

Jakaumien yhteensopivuuden testaaminen

Esimerkki 1

Pekka heitti noppaa testatakseen, tuleeko tulokseksi eri silmälukuja yhtä usein. Hän heitti noppaa 120 kertaa ja sai eri silmälukuja seuraavat määrät:

Silmäluku	1	2	3	4	5	6
Frekvenssi	14	19	30	17	16	24

Avaa OpenOffice-tiedosto →

Vertaa eri silmälukujen esiintymiskertojen eroa tilanteeseen, jossa silmäluvut esiintyisivät yhtä usein. Näyttääkö siltä, että Pekan noppa antaisi eri silmälukuja yhtä usein?

Ratkaisu

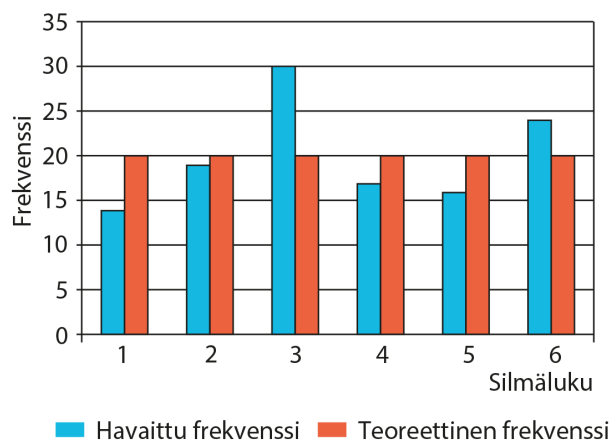
Jos 120 nopanheittokertaa olisi tuottanut tasaisesti eri silmälukuja, olisi jokainen silmäluku esiintynyt heittosarjassa

$$\frac{120}{6} = 20 \text{ kertaa.}$$

Lasketaan, kuinka paljon Pekan saamat eri silmälukujen esiintymismäärät poikkeavat teoreettisesta 20 kerrasta.

Silmäluku	Havaittu frekvenssi	Teoreettinen frekvenssi	Erotus
1	14	20	-6
2	19	20	-1
3	30	20	10
4	17	20	-3
5	16	20	-4
6	24	20	4

Eri silmälukujen esiintyminen 120 heitossa



Vaikka joidenkin silmälukujen kohdalla havaittu frekvenssi poikkeaa selvästi 20:stä, poikkeamat selittyvät mitä todennäköisimmin sattumalla.

χ^2 -yhteensopivuustesti

χ^2 -yhteensopivuustestillä pyritään selvittämään,

noudattaako havaittu jakauma jotakin oletettua teoreettista jakaumaa vai ei. Tällöin havaittujen frekvenssien ja vastaavien teoreettisten frekvenssien arvoista lasketaan testisuure χ^2 . Testisuuren arvo lasketaan seuraavasta kaavasta.

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i},$$

missä

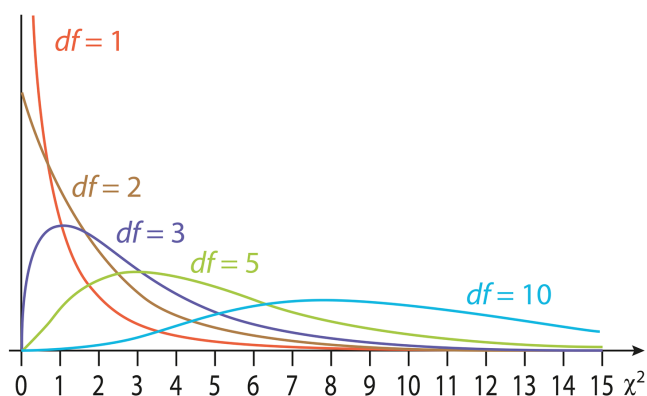
o_i on i :n luokan havaittu frekvenssi

e_i on i :n luokan teoreettinen frekvenssi

k on havaintoluokkien lukumäärä.

χ^2 luetaan khiin neliö tai khii toiseen.

Mikäli havaittu jakauma noudattaa oletettua teoreettista jakaumaa, testisuure χ^2 noudattaa χ^2 -jakaumaa. χ^2 -jakauman muoto riippuu vapausasteiden määrästä $k - 1$. Vapausasteiden määrä on siis yhtä pienempi kuin havaintoluokkien määrä. Vapausasteiden määrälle käytetään lyhennettä df (degrees of freedom).



Mitä enemmän havaitut frekvenssit poikkeavat teoreettisista frekvensseistä, sitä suuremman arvon testisuure χ^2 saa. Sitä arvoa, joka rajaa jakaumasta valittua luottamustasoa vastaavan osan, kutsutaan luottamustasoa vastaavaksi kriittiseksi arvoksi. Mikäli testisuureen arvo

ylittää kriittisen arvon, voidaan valitulla luottamustasolla päätellä, että havaittu jakauma ei noudata teoreettista jakaumaa.

Monien laskentaohjelmien tilastotoiminnoissa on myös χ^2 -testitoiminto, joka laskee testisuureen arvon ja antaa lisäksi riskitason p eli todennäköisyyden sille, että havaitun jakauman ja teoreettisen jakauman erot johtuvat sattumasta. Esimerkiksi 95 prosentin luottamustasolla riskitaso saa olla enintään 5 %, jotta havaitun jakauman ja teoreettisen jakauman eroa pidetään tilastollisesti merkitsevänä.

χ^2 -yhteensopivuustestiä voidaan käyttää, jos kaikki teoreettiset frekvenssit ovat suurempia kuin 1 ja enintään 20 % teoreettisista frekvensseistä on pienempiä kuin 5.

Avaa appletti →

Esimerkki 2

a) Laske testisuureen χ^2 arvo kaavan ja esimerkissä 1 esitettyjen havaittujen ja teoreettisten frekvenssien perusteella.

Silmäluku	Havaittu frekvenssi	Teoreettinen frekvenssi
1	14	20
2	19	20
3	30	20
4	17	20
5	16	20
6	24	20

Avaa OpenOffice-tiedosto →

Esimerkki 3: Kuvien arvo $\chi^2 = 8,9$ on kriittistä arvoa pienempi,

joten 120 nopanheitolla saadun jakauman voidaan katsoa noudattavan tasajakaumaa, jossa kaikki silmäluvut startanneiden hevosten saavuttamat voitot jakautuivat esiintyvät yhtä usein.

lähöradan mukaan seuraavasti:

	Voittojen lukumäärä	$(o_i - e_i)$	$\frac{(o_i - e_i)^2}{e_i}$
Rata 1	342	-6	$\frac{36}{20} = 1,8$
Rata 2	324	-6	$\frac{36}{20} = 1,8$
Rata 3	342	-6	$\frac{36}{20} = 1,8$
Rata 4	336	-6	$\frac{36}{20} = 1,8$
Rata 5	375	15	$\frac{225}{20} = 11,25$
Rata 6	291	-9	$\frac{81}{20} = 4,05$
Rata 7	219	-11	$\frac{121}{20} = 6,05$
Rata 8	171	-15	$\frac{225}{20} = 11,25$

merkittävästi. Voidaan siis luottaa, että noppa antaa eri silmäluokkia yhtä todennäköisesti.

Avaa OpenOffice-tiedosto

→ staus

Testisuureen χ^2 arvoksi saadaan

Testaa tilastollisesti, onko lähöradalla vaikutusta hevosen

a) $\chi^2 = 8,9$ on kriittistä arvoa pienempi, joten 120 nopanheitolla saadun jakauman voidaan katsoa noudattavan tasajakaumaa, jossa kaikki silmäluvut startanneiden hevosten saavuttamat voitot jakautuivat esiintyvät yhtä usein.

$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 1,8 + 0,05 + 5 + 0,45 + 0,8 + 0,8 = 8,9$.

Ratkaisu

Mitä enemmän havaitut frekvenssit o_i poikkeavat teoreettisista frekvensseistä e_i , sitä suuremman arvon jakaumasta, jossa voittoa hevosia löytyisi yhtä paljon eri

lähöradoilta, lasketaan tasajakaumaa noudattavat

teoreettiset frekvenssit.

b) Tapa 1

2400
Nopanhito
8300

ttinen arvo
ja
aan

Тара 1

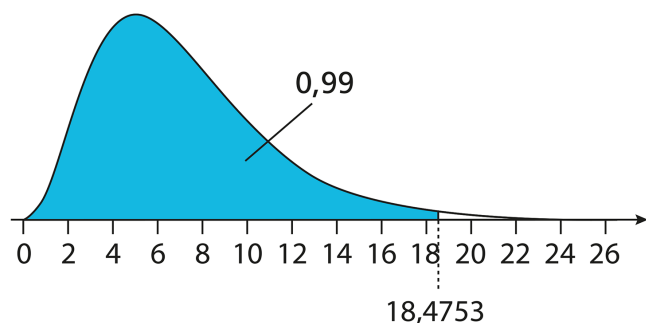
Lasketaan testisuureen χ^2 arvo havaittujen ja teoreettisten frekvenssien avulla.

$$\begin{aligned} \chi^2 &= \sum_{i=1}^8 \frac{(o_i - e_i)^2}{e_i} = \frac{(342 - 300)^2}{300} + \frac{(324 - 300)^2}{300} + \\ &\frac{(342 - 300)^2}{300} + \frac{(336 - 300)^2}{300} + \frac{(375 - 300)^2}{300} + \\ &\frac{(291 - 300)^2}{300} + \frac{(219 - 300)^2}{300} + \frac{(171 - 300)^2}{300} = 114,36 \end{aligned}$$

Eturivistä lähtee kahdeksan hevosta, joten mahdollisia voittajia eturivissä on kahdeksan ($k = 8$). Saatua testisuureen arvoa verrataan χ^2 -jakaumaan, jossa vapausasteiden määrä on $k - 1 = 8 - 1 = 7$.



Määritetään taulukkokirjan tai laskentaohjelman avulla kriittinen arvo χ^2 -jakaumalle, kun vapausasteiden määrä on 7 ja luottamustaso on 0,99. Kriittiseksi arvoksi saadaan 18,4753.



Laskettu testisuureen arvo 114,36 ylittää selvästi kriittisen arvon 18,4753. Tulos tarkoittaa sitä, että voittojen määrät eturivin eri lähtöradoilta poikkeavat tilastollisesti merkitsevästi tasaisesta jakaumasta. Voidaan sanoa, että eturivin kahdeksan lähtörataa eivät ole yhtä hyviä voittamisen kannalta.

Tapa 2

Käytetään laskentaohjelman χ^2 -testitoimintoa. Valitaan laskentaohjelmassa rivien määräksi 8 ja syötetään rivien havaitut ja teoreettiset frekvenssit laskentaohjelman taulukkopohjaan.

Laskentaohjelma antaa testisuureen arvon $\chi^2 = 114,36$ ja testin riskitasoksi $p \approx 0 = 0 \%$.

Riskitason mukaan on hyvin pieni todennäköisyys sille, että havaittujen ja teoreettisten frekvenssien erot johtuisivat sattumasta. Koska riskitaso on pienempi kuin 5 %, havaittu jakauma poikkeaa tasajakaumasta tilastollisesti merkitsevästi. Lähtöradalla on siis vaikutusta voittamiseen.

Tilastollisessa päättelyssä riskitaso on aina suurempi kuin 0.
Esimerkin tilanteessa riskitaso on niin pieni, että osa ohjelmista pyöristää sen arvoon 0.

Vastaus

Lähtöradalla on vaikutusta menestymiseen.

Esimerkki 4

Edellisenä vuonna pakollisen kemian kurssin arvosanat jakautuivat seuraavasti:

Arvosana	Osuus
4	3 %
5	10 %
6	24 %
7	27 %
8	19 %

9	11 %
10	6 %

Avaa OpenOffice-tiedosto

→

Seuraavana vuonna saman kurssin viiteen ryhmään osallistui yhteensä 130 opiskelijaa, jotka saivat arvosanoja seuraavat määrät:

Arvosana	Lukumäärä
4	6
5	15
6	30
7	26
8	23
9	18
10	12
Yhteensä	130

Testaa 5 prosentin riskitasolla, onko arvosanajakauma muuttunut aiemmasta.

Ratkaisu

Verrataan annettuja arvosanoja edellisen vuoden arvosanajakaumaan.

Lasketaan edellisen vuoden arvosanajakauman mukaiset teoreettiset frekvenssit, joiden mukaan 130 opiskelijan arvosanat jakautuisivat täsmälleen edellistä vuotta vastaavasti.

Arvosana	Frekvenssi	Teoreettinen frekvenssi
----------	------------	-------------------------

4	6	$0,03 \cdot 130 = 3,9$
5	15	$0,1 \cdot 130 = 13,0$
6	30	$0,24 \cdot 130 = 31,2$
7	26	$0,27 \cdot 130 = 35,1$
8	23	$0,19 \cdot 130 = 24,7$
9	18	$0,11 \cdot 130 = 14,3$
10	12	$0,06 \cdot 130 = 7,8$
Yhteensä	130	130

Tapa 1

Lasketaan testisuureen χ^2 arvo havaittujen ja teoreettisten frekvenssien avulla.

$$\chi^2 = \sum_{i=1}^7 \frac{(o_i - e_i)^2}{e_i} = \frac{(6 - 3,9)^2}{3,9} + \frac{(15 - 13,0)^2}{13,0} + \frac{(30 - 31,2)^2}{31,2} + \frac{(26 - 35,1)^2}{35,1} + \frac{(23 - 24,7)^2}{24,7} + \frac{(18 - 14,3)^2}{14,3} + \frac{(12 - 7,8)^2}{7,8} = 7,179759... \approx 7,1798$$



Eri arvosanoja on seitsemän. Saatua testisuureen arvoa verrataan χ^2 -jakaumaan, jossa vapausasteiden määrä on $k - 1 = 7 - 1 = 6$.

Määritetään taulukkokirjan tai laskentaohjelman avulla kriittinen arvo χ^2 -jakaumalle, kun vapausasteiden määrä on 6 ja luottamustaso on 0,95. Kriittiseksi arvoksi saadaan 12,5916.

Laskettu testisuureen arvo 7,1798 ei ylitä kriittistä arvoa 12,5916. Tulos tarkoittaa sitä, että annetut arvosanat eivät poikkea tilastollisesti merkitsevästi edellisen vuoden arvosanajakaumasta.

Tapa 2

Käytetään laskentaohjelman χ^2 -testitoimintoa. Valitaan laskentaohjelmassa rivien määräksi 7 ja syötetään arvosanojen frekvenssit ja teoreettiset frekvenssit laskentaohjelman taulukkopohjaan.

Laskentaohjelma antaa testisuureen arvon $\chi^2 = 7,1798$ ja testin riskitasoksi $p \approx 0,3045 = 30,45\%$.

Riskitason mukaan on 30,45 prosentin todennäköisyys, että annettujen arvosanojen jakauman erot verrattuna aiempaan arvosanajakaumaan johtuvat sattumasta. Koska riskitaso on suurempi kuin 5 %, annetut arvosanat eivät poikkea tilastollisesti merkitsevästi edellisen vuoden arvosanajakaumasta.

Vastaus

Arvosanajakauma ei poikkea merkitsevästi aiemmasta.

χ^2 -riippumattomuustesti

Ristiintaulukoinnin avulla voidaan selvittää, poikkeako kaksi ryhmää tai useampi ryhmä toisistaan tutkittavan ominaisuuden suhteen. Tällöin ristiintaulukoinnin antamia havaittuja frekvenssejä verrataan teoreettisiin frekvensseihin, joiden mukaan eri ryhmien jakaumat vastaavat toisiaan. Havaituista ja teoreettisista frekvensseistä lasketaan testisuure χ^2 samaan tapaan kuin χ^2 -yhteensopivuustestissä.

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i},$$

missä

o_i on i :nnen luokan havaittu frekvenssi

e_i on i :nnen luokan teoreettinen frekvenssi

k on luokkien lukumäärä, joka on sarakkeiden ja rivien lukumäärien tulo.

Kahden muuttujan muodostamasta ristiintaulukoinnista saatua testisuureen χ^2 arvoa verrataan χ^2 -jakaumaan, jossa vapausasteiden määrä on $(s - 1)(r - 1)$, missä s on sarakkeiden ja r rivien lukumäärä ristiintaulukoinnissa.

χ^2 -riippumattomuustestiä voidaan käyttää, jos kaikki teoreettiset frekvenssit ovat suurempia kuin 1 ja enintään 20 % teoreettisista frekvensseistä on pienempiä kuin 5.

Esimerkki 5

Korkeakouluopiskelijoille tehdyssä kyselyssä kysyttiin, onko opiskelijalla omasta mielestään työelämässä riittävä ruotsin kielen taito. Ristiintaulukointi kuvaa vastausten jakauman erikseen miesten ja naisten osalta.

	Mies	Nainen	Yhteensä
Kyllä	35	69	104
Ei	46	50	96
Yhteensä	81	119	200

Avaa OpenOffice-tiedosto →

Selvitä 95 prosentin luottamustasolla, onko miesten ja naisten ruotsin kielen osaamisessa tilastollisesti merkitsevää eroa.

Ratkaisu

Tapa 1

Taulukon perusteella vastaajia oli yhteensä 200, joista miehiä oli 81 ja naisia 119. Kaikista vastaajista ruotsin kieltä osaavien osuus oli

$$\frac{104}{200} = 0,52 = 52 \, \%.$$

Mikäli miesten ja naisten välillä ei olisi eroa ruotsin kielen osaamisessa, olisi kyselyn miesvastaajista ruotsin kieltä osannut

$$0,52 \cdot 81 = 42,12$$

ja naisvastaajista

$$0,52 \cdot 119 = 61,88.$$

Tällöin ristiintaulukointi sisältäisi seuraavat luvut.

	Mies	Nainen	Yhteensä
Kyllä	42,12	61,88	104
Ei	38,88	57,12	96
Yhteensä	81	119	200

Havaittujen ja teoreettisten frekvenssien sarake- ja rivisummat ovat samat.

Verrataan kyselyn antamia havaintomääriä näihin teoreettisiin frekvensseihin, joiden mukaan eroa ruotsin kielen osaamisessa ei olisi. Lasketaan testisuureen χ^2 arvo havaittujen ja teoreettisten frekvenssien avulla.

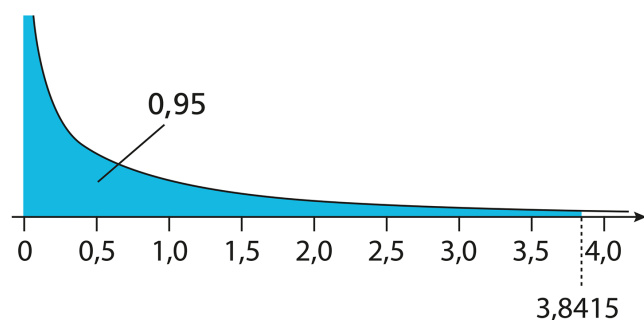
$$\chi^2 = \sum_{i=1}^4 \frac{(o_i - e_i)^2}{e_i} = \frac{(35 - 42,12)^2}{42,12} + \frac{(69 - 61,88)^2}{61,88} + \frac{(46 - 38,88)^2}{38,88} + \frac{(50 - 57,12)^2}{57,12} = 4,21418... \approx 4,214$$

Kahden sarakkeen ja kahden rivin taulukosta saatua χ^2 arvoa verrataan χ^2 -jakaumaan, jonka vapausasteiden määrä on

$$(s - 1)(r - 1) = (2 - 1)(2 - 1) = 1.$$

Määritetään taulukkokirjan tai laskentaohjelman avulla kriittinen arvo χ^2 -jakaumalle, kun vapausasteiden määrä

on 1 ja luottamustaso on 0,95. Kriittiseksi arvoksi saadaan 3,8415.



Laskettu testisuureen arvo 4,214 ylittää kriittisen arvon. Tulos tarkoittaa sitä, että havaitut frekvenssit poikkeavat tilastollisesti merkitsevästi teoreettisista frekvensseistä. Toisin sanoen eroa miesten ja naisten ruotsin kielen osaamisessa voidaan pitää luotettavana. Päätelmään liittyvä riskitaso eli todennäköisyys, että ero johtuisi sattumasta on alle 5 %.

Tapa 2

Käytetään laskentaohjelman χ^2 -testitoimintoa. Valitaan laskentaohjelmassa rivien ja sarakkeiden määräksi 2 ja syötetään havaitut frekvenssit laskentaohjelman taulukkopohjaan.

Laskentaohjelmissa on yleensä χ^2 -testille kaksi eri toimintoa, joista toista käytetään χ^2 -yhteensopivuustestissä ja toista χ^2 -riippumattomuustestissä. Nimitykset vaihtelevat ohjelman mukaan.

Laskentaohjelma antaa testisuureen arvon $\chi^2 = 4,2142$ ja testin riskitasoksi $p \approx 0,0401 = 4,01$ %.

Riskitason mukaan on noin 4 prosentin todennäköisyys sille, että havaittujen ja teoreettisten frekvenssien erot johtuvat sattumasta. Koska riskitaso on pienempi kuin 5 %, havaitut frekvenssit poikkeavat teoreettisista frekvensseistä tilastollisesti merkitsevästi. Otoksessa

havaittua eroa miesten ja naisten ruotsin kielen osaamisessa voidaan pitää luotettavana.

Vastaus

Ruotsin kielen osaamisessa miesten ja naisten välillä on tilastollisesti merkitsevä ero.

Esimerkki 6

Ensimmäisen ja toisen vuoden opiskelijat ottivat kantaa väitteeseen ”Kouluruoka on maukasta”. Ristiintaulukointi kuvaa vastausten jakautumisen erikseen ensimmäisen ja toisen vuoden opiskelijoiden osalta.

	1. vuosi	2. vuosi	Yhteensä
Samaa mieltä	18	27	45
Jokseenkin samaa mieltä	55	38	93
Jokseenkin eri mieltä	21	12	33

Eri mieltä	8	5	13
Yhteensä	102	82	184

Avaa OpenOffice-tiedosto →

Selvitä 95 prosentin luottamustasolla, poikkeavatko ensimmäisen ja toisen vuoden opiskelijoiden mielipiteet toisistaan.

Ratkaisu

Tapa 1

Lasketaan teoreettiset frekvenssit, joiden mukaan 1. ja 2. vuoden opiskelijat jakautuvat neljään luokkaan samassa suhteessa kokonaisjakauman mukaisesti.

	1. vuosi	2. vuosi	Yhteensä
Samaa mieltä	$\frac{45}{184} \cdot 102 =$ 24,9457	$\frac{45}{184} \cdot 82 =$ 20,0543	45
Jokseenkin samaa mieltä	$\frac{93}{184} \cdot 102 =$ 51,5543	$\frac{93}{184} \cdot 82 =$ 41,4457	93
Jokseenkin eri mieltä	$\frac{33}{184} \cdot 102 =$ 18,2935	$\frac{33}{184} \cdot 82 =$ 14,7065	33
Eri mieltä	$\frac{13}{184} \cdot 102 =$ 7,2065	$\frac{13}{184} \cdot 82 =$ 5,7935	13
Yhteensä	102	82	184

Lasketaan testisuureen χ^2 arvo havaittujen ja teoreettisten frekvenssien avulla.

$$\chi^2 = \sum_{i=1}^8 \frac{(o_i - e_i)^2}{e_i} = \frac{(18 - 24,9457)^2}{24,9457} + \frac{(55 - 51,5543)^2}{51,5543} +$$

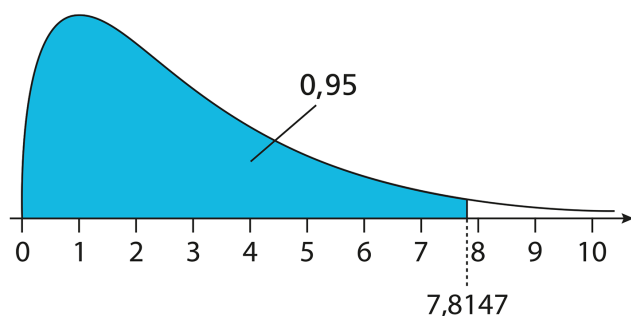
$$\frac{(21 - 18,2935)^2}{18,2935} + \frac{(8 - 7,2065)^2}{7,2065} + \frac{(27 - 20,0543)^2}{20,0543} +$$

$$\frac{(38 - 41,4457)^2}{41,4457} + \frac{(12 - 14,7065)^2}{14,7065} + \frac{(5 - 5,7935)^2}{5,7935} =$$

$$5,95084... \approx 5,951$$

Kahden sarakkeen ja neljän rivin taulukosta saatua testisuureen χ^2 arvoa verrataan χ^2 -jakaumaan, jonka vapausasteiden määrä on $(s - 1)(r - 1) = (2 - 1)(4 - 1) = 3$.

Määritetään taulukkokirjan tai laskentaohjelman avulla kriittinen arvo χ^2 -jakaumalle, kun vapausasteiden määrä on 3 ja luottamustaso on 0,95. Kriittiseksi arvoksi saadaan 7,8147.



Laskettu testisuureen arvo 5,951 sijoittuu χ^2 -jakauman luottamusvälille eli on pienempi kuin kriittinen arvo. Tulos tarkoittaa sitä, että havaitut frekvenssit eivät poikkea tilastollisesti merkitsevästi teoreettisista frekvensseistä. Toisin sanoen 1. ja 2. vuoden opiskelijoiden mielipiteet eivät eroa toisistaan tilastollisesti merkitsevästi.

Tapa 2

Valitaan laskentaohjelmassa rivien määräksi 4 ja sarakkeiden määräksi 2 ja syötetään havaitut frekvenssit laskentaohjelman taulukkopohjaan.

Laskentaohjelma antaa testisuureen arvon $\chi^2 = 5,9508$ ja testin riskitasoksi $p \approx 0,114 = 11,4 \%$. Riskitason mukaan on 11,4 prosentin todennäköisyys, että havaittujen ja teoreettisten frekvenssien erot johtuvat sattumasta. Koska riskitaso on suurempi kuin 5 %, havaittujen ja teoreettisten frekvenssien ero ei ole tilastollisesti merkitsevä, joten 1. ja 2. vuoden opiskelijoiden mielipiteissä ei katsota olevan eroa.

Vastaus

1. ja 2. vuoden opiskelijoiden mielipiteiden välillä ei ole eroa.

Avainkäsitteet

- χ^2 -testisuure
- χ^2 -jakauma
- vapausasteiden lukumäärä
- χ^2 -yhteensopivuustesti
- χ^2 -riippumattomuustesti