

3.2 Vapausasteet, merkitsevyys ja khiin neliön oletukset

Khiin neliö -testin tilastollinen merkitsevyys saadaan vertaamalla havaintoaineistosta saatua χ^2 :n kertymäarvoa teoreettiseen χ^2 -jakaumaan. Teoreettinen χ^2 -jakauma on jatkuvakestoinen ja sen arvo riippuu taulukon vapausasteista. Mutta mitä vapausasteet tarkoittaa? Asiaa voisi lähestyä seuraavasti: jos tunnemme taulukon rivisummat ja sarakesummat, miten monen solun arvo ristiintaulukoidussa taulukossa voi vaihdella vapaasti. Jos taulukossa on kaksi riviä ja kaksi saraketta, kuten tupakointia koskevassa aineistossamme, tällöin ainoastaan yhden solun arvo voi vaihdella vapaasti, muiden solujen arvo määräytyy täysin tuon yhden solun sekä rivisummien ja sarakesummien perusteella. Jos taulukon vapausasteet ovat 2, se tarkoittaa sitä, että kahden solun arvo voi vaihdella vapaasti, mutta muut arvot voi laskea niiden kahden vapaasti vaihtelevan solun sekä rivi- ja sarakesummien perusteella. Vapausasteet khiin neliön tapauksessa viittaavat siis siihen, miten monen solun arvo voi vaihdella vapaasti. Vapausasteiden määrä khiin neliölle saadaan laskettua kaavalla:

$$(\text{rivienmäärä}-1) * (\text{sarakkeidenmäärä}-1)$$

Rivien määrä ja sarakkeiden määrä tässä tarkoittavat ristiintaulukoidun taulukon rivien ja sarakkeiden määrää. Kuvitteelliselle aineistollemme rivien määrä on 2 ja sarakkeiden määrä on 2, ja tämän vuoksi sen vapausasteet = 1.

Tutkimuksessa on osoitettu, että kun havaintoaineiston otos on tarpeeksi suuri, sen jakauma approksimoi teoreettista khiin neliö -jakaumaa. Näin ollen, vaikka khiin neliö -testi on epäparametrinen testi, sillä on siis oletuksia. Yksi näistä oletuksista on se, että otoskoon täytyy olla tarpeeksi suuri, muuten p-arvon approksimointi teoreettisesta khiin neliö -jakaumasta ei ole luotettava. Riittävä otoskoko määritellään seuraavasti (ks. Cochran 1952):

- 2x2 taulukossa jokaisen solun odotusarvo (huom! ei havaintoarvo) on oltava vähintään viisi
- suuremmissa taulukoissa 80 prosentissa soluista odotusarvon on oltava vähintään viisi ja kaikissa soluissa odotusarvon on oltava suurempi kuin nolla.

Huomaa vielä, että **khiin neliö -testin rajoitus koskee odotusarvoja, ei havaintoarvoja**. Tässä kannattaa olla erittäin tarkkana, koska usein tulee huolestuttua, jos havaintoarvot ovat pieniä, vaikka pitäisi pikemminkin tarkastella havaintoarvojen suuruutta. Khiin neliö -testiä voi siis käyttää, vaikka havaintoarvoissa jokin frekvenssi olisi pienempi kuin viisi, kunhan vain odotusarvot ovat tarpeeksi suuret.

Toinen oletus/rajoitus khiin neliö -testille on sen suuntaisuus. Khiin neliö -testin avulla ei voi suoraan päätellä, onko muuttujien välillä suora suhde vai esiintyykö niiden välillä käänteinen suhde. Tämän arvioimiseksi on käytettävä muita keinoja, kuten Pearsonin residuaaleja tai vetokertoimien suhdetta. Näihin voit tutustua lisälukemistosta (mm. Gries 2013: 186–187).

Khiin neliö -testissä oman havaintotaulukon p-arvo lasketaan siis teoreettisen jakauman perusteella, kun on ensin varmistuttu odotusarvojen riittävästä suuruudesta. Testissä otoksen pohjalta laskettua khiin neliön kertymäärvä verrataan vapausasteiltaan vastaavaan teoreettiseen jakaumaan ja arvioidaan, miten poikkeava kertymäärvä on suhteessa jakauman muihin arvoihin. Jos taulukon vapausasteet = 1, tällöin khiin neliön kertymäärvä 3,8415 on ns. kriittinen arvo, eli sitä vastaava p-arvo on 0,05. Eli jos kyseessä on 2x2-taulukko ja khiin neliön kertymäärvä on suurempi kuin 3,8415, tällöin muuttujien välillä esiintyy tilastollisesti merkitsevä riippuvuus ja p-arvo on pienempi kuin 0,05.

Tupakointiaineistolle χ^2 :n kertymäärvä on pyöristettynä 5,77. Tämä on selvästi suurempi kuin kriittinen arvo 3,84. Näin ollen voidaan todeta, että kuvitteellisessa aineistossa tupakoinnin ja sukupuolen välillä esiintyy tilastollisesti merkitsevä riippuvuus. Saamaamme kertymäärvä vastaava tarkka p-arvo on 0,016, joka saadaan englanninkielisessä EXCEL:issä funktion CHISQ.DIST.RT avulla (suomenkielisessä versiossa funktio on CHINELIÖ.JAKAUMA.OH). Funktiolle annetaan kaksi arvoa, jotka erotetaan toisistaan puolipisteellä: khiin neliön kertymäärvä; vapausasteet).

Tilastollisia tuloksia raportoidessa on tärkeä ilmoittaa testisuure, vapausasteet sekä p-arvo. Saamamme tulos voidaan siis raportoida seuraavasti: aineiston perusteella henkilön tupakoiminen riippuu hänen sukupuolestaan tilastollisesti merkitsevällä tavalla ($\chi^2 = 5,77$; $df = 1$; $p = 0,016$).