnosti $\alpha =$

b) Test intervalem spolehlivosti

	dh	
1	-1.028277	20
		21

Interval spolehlivosti má tvar Protože, H_0
na hladině významnosti $\alpha =$

c) Test p -hodnotou

	p	
1	0.5607508	22
		23

P -hodnota vyšla Protože p -hodnota, H_0 na hladině významnosti $\alpha =$

Interpretace výsledků: Znaky léčba a stav pacienta jsou / nejsou stochasticky nezávislé.

Příklad 10.3. Pearsonův χ^2 test, Fisherův faktoriálový test, test podílu šancí

V průzkumu o kuřáctví bylo dotázáno 92 osob. Z 64 mužů jich kouří 19 a z 28 žen jich kouří 6. Na hladině významnosti $\alpha = 0.05$ testujte hypotézu, že kouření se vyskytuje stejně často u mužů a žen. Použijte Pearsonův chí-kvadrát test, Fisherův faktoriálový test i test podílem šancí.

Řešení příkladu 10.3

Pearsonův χ^2 test

- H_0 : Znaky pohlaví a preference stochasticky nezávislé.
- H_1 : Znaky pohlaví a preference stochasticky nezávislé.
- Hladina významnosti $\alpha =$

Podmínka dobré aproximace

	muz	zena	
kurak	17.3913	7.608696	24
nekurak	46.6087	20.391304	25
			26

Podmínky dobré aproximace splněny. Všechny teoretické četnosti jsou než 5.

a) Test kritickým oborem

	t0	q	
1	0.6713646	3.841459	27
			28

Hodnota testovací statistiky K je Kritický obor má tvar
 Protože, H_0 na hladině významnosti $\alpha =$

b) Test p -hodnotou

	p	
1	0.4125762	29
		30

P -hodnota vyšla Protože p -hodnota, H_0 na hladině významnosti $\alpha =$

Fisherův faktoriálový test

Fisher's Exact Test for Count Data		
data:	data	31
p-value	= 0.4576	32
alternative hypothesis:	true odds ratio is not equal to 1	33
95 percent confidence interval:		34
	0.498056 5.398695	35
sample estimates:		36
odds ratio		37
	1.54109	38
		39
		40
		41

Protože p -hodnota=..... je než $\alpha = 0.05$, nulovou hypotézu H_0 o nezávislosti kouření a pohlaví na hladině významnosti $\alpha =$

Podíl šancí

- H_0 : →
- H_1 : →
- Hladina významnosti α =
- Podmínky dobré aproximace jsou splněny (viz Pearsonův χ^2 test).

Výpočet (logaritmu) podílu šancí

	OR	lnOR
1	1.548148	0.4370595

42
43

Podíl šancí OR = Logaritmus podílu šancí $\ln(OR)$ =

a) **Test kritickým oborem**

	t0	q1	q2
1	0.8158654	-1.959964	1.959964

44
45

Hodnota testovací statistiky t_0 je Kritický obor má tvar
Protože, H_0 na hladině významnosti α =

b) **Test intervalem spolehlivosti**

Proti alternativě postavíme IS.

	dh	hh
1	-0.6128942	1.487013

46
47

Interval spolehlivosti má tvar Protože, H_0
na hladině významnosti α =

c) **Test p-hodnotou**

	p
1	0.4145771

48
49

P -hodnota vyšla Protože p -hodnota, H_0 na hladině významnosti α =

Interpretace výsledků: Znaky pohlaví a preference jsou / nejsou stochasticky nezávislé.