UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Nika Mirnik

Testiranje homogenosti in neodvisnosti v kontingenčnih tabelah reda 2×2

Delo diplomskega seminarja

Mentor: doc. dr. Jaka Smrekar

Ljubljana, 2013

Kazalo

Tabele

Testiranje homogenosti in neodvisnosti v kontingenčnih tabelah reda 2×2

Povzetek

Delo diplomskega seminarja nas vodi skozi ključne elemente statističnega testiranja. Opisuje testiranje hipotez, pri čemer opozarja na možnost napake, ki se lahko pojavi ob potrditvi oz. zavrnitvi ničelne hipoteze, in kakšno vlogo ima stopnja značilnosti pri tej napaki. Opisuje tudi formulacijo testne statistike in različne metode vzorčenja, ki jih lahko uporabimo. Predstavljena je tudi teorija testiranja neodvisnosti in homogenosti za 2×2 kontingenčno tabelo. Test neodvisnosti preverja hipotezo, da sta dve dekompoziciji neodvisni druga od druge. Test homogenosti pa preverja, ali sta dekompoziciji homogeni glede na opažene frekvence v obeh kategorijah. Diplomsko delo se nato osredotoči na χ^2 testa neodvisnosti in homogenosti. Opisuje in razloži predpostavke, na katerih temelji test, in nas vodi skozi računski del testa. Pokaže, kako potrdimo oz. ovržemo hipotezo s pomočjo χ^2 porazdelitve. Sledi Fisherjev eksaktni test, ki privzema drugačne predpostavke kot χ^2 test in uvede tudi enostransko alternativno hipotezo. Diplomsko delo nas vodi skozi računski proces in razloži, kako test uporablja hipergeometrijsko porazdelitev za vrednotenje ničelne hipoteze. Na koncu primerjam oba statistična testa in pri tem opozorim na razlike med njima in na primernost posameznega testa glede na dano velikost populacije.

Testing for homogeneity and independence in 2×2 *contingency tables*

Abstract

This graduation thesis walks us through key elements of statistical testing in general. It describes the logical framework of hypothesis testing, drawing attention to the problem of errors that can be made rejecting or confirming the null hypothesis and the role the level of significance plays in that error. It also describes the formulation of the test statistics and different methods of sampling one can use. The theory of testing for independence and homogeneity on 2×2 tables is presented, for testing in general. The statistical test of independence on a 2×2 table evaluates the general hypothesis that the two variables are independent of one another. The statistical test of homogeneity on 2×2 table evaluates whether or not the two samples are homogeneous with respect to the proportion of observations in each of the two categories. The thesis then focuses on χ^2 tests of independence and homogeneity. It describes and explains the assumptions on which the test is based. It leads us through the test itself and shows how we can confirm or reject the null hypothesis with the help of χ^2 distribution. This is followed by Fisher's exact test which requires different assumptions than the χ^2 test and introduces the use of directional alternative hypothesis. The thesis leads us through the calculation process explaining how this test uses hypergeometric distribution to evaluate the null hypothesis. The thesis is concluded by comparing the two statistical test, pointing out the differences between them and suitability of a test according to the sample size.
Math. Subj. Class. (2010): 62F03
Ključne besede: Testiranje hipotez, testna statistika, vzorčenje, neodvisnost, homogenost, hi kvadrat test, Fisherjev eksaktni test
Keywords: Hypothesis testing, test statistics, sampling, independence, homogeneity, chi square test, Fisher's exact test

1 Uvod

Recimo, da na neki končni populaciji obravnavamo dve lastnosti. Na primer, na populaciji študentov tretjih letnikov na FMF obravnavamo „spol“ (z možnima vrednostma ženski, moški) in „Zoisovo štipendijo“ (z možnima vrednostma ima štipendijo, nima štipendije).
Iz populacije potem naključno izberemo N študentov in jih razvrstimo glede na obe lastnosti na sledeč način:

	ženski	moški
ima štipendijo	a_{11}	a_{12}
nima štipendije	a_{21}	a_{22}

Tabela 1: primer: štipendija

Opazimo, da velja: $a_{11} + a_{12} + a_{21} + a_{22} = N$. Taki razpredelnici pravimo kontingenčna tabela.

Če bi nas zanimalo, ali je tabela 1 konsistentna s predpostavko, da sta lastnosti spol in Zoisova štipendija neodvisni, potem bi uporabili statistične teste neodvisnosti.

Recimo, da želimo za populacijo vzeti vse študentke FMF. Iz te populacije izberemo 2m študentk in jih razdelimo na 2 skupini po m študentk. Ti dve skupini označimo z A in B. Potem dekleta prosimo, da se pogledajo v ogledalo. Skupino A postavimo pred ogledalo, ki naredi njihovo postavo bolj ozko. Skupino B pa postavimo pred navadno ogledalo. Potem jim lahko zastavimo vprašanje: „Ali se počutite samozavestne?“ Možna sta seveda odgovora da ali ne. Rezultate lahko razdelimo takole:

	A	B
da	a_{11}	a_{12}
ne	a_{21}	a_{22}

Tabela 2: primer: ogledalo

Opazimo, da velja: $a_{11} + a_{21} = a_{12} + a_{22} = m$. Vsote stolpcev in vrstic so sedaj enolično določene s strani vodje eksperimenta.

Torej, če bi nas zanimalo, ali je tabela 2 konsistentna s predpostavko, da ogledalo, ki naredi dekleta bolj ozka, ne vpliva na samozavest študentk, potem bi uporabili statistične teste homogenosti.

2 Začetki

Pionir na področju testov na 2×2 kontingenčnih tabelah je bil Karl Pearson. Leta 1900 je predstavil χ^2 test ujemanja, ki se je izkazal za zelo uporabnega v mnogih različnih kontekstih. Na žalost pa je Pearson spregledal dejstvo, da je potrebno odšteti po eno prostostno stopnjo za vsak dodatni parameter pridobljen iz naših podatkov. Ta napaka je bila najbolj kritična ravno pri 2×2 kontingenčnih tabelah, saj moramo na njih uporabiti χ^2 z eno prostostno stopnjo, ne s tremi, kot je mislil Pearson.

Naslednji se je s kontingenčnimi tabelami ukvarjal Udny Yule, ki je leta 1911 v svoji knjigi *Introduction to the Theory of Statistics* predstavil test uporaben tudi za 2×2 tabele. V tej knjigi še ni omenjal χ^2 testa, ampak do leta 1915 je neskladje med njegovim in Pearsonovim testom s tremi prostostnimi stopnjami postalo očitno. Kot rezultat tega je zasnoval 350 2×2 tabel in 100 4×4 tabel s pomočjo mehanske naprave, na podlagi česar je dobil neodvisne porazdelitve in jih nato primerjal z χ^2 porazdelitvijo. Rezultatov tega poskusa ni takoj objavil.

Potem je leta 1922 R. A. Fisher objavil članek, v katerem je izpostavil Pearsonovo napako. Kljub temu, da se Yule ni popolnoma strinjal s Fisherjevim dokazom, je skoraj simultano objavil rezultate svojega poskusa, ki je potrdil Fisherjeve rezultate. Pearson svoje napake ni takoj priznal, kar je zbudilo veliko polemik, kljub temu pa so Fisherjevi popravki hitro postali splošno sprejeti.

χ^2 test je v resnici le dober približek in ne bo popolnoma točen v primerih, kjer je pričakovana vrednost posamezne celice v kontingenčni tabeli majhna. V knjigi *Statistical Methods for Research Workers* je Fisher objavil pravilo, da mora biti pričakovana vrednost v celici vsaj 5. To pravilo je sicer ustrezno, ampak precej konzervativno za teste z več kot eno prostostno stopnjo.

Če je eksaktna porazdelitev nekega problema znana, lahko točnost χ^2 testa preverimo tako, da primerjamo učinkovitost testa z učinkovitostjo eksaktne porazdelitve na vrsti primerov. Leta 1933 se je F. Yates začel zanimati za primerjavo eksaktne binomske porazdelitve s χ^2 porazdelitvijo na 2×2 tabelah s predhodno določenimi vsotami vrstic in stolpcev. Rezultat te primerjave je bil Yatesov popravek χ^2 formule, ki ga je objavil leta 1934. Fisher je še isto leto izdal novo verzijo knjige *Statistical Methods for Research Workers*, v katero je dodal Yatesov popravek, istočasno pa je v njej predstavil tudi Fisherjev eksaktni test.

Statistiki so razvijali tudi nove teste za testiranje na 2×2 tabelah. Eden izmed omembe vrednih je test, ki ga je leta 1945 objavil G. A. Barnard. Trdil je, da je boljši od Fisherjevega eksaktnega testa, a se je kmalu izkazalo, da je Barnardov test boljši le v izrednih primerih. Natančnost Fisherjevega testa prevlada v večini primerov, poleg tega pa so izračuni za Barnardov test zahtevnejši od izračunov za Fisherjev test.

To je postavilo temelje testiranju na 2×2 kontingenčnih tabelah. Še vedno pa prihaja do polemik, da je Fisherjev test preveč konzervativen, ker predvideva fiksiranje robnih vrednosti v tabeli, kar je v realnosti težko doseči.

3 Splošno o testiranju

Imamo neko slučajno spremenljivko X, katere porazdelitev je vsaj delno neznana. Torej iščemo porazdelitev za X.
Najprej je potrebno konstruirati model. Privzeti model pomeni najprej predpisati dopustne porazdelitve za X. V tem smislu model enačimo z množico dopustnih porazdelitev oz. z množico dopustnih porazdelitvenih gostot. Na primer, če privzamemo, da je X normalno porazdeljena slučajna spremenljivka, ima gostoto porazdelitve oblike:

$$f_{(\mu,\sigma)}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}(x-\mu)^2},$$

kjer je $\mu \in \mathbb{R}$ in $\sigma \in (0, \infty)$.
Množica dopustnih porazdelitev je tedaj

$$\{f_{(\mu,\sigma)} \Big| (\mu, \sigma) \in \mathbb{R} \times (0, \infty)\}.$$

V splošnem množico dopustnih porazdelitev označimo z

$$F = \{f_\vartheta : \mathbb{R} \to \mathbb{R} \Big| \vartheta \in \Theta\},$$

kjer je Θ množica parametrov, s katero opišemo gostoto porazdelitve f_ϑ. V normalnem modelu je $\Theta = \mathbb{R} \times (0, \infty)$.

3.1 Hipoteze in napake

Privzemimo model $\{f_\vartheta \Big| \vartheta \in \Theta\}$ s parametričnim prostorom Θ in razdelimo $\Theta = H_0 \underline{\cup} H_1$.
Testiramo: $\vartheta \in H_0$ proti alternativi $\vartheta \in H_1$.

Zdaj pa konstruiramo test:

$$\phi(\text{vzorec}) = \begin{cases} 1, & \text{vzorec} \in B \\ 0, & \text{sicer.} \end{cases}$$

Če $\phi = 1, H_0$ zavržemo. Množici B pravimo območje zavrženja.
Če $\phi = 0, H_0$ ne zavržemo, torej jo sprejmemo.

Zavedati se moramo, da lahko ob koncu testa izberemo pravilno odločitev, lahko pa se pojavita napaki dveh tipov:

NAPAKA TIPA I: zavržemo H_0, čeprav drži:
v ta namen določimo stopnjo značilnosti α , ki je največja dovoljena verjetnost, da se nam ta napaka dogodi. To naredimo na sledeči način:

$$\sup_{\vartheta \in H_0} P_\vartheta(vzorec \in B) \leq \alpha$$

NAPAKA TIPA II: <u>sprejememo H_0, čeprav ne drži</u>

Tu pa moramo biti pazljivi: napaka tipa II je

$$\sup_{\vartheta \in H_1} P_\vartheta(vzorec \notin B) = 1 - \inf_{\vartheta \in H_1} P_\vartheta(vzorec \in B)$$

Če definiramo $\alpha' = \inf_{\vartheta \in H_1} P_\vartheta(vzorec \in B)$, potem je napaka tipa II velikosti $1 - \alpha'$ in ker $\alpha \to 0$, gre istočasno $1 - \alpha' \to 1$. Pri dovolj lepi množici Θ in dovolj zvezni porazdelitvi $\Theta = H_0 \underline{\cup} H_1$ pa velja $\alpha = \alpha'$. Zato je potrebno α izbrati pametno in previdno.

Posledice takih napak so odvisne od problema, s katerim se ukvarjamo. Zaželeno je, da se test izvede tako, da so možnosti za katerokoli izmed teh dveh napak minimalne.

3.2 Vzorčenje in testne statistike

3.2.1 Vzorčenje

Vzorčenje je izbiranje manjše skupine posameznikov iz populacije, ki s svojimi lastnostmi predstavljajo lastnosti celotne populacije. Taka manjša skupina se imenuje vzorec populacije. Če je vzorec reprezentativen, lahko z njim dobimo celo bolj zanesljive podatke kot s popisom populacije, saj obstaja verjetnost, da bodo administrativne napake zbiranja in vnosa velike količine podatkov popisa večje kot vzorčne napake.

Obstajata dve glavni metodi vzorčenja: vzorčenje s ponavljanjem in vzorčenje brez ponavljanja.

Pri vzorčenju s ponavljanjem zaporedoma neodvisno izbiramo elemente iz populacije in z njimi sestavljamo vzorec določene velikosti, recimo N. To pomeni, da lahko element iz populacije izberemo tudi dvakrat.

Pri vzorčenju brez ponavljanja pa naenkrat izberemo N različnih elementov iz populacije. V tem primeru enega elementa iz populacije ne moremo uporabiti dvakrat.

3.2.2 Testna statistika

Ko konstruiramo ničelno in alternativno hipotezo, stopnjo značilnosti in metodo vzorčenja, potrebujemo test, ki ga izvedemo s pomočjo testne statistike. To je taka funkcija $T : (vzorci) \to \mathbb{R}$, s katero opišemo območje zavrženja B. Tipično je $B = \left\{vzorec \middle| T(vzorec) > c \text{ ali } T(vzorec) < -c\right\}$. Tako število c tedaj imenujemo kritična vrednost testne statistike. Za določitev c moramo poznati (približno) porazdelitev slučajne spremenljivke $T(vzorec)$ pri pogoju, da velja ničelna hipoteza. Statistika T naj bi merila "odstopanje" od ničelne hipoteze. Pričakujemo, da je T blizu 0, če je bil vzorec generiran pri veljavnosti ničelne hipoteze.

3.3 Kontingenčna tabela

Definicija: kontingenčna tabela je tabela, v katero zapišemo frekvence enot, ki ustrezajo spremenljivkama v čelu in glavi tabele.

Kontingenčna tabela dimenzije $r \times c$ sestoji iz r vrstic in c stolpcev. Pri tem sta r in c celi števili, večji ali enaki 2. Število celic v taki tabeli dobimo tako, da pomnožimo vrednost r z vrednostjo c. Podatek v vsaki od celic v kontingenčni tabeli je pravzaprav število opažanj.

Kontingenčna tabela velikosti $r \times c$

		Spremenljivka stolpca						Vsote vrstic
		C_1	C_2	$\cdots$	C_j	$\cdots$	C_c	
	R_1	O_{11}	O_{12}	$\cdots$	O_{1j}	$\cdots$	O_{1c}	$O_{1\bullet}$
	R_2	O_{21}	O_{22}	$\cdots$	O_{2j}	$\cdots$	O_{2c}	$O_{2\bullet}$
Spremenljivka vrstic	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
	R_i	O_{i1}	O_{i2}	$\cdots$	O_{ij}	$\cdots$	O_{ic}	$O_{i\bullet}$
	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
	R_r	O_{r1}	O_{r2}	$\cdots$	O_{rj}	$\cdots$	O_{rc}	$O_{r\bullet}$
Vsota stolpcev		$O_{\bullet 1}$	$O_{\bullet 2}$	$\cdots$	$O_{\bullet j}$	$\cdots$	$O_{\bullet c}$	n

Tabela 3: Osnovni model za $r \times c$ kontingenčno tabelo

V tabeli je n opažanj. Vsaka celica se identificira z indeksom, ki sestoji iz dveh elementov. Prvi element nam pove vrstico, v kateri se nahajamo, drugi pa stolpec, v katerem se nahajamo. Potemtakem znak O_{ij} pomeni število opažanj v celici, ki se nahaja v i-ti vrstici in j-tem stolpcu. $O_{i\bullet}$ je število vseh opažanj v i-ti vrstici in pa $O_{\bullet j}$ je število vseh opažanj v j-tem stolpcu.

Mi se bomo ukvarjali le s tabelami velikosti 2×2. Tudi za te tabele velja vse zgoraj našteto, le da je vse skupaj bolj enostavno.

3.4 Test neodvisnosti in homogenosti

3.4.1 Neodvisnost

TEORIJA ZA 2×2 TABELE

Imamo neko populacijo Ω, ki jo razdelimo na dva načina:

$$\Omega = A_1 \dot\cup A_2$$

$$\Omega = B_1 \dot\cup B_2$$

Potem nas zanimajo sledeče verjetnosti: $P(A_i \cap B_j) = p_{ij} > 0$. Pozitivnost le teh je privzeta za naš test.

Predpostavka neodvisnosti:

$$P(A_i \cap B_j) = P(A_i) \cdot P(B_j)$$

Zdaj lahko zapišemo množico parametrov:

$$\Theta = \left\{ \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \middle| \; 0 < p_{ij} \; ; \; \sum p_{ij} = 1 \right\} \; ; \; \dim \Theta = 3$$

Dimenzijo Θ dobimo tako: matrika verjetnosti je dimenzije 4; od tega moramo odšteti število vezi. Imamo samo eno vez in sicer $\sum p_{ij} = 1$.
Zato je $\dim \Theta = 4 - 1 = 3$.

Kaj pa ničelna hipoteza?
$H_0 : p_{ij} = p_i \cdot q_j$, kjer je $p_i = P(A_i), q_j = P(B_j)$.
Logično je, da morata veljati sledeči 2 vezi: $p_1 + p_2 = 1$ in $q_1 + q_2 = 1$
Potem je:

$$H_0 = \left\{ \begin{bmatrix} p_1 \cdot q_1 & p_1 \cdot q_2 \\ p_2 \cdot q_1 & p_2 \cdot q_2 \end{bmatrix} \middle| \; 0 < p_i \; ; \; 0 < q_j \; ; \; p_1 + p_2 = q_1 + q_2 = 1 \right\} \; ; \; \dim H_0 = 2$$

Dimenzijo H_0 dobimo tako: matrika produktov verjetnosti je dimenzije 4; od tega moramo odšteti število vezi, ki je tokrat 2. Zato je $\dim H_0 = 4 - 2 = 2$.
Potem lahko s pomočjo tega izračunamo prostostne stopnje 3.5:

$$df = \dim \Theta - \dim H_0 = 3 - 2 = 1.$$

SLUČAJNI EKSPERIMENT

Za slučajni eksperiment (:= X) naključno izberemo $\omega \in \Omega$ in definiramo: $X(\omega) = (i, j)$, če je $\omega \in A_i \cap B_j$. To n-krat ponovimo.

3.4.2 Homogenost

TEORIJA ZA 2×2 TABELE

Podobno kot pri testiranju neodvisnosti opazujemo neko populacijo Ω (na primer „dekleta“), na njej pa proučujemo dve dekompoziciji, ki ju identificiramo s slučajnima spremenljivkama:

$$W_1 : \Omega \to \{1, 2\}$$

$$W_2 : \Omega \to \{1, 2\}$$

Na primer, za dekle $\omega \in \Omega$ je $W_1(\omega) = 1$ če se dekle pred ogledalom, ki zožuje postavo, počuti samozavestno, oziroma $W_1(\omega) = 2$, če se ne počuti samozavestno. Podobno je $W_2(\omega) = 1$, če se dekle pred običajnim ogledalom počuti samozavestno in $W_2(\omega) = 2$, če se ne.

Kasneje bomo pri hipotezi za Fisherjev eksaktni test uporabljali oznaki:

$$\pi_1 = P(W_1 = 1)$$

$$\pi_2 = P(W_2 = 1)$$

Konstruiramo SLUČAJNI EKSPERIMENT (:= X):

Naključno izberemo $\omega \in \Omega$ (dekle) ter neodvisno še $j \in \{1, 2\}$ (vrsto ogledala). Naj izbor ogledala opisuje slučaja spremenljivka J z možnima vrednostma 1 in 2. Postavimo:

$$X = (W_j(\omega), j) \in \{1, 2\} \times \{1, 2\}$$

Zanima nas sledeča verjetnost:

$$p_{ij} := P(X = (i, j)) = P(J = j \text{ in } W_j = i) = P(W_j = i) \cdot P(J = j),$$

kjer je zadnji enačaj posledica neodvisnosti obeh izborov.
Jasno je, da veljata vezi

$$p_{1j} + p_{2j} = P(J = j) \text{ za } j = 1, 2.$$

Pa si poglejmo sedaj množico parametrov:

$$\Theta = \left\{ \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \middle| \; p_{11} + p_{21} = P(J = 1) \text{ in } p_{12} + p_{22} = P(J = 2) \right\}.$$

Izračun dimenzije je podoben kot pri testu neodvisnosti v razdelku 3.4.1: matrika verjetnosti je zopet dimenzije 4; odštejemo še obe vezi in dobimo, da je $\dim \Theta = 4 - 2 = 2$.

Kaj pa ničelna hipoteza?
H_0 : obstajata števili q_1 in q_2 iz intervala $(0, 1)$, za kateri je $p_{ij} = q_i \cdot P(J = j)$, pri čemer mora veljati $q_1 + q_2 = 1$. Verjetnosti $P(J = 1)$ in $P(J = 2)$ sta seveda konstantni, saj način vzorčenja izberemo mi sami. Zato je dimenzija ničelne hipoteze enaka 1. S pomočjo tega lahko zopet izračunamo prostostne stopnje 3.5:

$$df = \dim \Theta - \dim H_0 = 2 - 1 = 1.$$

Glede na to, da velja $p_{ij} = P(W_j = i) \cdot P(J = j)$, iz enakosti $p_{ij} = q_i \cdot P(J = j)$ sledi $P(W_1 = i) = P(W_2 = i) = q_i$. Torej lahko ničelno hipotezo ekvivalentno opišemo z enakostjo

$$\pi_1 = \pi_2.$$

3.5 Prostostne stopnje

Kako izračunamo prostostne stopnje za test neodvisnosti in homogenosti, sem pokazala že v 3.4.1 in 3.4.2. Kaj pa prostostne stopnje v resnici so?

To je število neodvisnih vrednosti slučajne spremenljivke, ki se lahko v statističnih izračunih spreminjajo.

Prostostna stopnja je parameter sistema, ki mora imeti naslednje lastnosti:

- sistem je s pomočjo določitve parametra enolično določen
- kadar en parameter izpustimo, sistem ni več enolično določen
- vsak parameter lahko spremenimo, ne da bi pri tem spremenili drug parameter.

Formula za izračun prostostne stopnje, ki smo jo sicer že videli:

$$df_0 = \dim \Theta - \dim H_0.$$

Naj na tem mestu poudarim še sledeče: ne glede na to, da se je dim Θ za test neodvisnosti razlikovala od dimenzije za test homogenosti in da sta se testa razlikovala po številu vezi, so bile končne prostostne stopnje za oba testa enake. To vedno drži in velja tudi za $r \times c$ tabele, ne le za 2×2 tabele. To je tudi eden izmed razlogov, zakaj sta testa računsko identična.

3.6 Konstrukcija testne statistike

Pričakovana frekvenca celice (i,j) (označujemo jo z ε_{ij}) je frekvenca, ki jo pričakujemo v (i,j)-ti celici kontingenčne tabele, če ničelna hipoteza drži. Opažena frekvenca celice (i,j) (označujemo jo z o_{ij}) pa je frekvenca, ki smo jo opazili v tej celici v slučajnem eksperimentu. Merimo lahko razlike med opaženimi in pričakovanimi frekvencami v kontingenčni tabeli.

Pri veljavnosti H_0 pričakujemo $o_{ij} \approx \varepsilon_{ij}$ za vse celice.
Če pa velja H_1, pričakujemo, da o_{ij} bistveno odstopa od ε_{ij} vsaj za eno celico.

To lahko uporabimo za konstrukcijo testne statistike: $T = o_{ij} - \varepsilon_{ij}$, kritična vrednost $c = 0$. T pri tem meri odstopanje od ničelne hipoteze.
Ob končanem testu imamo samo dve možnosti: ali našo ničelno hipotezo potrdimo ali jo ovržemo. Pri tem lahko sprejmemo pravilno odločitev, lahko pa naredimo eno izmed dveh napak, ki sem jih omenila že v 3.1.

4 χ^2 testa neodvisnosti in homogenosti

4.1 Računski del

Da sploh lahko uporabljamo hi-kvadrat test, moramo prej zadostiti trem predpostavkam:

- nominalni podatki (frekvence) vključeni v analizo morajo biti podani za dve medsebojno izključujoči kategoriji.
- vzorčenje izvajamo s ponavljanjem (vračanjem) ali pa je množica, v kateri izvajamo vzorčenje, neskončna
- opažene frekvence za vsako celico v kontingenčni tabeli morajo biti velikosti vsaj 5.

Ker sta χ^2 test neodvisnosti in χ^2 test homogenosti računsko identična, vas bom vodila le skozi izračune za test neodvisnosti. Poglejmo si primer (poudarjam, da so številke v primeru izmišljene in niso pridobljene s podatki iz testov).

Imamo 200 študentov FMF za vzorec - 100 punc, 100 fantov. Vprašamo jih, katero izmed humorističnih serij raje gledajo, Prijatelje (Friends - to bom skrajšala v FR) ali Oh ta sedemdeseta (That 70's show - to pa bom skrajšala kar v 70).
Dobljene podatke vpišemo v tabelo in se vprašamo: Ali dobljeni podatki kažejo na to, da je priljubljenost teh dveh humorističnih serij odvisna od spola gledalcev?

	FR	70	Vsote vrstic
fantje	30	70	100
punce	60	40	100
Vsote stolpcev	90	110	200

Tabela 4: primer: serije

Poglejmo, kako se izračuna vrednost naše testne statistike χ^2:

Opažene frekvence (o - observed) so frekvence, ki smo jih opazili pri poskusu, torej frekvence iz prejšnje tabele. Z ε pa označujemo pričakovane frekvence (ε - expected). Kako pa dobimo te?

Formula za izračun: $\varepsilon_{ij} = \dfrac{O_{i\bullet}O_{\bullet j}}{n}$

Poglejmo si to na naših številkah:
$celica_{11} = \frac{100*90}{200} = 45$
$celica_{22} = \frac{100*110}{200} = 55$

Ko imamo o in ε, kako izračunamo vrednost χ^2?
To storimo s pomočjo sledeče formule:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} (\frac{(O_{ij} - \varepsilon_{ij})^2}{\varepsilon_{ij}})$$

Delne izračune imamo že kar v tabeli 5, tako da če pogledamo: 3. stolpec vsebuje razlike med o in ε, 4. stolpec vsebuje kvadrat te razlike, kar je ravno števec naše formule za χ^2, v zadnjem stolpcu pa kvadrat te razlike delimo še s pričakovano frekvenco, da dobimo ravno ulomek v naši formuli. Potem upoštevamo še vsoti (ne pozabimo na to, da pri nas velja $r = c = 2$), seštejemo elemente zadnjega stolpca in dobimo našo vrednost.

	o_{ij}	ε_{ij}	$o_{ij} - \varepsilon_{ij}$	$o_{ij} - \varepsilon_{ij})^2$	$\frac{(o_{ij}-\varepsilon_{ij})^2}{\varepsilon_{ij}}$
$celica_{11}$	30	45	-15	225	5.00
$celica_{12}$	70	55	15	225	4.09
$celica_{21}$	60	45	15	225	5.00
$celica_{22}$	40	55	-15	225	4.09
Vsote:	200	200	0		$\chi^2 = 18.18$

Tabela 5: izračuni

Vrednost naše testne statistike je torej $\chi^2 = 18.18$.

4.2 Dobljeni rezultati

Kakšni morajo biti dobljeni rezultati?

- Vsoti opaženih in pričakovanih frekvenc morata vedno biti identični.
- Pomembno je tudi to, da ko seštejemo tretji stolpec tabele 5, torej ko seštejemo razlike med pričakovanimi in opaženimi frekvencami, moramo dobiti 0.
- Ker vse razlike med o in ε v četrtem stolpcu tabele 5 kvadriramo, mora biti vsota 5. stolpca, ki je tudi vrednost χ^2, vedno nenegativna (= 0 bo samo v primeru, ko bo veljalo $o = \varepsilon$ za vse celice naše tabele).

Če katerakoli od teh točk ne drži, to pomeni, da je prišlo do napake.

4.3 Interpretacija rezultatov

Dobljena vrednost $\chi^2 = 18.18$ se vrednoti s pomočjo tabele 6, v kateri najdemo vrednosti χ^2 porazdelitve za različne stopnje značilnosti in prostostne stopnje. Te vrednosti so kritične vrednosti C za naš test.

Prostostne stopnje (df)	χ^2 **vrednosti**									
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.63
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09
Stopnja značilnosti	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01

Tabela 6: χ^2 porazdelitev

Da znamo preveriti, kaj je rešitev, moramo najprej izračunati prostostne stopnje (3.5). Najlažje je to izračunati ne po formuli iz 3.5, temveč po poenostavljeni formuli za kontingenčne tabele:

$$df = (r-1)(c-1) = (\text{število stolpcev} - 1)(\text{število vrstic} - 1)$$

Ker se ukvarjamo samo z 2×2 kontingenčnimi tabelami, je df za nas vedno =1. To pomeni, da bomo ves čas gledali le prvo vrstico v tabeli 6.

Naslednja stvar, ki si jo moramo izbrati, je α = stopnja značilnosti. Omenjala sem jo že v 3.1 poglavju, ampak tokrat si jo bomo pogledali z nekoliko drugačnega vidika: α pravzaprav predstavlja verjetnost, da so odstopanja med dejanskimi in pričakovanimi frekvencami zgolj slučajna.
V jeziku napak: to je največja dovoljena napaka tipa I. Na tem mestu pa ne smemo pozabiti na komplementarnost: če postavimo zgornjo mejo na napako tipa I, ne moremo postaviti zgornje meje na napako tipa II.
Najbolj pogosti in smiselni α so tisti, ki so v tabeli 6 pobarvani zeleno ... Torej 10%, 5% in 1%.

Pa si poglejmo to na našem primeru:
$\chi^2 = 18.18$
df = 1

Za $\alpha = 5\%$ preberemo iz tabele 6 : $\chi^2 = 3.84$
Ker je naša izračunana vrednost testne statistike večja od vrednosti v tabeli oz. kritične vrednosti, to pomeni, da pristanemo v območju zavrženja B. Našo ničelno hipotezo zato zavrnemo pri stopnji značilnosti 0.05.

Podobno za $\alpha = 1\%$ preberemo iz tabele 6 : $\chi^2 = 6.63$
Tudi to je manj kot 18.18, zato lahko ničelno hipotezo zavrnemo tudi pri stopnji značilnosti 0.01.
Po zavrnitvi ničelne hipoteze pa lahko rečemo še, da so podatki konsistentni z alternativno hipotezo.

5 Fisherjev eksaktni test neodvisnosti in homogenosti

Fisherjev eksaktni test se priporoča v vseh primerih, ko je vzorec majhen, torej v vseh primerih, ko je $n < 20$. V resnici pa je test veljaven za vzorce vseh velikosti.

Da sploh lahko uporabljamo Fisherjev eksaktni test, moramo prej zadostiti trem predpostavkam:

- nominalni podatki (frekvence) vključeni v analizo morajo biti podani za dve medsebojno izključujoči kategoriji
- vzorčenje izvajamo s ponavljanjem (vračanjem) ali pa je množica, v kateri izvajamo vzorčenje, neskončna.
- vsote vrstic in stolpcev v tabeli so predhodno določene s strani raziskovalca oz. osebe, ki bo opravljala test.

Pri tem se je treba zavedati, da lahko zadnji predpostavki v praksi le redko zadostimo. A vendar zaradi te predpostavke lahko proučujemo le Fisherjev eksaktni test homogenosti, saj pri neodvisnosti vrstični vsoti nista fiksni.

5.1 Ničelna in alternativna hipoteza

Ničelna hipoteza: $H_0 : \pi_1 = \pi_2$
Dvostranska alternativna hipoteza: $H_1 : \pi_1 \neq \pi_2$
Enostranski alternativni hipotezi: $H_1 : \pi_1 > \pi_2$ ali $H_1 : \pi_1 < \pi_2$.
Dvostranska alternativna hipoteza je v resnici unija enostranskih.
Spet lahko ničelno in alternativno hipotezo postavimo le pod pogojem, da so vsote stolpcev in vrstic fiksirane. Če torej zadostimo zadnji predpostavki za Fisherjev test, nimamo kaj skrbeti, sicer pa moramo paziti na to, preden si hipoteze postavimo.

5.2 Računski del

Tabelo bomo pri Fisherjevem testu prikazovali nekoliko drugače, kot smo jo prej pri hi-kvadrat testu:

	A_1	A_2	
B_1	a	b	a+b
B_2	c	d	c+d
	a+c	b+d	n

Tabela 7: tabela za Fisherjev eksaktni test

Fisher je pokazal, da je verjetnost kakršnega koli takega nabora opaženih frekvenc dana s hipergeometrijsko porazdelitvijo.

Enačba za izračun verjetnosti:

$$P = \frac{\binom{a+c}{a}\binom{b+d}{b}}{\binom{n}{a+b}} = \frac{(a+c)!(b+d)!(a+b)!(c+d)!}{n!a!b!c!d!}$$

Pa si poglejmo to na primeru:
Imamo neko manjšo populacijo fantov in deklet, ki so razdeljeni glede na to, ali imajo rajši mačke ali pse. Recimo, da je naša hipoteza, da je delež ljudi, ki preferirajo mačke, večji pri dekletih kot pri fantih in želimo to testirati.

$$\pi_1 = \text{verjetnost, da naključno izbran fant preferira mačko.}$$

$$\pi_2 = \text{verjetnost, da naključno izbrano dekle preferira mačko.}$$

Pričakovana vrednost (matematično upanje) „elementa“ a v tabeli je $(a + c) \cdot \pi_1$, pričakovana vrednost „elementa“ b pa $(b + d) \cdot \pi_2$.

Test bi v tabeli izgledal tako:

	fantje	dekleta	
mačka	a = 1	b = 5	a+b = 6
pes	c = 5	d = 1	c+d = 6
	a+c = 6	b+d = 6	n = 12

Tabela 8: primer: hišni ljubljenčki

Poglejmo izračun verjetnosti:
$$P_1 = \frac{\binom{6}{1}\binom{6}{5}}{\binom{12}{6}} = \frac{[\frac{6!}{1!5!}][\frac{6!}{5!1!}]}{\frac{12!}{6!6!}} = 0.039$$
Po drugi enačbi:
$$P_1 = \frac{6!6!6!6!}{12!1!5!5!1!} = 0.039$$

Ta formula nam da eksaktno hipergeometrijsko verjetnost, da bomo opazili ravno ta nabor frekvenc, ravno to porazdelitev podatkov, pri fiksiranih stolpcih in vrsticah ter ničelni hipotezi, da je enako verjetno, da fant in dekle preferirata mačko oz. psa. Ker formula ne vsebuje verjetnosti, da moški preferira mačko oz. psa, ali verjetnosti, da ženska preferira mačko oz. psa, dobimo eksaktno hipergeometrijsko porazdelitev tudi, če ta verjetnost za moškega in žensko ni enaka. Zatorej lahko izračunamo eksaktno verjetnost kateregakoli nabora opaženih frekvenc za 12 fantov in deklet v štirih celicah naše kontingenčne tabele.

Ampak Fisher je pokazal, da lahko za izračun stopnje značilnosti naših podatkov upoštevamo le primere, kjer so vrstične vsote in vsote stolpcev enake kot v opaženem naboru frekvenc, in med temi le tiste primere, kjer je porazdelitev podatkov v tabeli enako ali bolj ekstremna kot opažena porazdelitev podatkov v tabeli.

V našem primeru bi bil samo en nabor rezultatov bolj ekstremen, in sicer sledeči:

	fantje	dekleta	
mačka	a = 0	b = 6	a+b = 6
pes	c = 6	d = 0	c+d = 6
	a+c = 6	b+d = 6	n = 12

Tabela 9: primer: hišni ljubljenčki - ekstrem

Pa izračunajmo verjetnosti:

$$P_2 = \frac{\binom{6}{0}\binom{6}{0}}{\binom{12}{6}} = \frac{[\frac{6!}{0!6!}][\frac{6!}{6!0!}]}{\frac{12!}{6!6!}} = 0.001$$

Oziroma:

$$P_2 = \frac{6!6!6!6!}{12!0!6!6!0!} = 0.001$$

Za izračun stopnje značilnosti, tj. verjetnosti, da bo naš nabor frekvenc vedno enak oziroma bolj ekstremen, če ničelna hipoteza drži, moramo sešteti P_1 in P_2. Torej:

$P_T = P_1 + P_2 = 0.039 + 0.001 = 0.04$ (P_T imenujemo družna verjetnost)

5.3 Dobljeni rezultati

Kakšni morajo biti dobljeni rezultati?

- Če je bila alternativna hipoteza dvostranska:
  potem mora biti vrednost $P_T \leq \frac{\alpha}{2}$
- Če je bila alternativna hipoteza enostranska:
  v tem primeru mora biti opažen nabor frekvenc konsistenten z enostransko hipotezo, poleg tega pa mora biti $P_T \leq \alpha$.

Če vzamemo $\alpha = 0.05$, potem naša dvostranska hipoteza ni potrjena, saj je $P_T = 0.04 > \frac{0.05}{2} = 0.025$.

Enostranska alternativna hipoteza $H_1 : \pi_1 < \pi_2$ je potrjena pri stopnji značilnosti 0.05, saj so dobljeni podatki konsistentni s hipotezo in $P_T = 0.04 \leq 0.05 = \alpha$. Od tu že jasno vidimo, da H_1 ni podprta pri $\alpha = 0.01$, saj je P_T večja od te vrednosti.

Enostranska alternativna hipoteza $H_1 : \pi_1 > \pi_2$ ne more biti potrjena pri nobeni stopnji značilnosti, saj podatki z njo niso konsistentni.

5.4 Razlaga

Zakaj mora biti formula za ocenjevanje dobljenih rezultatov v primeru dvostranske hipoteze drugačna kot v primeru enostranske?
Poglejmo si vse možne izide za opažene frekvence pri pogoju, da so vse robne vsote enake 6 in je naš n = 12. Ti izidi so zbrani v tabeli 10.

	fantje	dekleta	
mačka	0	6	6
pes	6	0	6
	6	6	12

IZID 1: P = 0.001

	fantje	dekleta	
mačka	1	5	6
pes	5	1	6
	6	6	12

IZID 2: P = 0.039

	fantje	dekleta	
mačka	2	4	6
pes	4	2	6
	6	6	12

IZID 3: P = 0.243

	fantje	dekleta	
mačka	3	3	6
pes	3	3	6
	6	6	12

IZID 4: P = 0.433

	fantje	dekleta	
mačka	4	2	6
pes	2	4	6
	6	6	12

IZID 5: P = 0.243

	fantje	dekleta	
mačka	5	1	6
pes	1	5	6
	6	6	12

IZID 6: P = 0.039

	fantje	dekleta	
mačka	6	0	6
pes	0	6	6
	6	6	12

IZID 7: P = 0.001

Tabela 10: tabela izidov

Če želimo oceniti $H_1 : \pi_1 < \pi_2$, nas iz tabele 10 zanimata le izid 1 in 2. Njuna družna verjetnost je $P_T = 0.04$, kar je manj kot vrednost $\alpha = 0.05$ (ki v resnici predstavlja najbolj ekstremnih 5 % naše vzorčne porazdelitve v enem izmed dveh repov porazdelitve).
Kot že rečeno, ker je $\alpha > P_T$, lahko to hipotezo potrdimo pri stopnji značilnosti 0.05.

Če pa želimo oceniti $H_1 : \pi_1 \neq \pi_2$, moramo poleg izida 1 in 2 upoštevati še izida 6 in 7, ki sta analogna ekstremna izida v nasprotnih repih hipergeometrijske porazdelitve. Če seštejemo verjetnosti izidov 1, 2, 6 in 7, dobimo $P_T = 0.08$, kar je zdaj verjetnost v obeh repih porazdelitve, da bomo dobili rezultate, ki bodo enaki ali bolj ekstremni od tistih, ki smo jih dobili v našem primeru.
Ker je po formuli, ki drži za enostransko alternativno hipotezo, $P_T > \alpha$, vidimo, da podatki ne podpirajo alternativne hipoteze. Rezultat je torej pravilen.

Če enačbo za enostransko alternativno hipotezo $P_T \leq \alpha$ na tem mestu delimo z 2 na obeh straneh, dobimo začetno enačbo za enostransko alternativno hipotezo, ki smo jo želeli dokazati. Ko bomo P_T delili z dva, bomo v resnici storili le to, da bomo upoštevali le eno stran hipergeometrijske porazdelitve, le en njen rep. To pa seveda lahko storimo, saj vemo, da je ta porazdelitev simetrična.

Če si zdaj za preverjanje rezultatov pogledamo še ocenitev $H_1 : \pi_1 > \pi_2$:
V tem primeru nas zanimata le izid 6 in 7. Združena verjetnost za ta dva izida je 0.04. To zopet primerjamo z α in vidimo, da je $\alpha > P_T$. Kljub temu sicer zadovoljivemu rezultatu pa vidimo, da podatki niso konsistentni s $H_1 : \pi_1 > \pi_2$, zato te hipoteze ne moremo potrditi.

5.5 Fisherjev eksaktni test v praksi: primer „Lady tasting tea“

Eden prvih statističnih testov kadarkoli opravljenih je t. i. „Lady tasting tea“ primer, oziroma v prevodu, primer gospodične, ki poskuša čaj.

Zgodba se začne s čajanko, ki se zaradi preprostega dogodka spremeni v eno najbolj slavnih čajank vseh časov. Tisto popoldne naj bi Fisher nalil skodelico čaja in jo ponudil dekletu, ki je sedelo poleg njega, Dr. B. Muriel Bristol. Dekle je ponudbo zavrnilo, rekoč, da bi raje skodelico, v katero je bilo najprej vlito mleko in šele potem čaj, saj je okus boljši. Fisher je njeno trditev oklical za nesmisel, ona pa mu je zagotavljala, da je razlika več kot opazna. Iz ozadja se je oglasil drug gost čajanke, William Roach, in predlagal, da se gospodično testira.

V kar najkrajšem času so sestavili preprost test, ki naj bi določil, ali gospodična Bristol res prepozna razliko med čajem z mlekom in mlekom s čajem. Fisher ji je podal 8 skodelic. V štirih je bilo mleko nalito prej kot čaj in v štirih kasneje. H. Fairfield Smith, Fisherjev prijatelj, je leta kasneje povedal, da je gospodična „uganila“ vseh 8 skodelic.

Glede na to, da je nekaj let po tem dogodku Fisher predstavil svoj eksaktni test, smo lahko skoraj prepričani, da ga je navdihnila prav gospodična s čajem. Na glavna vprašanja, ki si jih mora statistik postaviti, preden izvede tak test, opozarja [6]:

- *Kaj lahko storimo glede naključnih spremenljivk, kot so temperatura, sladkost in podobno?*
  V idealnem svetu bi bile skodelice identične, če izvzamemo vrstni red nalivanja čaja. V realnosti pa je te naključne spremenljivke nemogoče izkoreninit. Največ, kar lahko storimo, je, da jih minimiziramo z vzpostavitvijo naključnega vrstnega reda skodelic.

- *Koliko skodelic uporabiti? V kakšnem vrstnem redu naj bodo prinešene?*
  Glavna ideja je, da naj bi število in vrstni red skodelic omogočala gospodični priložnost, da dokaže svojo resnično sposobnost ločevanja. Po drugi strani pa naj bi preprečevala morebitnemu prevarantu, da zlahka loči skodelice med sabo.

- *Kakšne zaključke lahko potegnemo iz popolnega rezultata? Kaj lahko zaključimo v primeru, da se je gospodična enkrat ali celo večkrat zmotila?*
  Če gospodična ne bi bila zmožna ločevanja med vrstnima redoma nalivanja, potem mora test delovati tako, da je kar najmanj verjetno, da bo pravilno uganila, če bo ugibala. Podobno: če dejansko ima sposobnost razločevanja, potem je morda nerazumno, da od nje zahtevamo popoln rezultat, brez vsake napake, da bi razločili njeno sposobnost od nekoga, ki preprosto ugiba.

Celoten primer gospodične, ki poskuša čaj, lahko povzamemo s kontingenčno tabelo:

		Dejanski vrstni red nalivanja		
		Najprej čaj	Najprej mleko	
Gospodična pravi	Najprej čaj	a	b	a+b
	Najprej mleko	c	d	c+d
		a+c	b+d	n

Tabela 11: primer: lady tasting tea

n je število skodelic čaja. $a + c$ so skodelice, v katere je bil najprej vlit čaj, $a + b$ pa so skodelice, za katere gospodična trdi, da je bil vanje najprej vlit čaj. Če lahko gospodična dejansko okusi razliko, potem bosta b in c majhna. Če pa razlike v resnici ne okusi, bosta a in c približno enaka.

Zdaj pa recimo, da bomo gospodični ponudili 8 skodelic čaja - 4, v katere je bilo najprej vlito mleko in 4, v katere je bil najprej vlit čaj, in ji to tudi povemo. Skodelice ji prinašamo po nekem naključnem vrstnem redu. Potem velja:

$$a + b = a + c = c + d = b + d = 4$$

To pomeni, da v trenutku, ko določimo eno izmed teh spremenljivk, so ostale tri določene kot

$$b = 4 - a,\ c = 4 - a \text{ in } d = a.$$

V tem primeru bo $a + b = a + c$ ne glede na število skodelic (n), ker bo gospodična vedela, v koliko skodelic je bil najprej nalit čaj.

Sposobnost razločanja skodelic med sabo pa lahko najbolje ugotovimo z naključnim vrstnim redom skodelic. Če nima nobene sposobnosti okušanja razlike, potem naključni vrstni red zagotavlja, da so tiste štiri skodelice, ki jih ona izbere kot „Najprej čaj“, enako verjetne kot katerekoli štiri izmed osmih skodelic.
Možnosti je torej $\binom{8}{4} = 70$. Zaradi najključnega vrstnega reda je vseh 70 možnosti enako verjetnih. Nekdo brez sposobnosti razločanja skodelic ima torej 1/70 možnosti, da ugane.

Izkaže se, da ima a iz naše tabele (torej število pravilno uganjenih skodelic, v katere je bil najprej vlit čaj) hipergeometrijsko porazdelitveno funkcijo. Zatorej veljajo sledeče verjetosti:

Število pravilnih ugibanj	verjetnost
0	1/70
1	16/70
2	36/70
3	16/70
4	1/70

Tabela 12: lady tasting tea - verjetnosti

Pri hipergeometrijski porazdelitvi pa uporabljamo ravno Fisherjev eksaktni test.

6 Primerjava testov

6.1 χ^2 test proti Fisherjevem eksaktnem testu

V realnosti χ^2 porazdelitev zagotavlja le približek eksaktni vzorčni porazdelitvi za kontingenčno tabelo. Približek zato, ker uporabimo χ^2 porazdelitev, ki je zvezna verjetnostna porazdelitev, za aproksimacijo diskretne verjetnostne porazdelitve. Natančnost χ^2 aproksimacije narašča, ko naš vzorec narašča, in razen za primere, ko je vzorec zelo majhen, dobimo odlično aproksimacijo eksaktni vzorčni porazdelitvi.
Problem se pojavi, ko računamo natančne verjetnosti za 2×2 kontingenčno tabelo na zelo majhnem vzorcu. Takrat se kaj zlahka zgodi, da katera izmed frekvenc v kontingenčni tabeli pade pod 5. Predpostavka za χ^2 test takrat pade in izračunani podatki niso več nujno pravilni. V takih primerih uporabimo Fisherjev eksaktni test.

Fisherjev test je v nasprotju s χ^2 testom eksaktni test. Namesto da se zanaša na aproksimacijo, ki v limiti, ko gre velikost vzorca proti neskončno, postane točna, je z njim mogoče natančno izračunati stopnjo značilnosti in odstopanje podatkov od ničelne hipoteze. Fisherjev test je neobčutljiv na velikost frekvenc v celicah. V primeru velikih vzorcev je računski proces zelo obsežen, skoraj neobvladljiv. To pa so situacije, kjer raje uporabimo hi-kvadrat test.

6.2 Primerjava testov na primeru

Da bomo jasno videli razliko med obema testoma, bomo opravili oba testa na enem primeru. Obravnavali bomo primer Hišni ljubljenčki 8, saj imamo vse izračune za Fisherjev eksaktni test že opravljene. Na tem istem primeru bomo sedaj opravili tudi χ^2 test.
Rezultati hi-kvadrat testa:

	O_{ij}	ε_{ij}	$O_{ij} - \varepsilon_{ij}$	$(O_{ij} - \varepsilon_{ij})^2$	$\frac{(O_{ij}-\varepsilon_{ij})^2}{\varepsilon_{ij}}$
$celica_{11}$	1	3	-2	4	1.33
$celica_{12}$	5	3	2	4	1.33
$celica_{21}$	5	3	2	4	1.33
$celica_{22}$	1	3	-2	4	1.33
Vsote:	12	12	0		$\chi^2 = 5.31$

Tabela 13: Hišni ljubljenčki - hi-kvadrat test

Spet pogledamo v tabelo 6:

- $H_1 : \pi_1 \neq \pi_2$
  pri $\alpha = 0.05$ je tabelirna vrednost $\chi^2_{.05} = 3.84$. Ker je naša izračunana vrednost večja od tabelirne vrednosti, to pomeni, da lahko potrdimo $H_1 : \pi_1 \neq \pi_2$ pri stopnji značilnosti 0.05.
  Ne moremo pa je podpreti pri stopnji značilnosti 0.01, saj je tabelirna vrednost $\chi^2_{.01} = 6.63$ večja od naše izračunane vrednosti.

- $H_1 : \pi_1 < \pi_2$
  pri $\alpha = 0.05$ je tabelirna vrednost $\chi^2_{.90} = 2.71$. Naša izračunana vrednost je spet večja od tabelirne vrednosti, zato lahko to hipotezo potrdimo pri stopnji značilnosti 0.05.
  Pri $\alpha = 0.01$ pa je tabelirna vrednost $\chi^2_{.98} = 5.43$, ki pa je le za las večja od naše izračunane vrednosti. Tako tudi $H_1 : \pi_1 < \pi_2$ ni podprta pri stopnji značilnosti 0.01.

Primerjava rezultatov:
Če uporabimo Fisherjev eksaktni test, potem $H_1 : \pi_1 \neq \pi_2$ ni podprta pri 0.05, v primeru uporabe χ^2 testa pa je.
Oba testa omogočata raziskovalcu, da potrdi ničelno hipotezo pri 0.05, če uporabljamo $H_1 : \pi_1 < \pi_2$.
Če pa si pogledamo situacijo pri 0.01, opazimo, da je v primeru uporabe χ^2 testa hipoteza blizu tega, da bi bila potrjena, v primeru uporabe Fisherjevega eksaktnega testa pa je do tega zelo daleč.

Primerjava kaže na to, da v primeru majhnega vzorca χ^2-kvadrat test podcenjuje dejansko verjetnost povezano z opaženimi frekvencami. Posledično se s tem poveča verjetnost, da bomo zagrešili napako tipa I.

Literatura

[1] David J. Sheskin, *Handbook of parametric and non-parametric statistical procedures - 3ed*, Chap-man & Hall / CRC, (2004).

[2] A. Martin Andres, I. Herranz Tejedor, A. Silva Mato, *The Wilcoxon, Spearman, Fisher,*χ^2*-, Student and Pearson Tests and* 2×2 *Tables*, Journal of the Royal Statistical Society. Series D (The Statistician) **44** (1995), no. 4, strani 441-450

[3] Erich Leo Lehmann, *Testing statistical hypotheses - 3ed*, Springer Science+Business Media, Inc., (2005).

[4] F. Yates, *Test of Significance for* 2×2 *Contingency Tables*, Journal of the Royal Statistical Society, Series A (General) **147** (1984), no.3, strani 426-463.

[5] Douglas S. Safer and Zhiyi Zhang, *Introductory Statistics - v. 1.0*, [ogled 26.5.2013], dostopno na http://2012books.lardbucket.org/.

[6] Lt. Col. Rod Sturdivant, Ph.D., U.S. Military Academy, Dept. of Mathematical Sciences, *Lady tasting tea*, [ogled 29.4.2012], dostopno na http://www.dean.usma.edu/math/people/sturdivant/images/MA376/dater/ladytea.pdf.