OTOSKOON MÄÄRÄÄMINEN

Kappaleessa 3.2 on todettu, että otoskoko vaikuttaa luottamusvälin pituuteen: *Mitä suurempi on otoskoko, sitä kapeampi on luottamusväli.* Tämän luottamusvälin ominaisuuden takia on mahdollista määrätä otoksesta saatavan luottamusvälin pituus ennen otoksen poimimista. Tämä tapahtuu valitsemalla otoskoko, joka tuottaa halutun pituisen luottamusvälin. Edellytyksenä sille, että tarvittava otoskoko voidaan määrätä ennen otoksen poimista on se, että perusjoukon hajonnan (tai varianssin) suuruudesta on olemassa *ennakkotietoa*.

Oletetaan siis, että perusjoukon hajonnan suuruudesta on käytössä ennakkotietoa. Ennakkotieto voi olla peräisin aikaisemmista tutkimuksista. Käyttämällä hyväksi ennakkotietoa voimme käyttäytyä ikäänkuin tuntisimme perusjoukon hajonnan.

Odotusarvon μ luottamusväli luottamustasolla $1-\alpha$ on muotoa

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}},$$

jossa $\bar{X}$ on otoskeskiarvo, $z_{\alpha/2}$ on luottamustasoon $1-\alpha$ liittyvä luottamuskerroin, ja σ on perusjoukon ennakkotiedon perusteella tunnetuksi oletettu standardipoikkeama. Luottamuskerroin $z_{\alpha/2}$ toteuttaa yhtälön $\Phi(-z_{\alpha/2}) = \alpha/2$, jossa Φ on standardoidun normaalijakauman kertymäfunktio.

Oletetaan, että luottamusväliksi halutaan

$$\bar{X} \pm a,$$

jolloin luottamusvälin pituus on $2a$. Tästä ja luottamusvälin lausekkeesta saadaan yhtälö

$$z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = a,$$

josta n voidaan ratkaista. Tarvittava otoskoko on siten

$$n = \left(\frac{z_{\alpha/2}\sigma}{a}\right)^2.$$

Huomaa, että tarvittava otoskoko ei riipu perusjoukon tuntemattomasta odotusarvosta, mutta riippuu perusjoukon hajonnasta, josta oletuksen mukaan on käytössä ennakkotietoa.

Tarvittavan otoskoon kaavasta nähdään seuraavaa:

- Mitä kapeampi luottamusväli halutaan, sitä enemmän havaintoja tarvitaan.
- Mitä suurempi on perusjoukon hajonta, sitä enemmän havaintoja tarvitaan.
- Mitä suurempi on valittu luottamustaso, sitä enemmän havaintoja tarvitaan.

Lienee selvää, että kaava antaa ainoastaan *suuruusluokan* tarvittavalle otoskoolle. Kun luottamusväli määrätään otoksesta, jonka koko on määrätty ym. kaavalla, luottamusvälin pituudeksi ei yleensä saada täsmälleen lukua $2a$. On kuitenkin odotettavissa, että otoksesta määrätyn luottamusvälin pituus ei poikkea kovin paljon halutusta pituudesta, jos ennakkotieto perusjoukon hajonnan suuruudesta ei poikkea kovin paljon otoksesta määrätystä otoshajonnasta.

ESIMERKKI 1.

Erään psykologisen kokeen kohteena on koehenkilön reaktioaika tiettyyn ärsykkeeseen. Aikaisemmista tutkimuksista tiedetään, että reaktioaika on normaalijakautunut niin, että jakauman hajonta suuruusluokaltaan 0.5 s.

Kuinka suuri koehenkilöiden joukko on testattava, jotta olisi mahdollista sanoa, että otoksesta määrätty reaktioajan keskiarvo ei poikkea satunnaisesti valitun henkilön odotettavissa olevasta reaktioajasta enempää kuin 0.25 s

(a) 95%:n,

(b) 99%:n

todennäköisyydellä?

(a) Koska $\Phi(-1.96) = 0.025$,

$$n = \left(\frac{1.96 \times 0.5}{0.25}\right)^2 \approx 15.4.$$

Tarvittava havaintojen lukumäärä on siten vähintään 16.

(b) Koska $\Phi(-2.58) = 0.005$,

$$n = \left(\frac{2.58 \times 0.5}{0.25}\right)^2 \approx 26.6.$$

Tarvittava havaintojen lukumäärä on siten vähintään 27. ●

Esimerkistä nähdään seuraavaa: Mitä korkeammalla todennäköisyydellä valitun mittaisen välin halutaan peittävän perusjoukon odostusarvon, sitä suurempi otos on poimittava.

Tässä esitetty tarkastelu selventää myös *tehokkuuden* käsitettä. Kappaleessa 3.1.2 on todettu, että sekä otoskeskiarvo $\overline{X}$ että otosmediaani *Md* ovat normaalijakauman odotusarvon μ harhattomia estimaattoreita. Voidaan osoittaa, että otosmediaanin varianssi on suurempi kuin otoskeskiarvon varianssi:

$$D^2(\overline{X}) = \frac{\sigma^2}{n} < \frac{\pi}{2}\frac{\sigma^2}{n} = D^2(Md).$$

Tämä merkitsee seuraavaa: Jos odotusarvolle μ halutaan muodostaa otosmediaania käyttäen samanlevyinen luottamusväli kuin otoskeskiarvoa käyttäen, tarvitaan $\pi/2$ kertaa enemmän havaintoja. Otoskeskiarvo on tässä mielessä normaalijakauman odotusarvon estimaattorina *tehokkaampi* kuin otosmediaani: Se käyttää tehokkaammin havaintoja.

3.2.3 SUHTEELLISEN OSUUDEN LUOTTAMUSVÄLI

Olkoon A tapahtuma, joka kuvaa sitä onko perusjoukon alkiolla omininaisuus $\mathscr{P}$. Olkoon tapahtuman A todennäköisyys $\mathrm{P}(A) = p$. Poimitaan perusjoukosta yksinkertainen satunnaisotos, jonka koko on n. Olkoon satunnaismuuttuja X niiden alkioiden frekvenssi, joilla on ominaisuus $\mathscr{P}$ ja olkoon

$$P = \frac{X}{n}$$

vastaava suhteellinen frekvenssi.

Luottamusvälin konstruktio perustuu siihen, että suhteellinen frekvenssi P noudattaa suurissa otoksissa approksimatiivisesti normaalijakaumaa $\mathrm{N}(p, \sqrt{pq/n})$, jossa $q = 1 - p$. Tästä voidaan päätellä, että standardoitu muuttuja

$$Z = \frac{P-p}{\sqrt{pq/n}} \sim_a \mathrm{N}(0,1).$$

Määritellään väli

$$P - z_{\alpha/2}\sqrt{\frac{PQ}{n}}, P + z_{\alpha/2}\sqrt{\frac{PQ}{n}},$$

jossa $-z_{\alpha/2}$ on piste, joka erottaa standardoidun normaalijakauman vasemmalle hännälle todennäköisyysmassan, jonka suuruus on $\alpha/2$. Symmetrian takia $+z_{\alpha/2}$ on piste, joka erottaa standardoidun normaalijakauman oikealle hännälle todennäköisyysmassan, jonka suuruus on $\alpha/2$. Siten $z_{\alpha/2}$ on luottamustasoon $1-\alpha$ liittyvä luottamuskerroin. Jos siis $\Phi(z)$ on standardoidun normaalijakauman kertymäfunktio niin,

$$\Phi(-z_{\alpha/2}) = \alpha/2,$$

$$\Phi(+z_{\alpha/2}) = 1 - \alpha/2.$$

Huomaa, että tuntemattomat p ja $q = 1 - p$ estimaattorin P hajonnan lausekkeessa $\mathrm{D}(P) = \sqrt{pq/n}$ on korvattu estimaattoreillaan P ja $Q = 1 - P$.

Voidaan osoittaa, että ym. väli on parametrin p *approksimatiivinen* luottamusväli luottamustasolla $1 - \alpha$. Väli siis peittää tuntemattoman parametrin arvon p *likimäärin* todennäköisyydellä $1 - \alpha$: *Approksimatiivisesti* pätee

$$\mathrm{P}(P - z_{\alpha/2}\sqrt{\frac{PQ}{n}} \le p \le P + z_{\alpha/2}\sqrt{\frac{PQ}{n}}) = 1 - \alpha.$$

Luottamusväli on symmetrinen keskipisteensä P suhteen ja sekä välin keskipiste että pituus vaihtelevat satunnaisesti otoksesta toiseen. Eri otoksista lasketuista väleistä approksimatiivisesti $100\times(1-\alpha)\%$ kuitenkin peittää tuntemattoman p:n arvon.

Todettakoon lopuksi, että luottamusvälin pituudella on seuraavat ominaisuudet:

Luottamusväli *kapenee*, jos

- havaintojen lukumäärä n kasvaa,
- luottamustasoa $1 - \alpha$ pienennetään.

Luottamusväli *levenee*, jos

- luottamustasoa $1-\alpha$ kasvatetaan.

Luottamusvälin pituus riippuu myös suhteellisesta frekvenssistä P. Luottamusväli on leveimmillään silloin, kun $P = 0.5$ ja se kapenee, kun P lähestyy kumpaa tahansa ääriarvoaan 0 tai 1. Tämä nähdään siitä, että estimaattorin P varianssin $D^2(P) = p(1-p)/n$ estimaattori

$$\frac{P(1-P)}{n}$$

on P:n funktiona alaspäin aukeavan *parabelin* kaari välillä $[0,1]$:

$$f(P) = \frac{1}{n}P - \frac{1}{n}P^2,\ P \in [0,1].$$

Tämä parabelin kaari saa maksiminsa pisteessä $P = 0.5$ ja miniminsä pisteissä $P = 0$ ja $P = 1$.

Sen sijaan *suhteellinen virhe*

$$\frac{\sqrt{P(1-P)}}{nP},\ P \in [0,1]$$

on suurimmillaan silloin, kun P on lähellä arvo 0 ja pienenee, kun P kasvaa.

ESIMERKKI 1.

Ennen vaaleja on tapana julkaista kyselytutkimuksia äänestäjien mielipiteistä. Kyselyissä ei kuitenkaan käytetä yksinkertaista satunnaisotantaa, vaan menetelmiä, jotka takaavat eri äänestäjäryhmien edustavuuden otoksessa. Kannatusprosentteja on myös tapana korjata edellisiin vaaleihin liittyvillä tiedoilla. Tutkimuksia tekevien yritysten käyttämissä menetelmissä on lukuisia yksityiskohtia, joita ne pitävät liikesalaisuuksina. Tämä merkitsee sitä, että ulkopuolinen ei pysty määräämään puolueiden todellisille kannatusosuuksille luottamusvälejä. Koska luottamusväliarvioita ei useinkaan julkaista, on yleisön vaikea arvioida peräkkäisissä tutkimuksissa tapahtuvien kannatusosuuksien muutosten merkitystä.

Määräämällä kannatusosuuksille tavanomaiset 95%:n luottamusvälit sen (virheellisen) oletuksen perusteella, että otos on poimittu yksinkertaisella satunnaisotannalla, saadaan kuitenkin jonkinlainen käsitys kannatusosuusarvioihin sisältyvästä epävarmuudesta.

Seuraavassa taulukossa verrataan viimeisen ennen vuoden 1995 eduskuntavaaleja julkaistun kyselytutkimuksen tuloksia vaalien tulokseen. Kysely on ollut eräässä mielessä onnistunut, jos vaalien tulokset sopivat kyselytutkimuksen tuloksille määrättyjen luottamusvälien sisään. On kuitenkin syytä huomata, että vertailu on tarkkaan ottaen oikeutettu vain sillä edellytyksellä, että äänestäjät eivät ole muuttaneet mielipiteitään kyselyn ja vaalien välissä. Myös vaaleissa nukkuvat saattavat aiheuttaa vertailuun harhaa.

Otoskoko n oli kyselyssä noin 2000.

Taulukko 1. Puolueiden kannatusosuudet viimeisessä ennen vuoden 1995 eduskuntavaaleja tehdyssä kyselyssä ja vaaleissa.

Puolue	Kannatusosuus kyselyssä (%)	95%:n luottamus-väli	Vaalien tulos
Keskusta	16.6	15.0,18.2	19.9
Sdp	30.4	28.3,32.4	28.3
Kok	16.9	15.3,18.5	17.9
Vas	11.8	10.4,13.2	11.2
Rkp	4.9	4.0,5.8	5.1
Vihr	7.9	6.7,9.1	6.5
Skl	3.1	2.3,3.9	3.0
Smp	0.9	0.5,1.3	1.3
Lkp	0.7	0.3,1.1	0.6
Nuors	4.0	3.1,4.8	2.8

Kuten taulukosta näkyy, vaalien tulos mahtuu luottamusvälin sisään kaikissa muissa paitsi Keskustan, Vihreiden ja Nuorsuomalaisten tapauksissa. Virhe oli suurin Keskustan tapauksessa.

Tutkimus antoi siis melko hyvän ennakkokuvan äänestäjien mielipiteistä vaaleissa Keskustan kannatusta lukuunottamatta. ●

OTOSKOON MÄÄRÄÄMINEN

Tutkimuksen kohteena olevan ominaisuuden $\mathscr{P}$ omaavien perusjoukon alkioiden suhteelliselle osuudelle p halutaan usein tuottaa luottamusväli, jonka pituus on valittu etukäteen. Tämä perustuu siihen, että suhteellisen osuuden luottamusvälin pituus riippuu *otoskoosta*. Jos suhteellisen osuuden suuruusluokasta on käytettävissä *ennakkotietoa*, saadaan tarvittava otoskoko määrätyksi. Ennakkotieto voi olla peräisin aikaisemmista tutkimuksista. Koska luottamusväli on leveimmillään silloin, kun suhteellinen frekvenssi $P = 0.5$, saadaan tarvittavalle otoskoolle samaa tarkastelua käyttäen yläraja myös sellaisissa tilanteissa, joissa tällaista ennakkotietoa ei ole käytettävissä.

Suhteellisen osuuden p luottamusväli on muotoa

$$P \pm z_{\alpha/2}\sqrt{\frac{P(1-P)}{n}},$$

jossa $z_{\alpha/2}$ on luottamustasoon $1-\alpha$ liittyvä luottamuskerroin. Luottamuskerroin $z_{\alpha/2}$ toteuttaa yhtälön $\Phi(-z_{\alpha/2})=\alpha/2$, jossa Φ on standardoidun normaalijakauman kertymäfunktio.

Oletetaan, että luottamusväliksi halutaan

$$P \pm a,$$

jolloin luottamusvälin pituus on $2a$.

Käyttämällä ennakkotietoa suhteellisesta osuudesta p, voidaan luottamusvälin lausekkeessa oleva neliöjuurilauseke korvata suhteellisen osuuden P standardipoikkeamalla

$$\sqrt{\frac{p(1-p)}{n}},$$

jossa p on ennakkotiedon mukainen arvo ominaisuuden $\mathscr{P}$ omaavien perusjoukon alkioiden suhteelliselle osuudelle.

Tästä ja luottamusvälin lausekkeesta saadaan yhtälö

$$z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}=a,$$

josta tarvittava otoskoko n voidaan ratkaista. Tarvittava otokoko on siten

$$n=\left(\frac{z_{\alpha/2}\sqrt{p(1-p)}}{a}\right)^2.$$

Tarvittavan otoskoon yläraja n_{max} saadaan sijoittamalla tähän $p = 0.5$:

$$n_{max}=\left(\frac{z_{\alpha/2}}{2a}\right)^2.$$

ESIMERKKI 2.

Erääseen kuntaan on suunnitteilla ydinvoimala. Kunnan johto haluaa selvittää asukkaiden mielipiteet ilman kansanäänestystä. Siksi asukkaiden mielipiteitä päätetään tutkia poimimalla asukkaiden joukosta yksinkertainen satunnaisotos, johon poimituilta asukkailta kysytään kannattavatko he ydinvoimalan rakentamista kuntaan vai ei. Rakentamisen kannattajia ja vastustajia arvioidaan olevan kunnassa suunnilleen yhtä paljon.

Koska asia on tärkeä, halutaan 99%:n varmuus sille, että kannattajien todellinen suhteellinen osuus ei eroa enempää kuin 1:n prosenttiyksikön verran otoksessa havaitusta kannattajien osuudesta.

Koska $\Phi(-2.58) = 0.005$,

$$n_{max}=\left(\frac{2.58}{2\times 0.01}\right)^2=16641.$$

Näin suuria otoksia ei yleensä ole varaa eikä mahdollista kerätä. Sitä paitsi ko. kunnassa voi olla jopa vähemmän asukkaita kuin otokseen vaaditaan! ●

Tästä esimerkistä nähdään, että tilanteissa, joissa tutkitaan ihmisten mielipiteitä kysymykseen, joka jakaa ihmiset kahteen suunnilleen yhtä suureen ryhmään, tarvitaan erittäin paljon havaintoja ennenkuin voidaan olla varmoja, että todelliset ja otoksesta määrätyt suhteelliset osuudet eivät poikkea kovin paljoa toisistaan. Tämä merkitsee sitä, että tilanteissa, joissa vaa'an kallistumisesta puoleen tai toiseen on merkittäviä seurauksia, todellisista suhteellisista osuuksista ei voida saada kovin helposti varmuutta käyttämällä otantaa.

On syytä huomata, että otoskoko voi olla paljonkin pienempi, jos ryhmien koot perusjoukossa poikkeavat selvästi maksimilevyisen luottamusvälin tuottavasta arvosta 0.5. Tätä vahvistaa myös se, että luottamusvälin ei tällöin tarvitsekaan olla lyhyt, koska ryhmien suhteelliset osuudet perusjoukosta poikkeavat toisistaan selvästi!

4. TESTAUS

4.1 MERKITSEVYYSTESTIT

4.1.1 JOHDANTO

Empiirisen tutkimuksen päämääränä on tehdä havaintoihin perustuvia *johtopäätöksiä* tutkimuksen kohteena olevasta ilmiöstä. Empiirisessä tutkimuksessa on käytössä kaksi erilaista lähestymistapaa, joita molempia luonnehditaan lyhyesti seuraavassa.

Eksploratiivista (engl. *explore*, tehdä tutkimusmatka) lähestymistapaa käytetään silloin, kun tutkimuksen kohde on niin uusi tai huonosti tunnettu, että ensisijaisena päämääränä on selvittää kohteen ominaisuudet. Tätä lähestymistapaa käytetään ennen tutkimuksen kohdetta koskevan teorian rakentamista ns. *raivaustutkimuksissa*. Raivaustutkimuksissa tärkeimmät työvaiheet ovat tutkimusaineiston kerääminen ja kerättyyn aineistoon kaivautuminen, jolloin päämääränä on selvittää mitä aineistosta löytyy. Tilastollisessa tutkimuksessa eksploratiivista lähestymistapaa edustavat selvimmin *aineiston kuvailussa käytettävät menetelmät*.

Konfirmatorisessa (engl. *confirm*, vahvistaa) lähestymistavassa tutkimuksen kohde tunnetaan niin hyvin, että siitä on olemassa teoria, joka tuottaa kohdetta koskevia oletuksia ja ennusteita. Tutkimus pyrkii tällöin selvittämään ovatko tutkimuksen kohteesta tehdyt oletukset sopusoinnussa havaintojen kanssa tai tuottaako teoria oikeaan osuvia ennusteita. Tilastollisessa tutkimuksessa konfirmatorista lähestymistapaa edustavat selvimmin *tilastolliset testit*.

Kaikkien tieteiden historiassa voidaan nähdä eksploratiivisen ja konfirmatorisen lähestymistavan vuorottelua. Eksploratiivista lähetymistapaa käytetään tieteen raivaustutkimusvaiheessa. Tämä vaihe johtaa teorian formulointiin. Konfirmatorisessa vaiheessa teoria vahvistetaan empiirisesti. Tiedon karttuessa löydetään kuitenkin yleensä empiirisiä tosiseikkoja, jotka eivät sovi teoriaan. Tämä pakottaa teorian uudistamiseen ja uuteen raivaustutkimusvaiheeseen.

Edellisessä luvussa tarkasteltiin perusjoukkoa kuvaavien parametrien *estimointia*. Tällöin todettiin, että parametrien estimaattoreina käytetään perusjoukosta kerättyä havaintoaineistoa kuvaavia otostunnuslukuja. Estimaattoreiden ominaisuudet kiteytettiin niiden otosjakaumiin. Tällöin kiinnitettiin erityistä huomiota siihen satunnais- eli otosvaihteluun, joka tekee perusjoukon parametreista tehtävät johtopäätökset epävarmoiksi. Epävarmuuden astetta voidaan kuvata parametrin luottamusvälin avulla. Luottamusväli konstruoinnissa käytetään apuna parametria estimoivan otostunnusluvun otosjakaumaa.

Tämän luvun aiheena ovat *tilastolliset testit*. Kuten edellä todettiin, empiirisen tieteellisen tutkimuksen konfirmatorisessa vaiheessa tutkimuksen kohdetta koskevia oletuksia pyritään testaamaan selvittämällä ovatko oletukset sopusoinnussa tutkimuksen kohteesta kerättyjen havaintojen kanssa. Jotta tilastollisia testejä voisi käyttää, oletukset on muotoiltava jonkin perusjoukon ominaisuuksia kuvaavan tilastollisen mallin parametreja koskeviksi oletuksiksi. Tilastollinen testaus kohdistuu näihin parametreihin.

Tilastollisessa testauksessa olennaisin työvaihe on löytää sopiva *testisuure* tutkimuksen kohteesta tehdylle oletukselle. Testisuureen tehtävänä on mitata tehdyn oletuksen ja havaintojen yhteensopivuutta. Testisuureella on aina jokin *normaaliarvo*, joka kuvastaa sitä, että testattava oletus ja havainnot sopivat (täydellisesti) yhteen. Jos testisuureen arvo poikkeaa normaaliarvostaan kyllin paljon, joudutaan testattava oletus asettamaan epäilyksen alaiseksi. Tällöin pyritään selvittämään onko testisuureen poikkeama normaaliarvostaan niin suuri, että poikkeamaa ei voida pitää sattuman tuottamana. Tällöin poikkeamaa voidaan pitää *tilastollisesti merkitsevänä*.

Edellisen luvun johdannossa viitattiin jo siihen, että *tilastollisessa päättelyssä johtopäätökset eivät ole tavanomaisessa mielessä varmoja*. Tämä johtuu tilastollisiin tutkimusaineistoihin sisältyvästä satunnaisvaihtelusta. Tämä on totta myös tilastollisista testeistä: Tilastollisen testi voi johtaa virheelliseen johtopäätökseen. Onneksi johtopäätösten teossa tehtävien *virheiden todennäköisyyksille* voidaan kuitenkin antaa arvio.

4.1.2 MERKITSEVYYSTESTIT

Otetaan tarkastelun kohteeksi seuraava esimerkki:

ESIMERKKI 1.

Psykologisessa kokeessa haluttiin selvittää melun vaikutusta älylliseen suorituskykyyn. Koehenkilöt pantiin suorittamaan helppoja aritmeettisia laskutoimituksia annetussa ajassa joko meluisassa tai meluttomassa työskentely-ympäristössä ja mittauksen kohteeksi otettiin tehtyjen virheiden lukumäärä.

Koetulokset on ilmoitettu seuraavassa taulukossa:

Taulukko 1. Virheiden lukumäärien keskiarvot ja hajonnat meluttomassa ja meluisassa työskentely-ympäristössä.

Ryhmä	Ryhmän koko	Virheiden lkm:n keskiarvo	Virheiden lkm:n keskihajonta
Meluton ympäristö	16	14.4	2.28
Meluisa ympäristö	16	16.5	2.45

Keskiarvotietojen perusteella näyttää siltä, että koehenkilöt suoriutuivat meluttomassa ympäristössä laskutoimituksista paremmin. Miten tämän havainnon merkitystä tutkitaan? Onko havaittu ero niin suuri, että sen perusteella voidaan todella sanoa, että meluttomassa ympäristössä työskentelevät selviytyvät paremmin älyllisistä tehtävistä vai onko havaittu ero sellainen, että se voisi olla seurauksena *satunnaisvaihtelusta*? Tilastollinen *testiteoria* pyrkii vastaamaan näihin kysymyksiin. ●

Testausasetelmassa joudutaan aina ensin kiinnittämään jokin *yleinen hypoteesi*, joka sisältää oletukset tutkittavan ominaisuuden jakautumisesta perusjoukossa ja käytetystä otantamenettelystä. Yleinen hypoteesi jätetään usein esittelemättä testausasetelmaa kuvattaessa. Se on kuitenkin olennainen osa asetelmaa.

YLEINEN HYPOTEESI

Testausasetelmasta tehtävät yleiset oletukset muodostavat *yleisen hypoteesin* H.

Esimerkissämme yleiseksi hypoteesiksi otetaan seuraavat oletukset: Satunnaisesti valitun henkilön tekemien virheiden lukumäärä noudattaa sekä meluttomassa että meluisassa ympäristössä työskenneltäessä normaalijakaumaa. Lisäksi eri henkilöiden tekemät virheet jakautuvat toisistaan riippumattomasti.

Esimerkissämme asetamme satunnaismuuttujan X_{1i} kuvaamaan henkilön i tekemien virheiden lukumäärää meluttomassa ympäristössä ja satunnaismuuttujan X_{2j} kuvaamaan henkilön j tekemien virheiden lukumäärää meluisassa ympäristössä. Yleinen hypoteesi H sisältää seuraavat kolme oletusta:

1. $X_{1i} \sim N(\mu_1, \sigma_1^2)$,
2. $X_{2j} \sim N(\mu_2, \sigma_2^2)$,
3. X_{1i} ja X_{2j} ovat riippumattomia kaikille i ja j.

Huomaa, että riippumattomuushypoteesiin sisältyy kaksi erillistä osaa:

- Henkilön A tekemien virheiden lukumäärä ei riipu henkilön B tekemien virheiden lukumäärästä kummankaan ryhmän *sisällä*.
- Henkilön A tekemien virheiden lukumäärä ei riipu henkilön C tekemien virheiden lukumäärästä ryhmien *välillä*.

Jälkimmäinen ehto merkitsee mm. sitä, että sellainen tutkimusatelma suljetaan pois, jossa samoja koehenkilöitä käytetään molemmissa testeissä. Jos samoja koehenkilöitä käytetään molemmissa testeissä, joudutaan ns. *parivertailutilanteeseen*, jota ei voi testata alla esitettävällä tavalla. Parivertailutilanteelle esitetään testi myöhemmin.

Oletusta, jota halutaan testata kutsutaan nollahypoteesiksi:

NOLLAHYPOTEESI

Testattava oletus muodostaa *nollahypoteesin* H_0.

Yleistä hypoteesia ei aseteta *nollahypoteesia testattaessa* kyseenalaiseksi. Toinen asia on se, että myös yleisen hypoteesin oletuksia voidaan ja myös pitää testata. Esimerkkimme yleiseen hypoteesiin sisältyviä oletuksia tehtyjen virheiden jakaumien normaalisuudesta voidaan testata myöhemmin esitettävällä χ^2-yhteensopivuustestillä.

Esimerkissämme nollahypoteesiksi otetaan se, että työskentely-ympäristön meluisuus tai meluttomuus *ei* vaikuta kokeessa tehtyjen virheiden määrään. Siten

$$H_0: \quad \mu_1 = \mu_2.$$

Nollahypoteesi on lähes aina muotoa

ei eroa

tai

ei vaikutusta.

Tilastollinen testaus perustuu siihen, että nollahypoteesi asetetaan havainnosta saatuja tietoja vastaan. Jos havainnoista saatavat todisteet nollahypoteesia vastaan ovat kyllin voimakkaita, nollahypoteesi hylätään.

MERKITSEVYYSTESTI

Merkitsevyystesti on *päätössääntö*, joka kertoo onko nollahypoteesi hylättävä vai ei.

Testin tekeminen perustuu *testisuureeseen*, joka *mittaa* havaitun maailmantilan todennäköisyyttä, jos nollahypoteesi pätee. Testisuureella on aina jokin normaaliarvo, josta poikkeaminen viittaa siihen, että nollahypoteesi ei päde. Testi suunnitellaan siten, että se vahvistaa havaintoaineistoon sisältyviä todisteita nollahypoteesia vastaan.

Esimerkissämme testisuure on luonteva perustaa otoskeskiarvojen erotukseen: Jos erotus on "suuri", nollahypoteesi hylätään. Jos taas erotus on "pieni", nollahypoteesi jätetään voimaan. Varsinaisena ongelmana testauksessa on se, millä perusteella tehdään johtopäätös, että erotus on "suuri" tai "pieni". Testisuure konstruoidaan juuri tähän tehtävään.

Esimerkissämme testisuure perustetaan siis erotukseen

$$\overline{X}_1 - \overline{X}_2$$

jossa

$$\overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i}$$

on meluttomassa ympäristössä työskennelleiden koehenkilöiden virheiden keskiarvo ja

$$\overline{X}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} X_{2j}$$

on meluisassa ympäristössä työskennelleiden koehenkilöiden virheiden keskiarvo.

Jos erotus saa arvon nolla, ei ryhmien suorituskyvyillä ole eroa. Jos erotus poikkeaa tästä normaaliarvosta kyllin paljon, on nollahypoteesi asetettava kyseenalaiseksi. Erotuksen suuruuden mittaaminen on järkevää perustaa erotuksen hajontaan. Aikaisemmin esitettyjen tulosten nojalla tiedetään, että

$$\overline{X}_1 - \overline{X}_2 \sim \mathrm{N}(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}).$$

Erityisesti otoskeskiarvojen erotuksen odotusarvo ja varianssi saadaan kaavoista

$$\mathrm{E}(\overline{X}_1 - \overline{X}_2) = \mu_1 - \mu_2,$$

$$\mathrm{D}^2(\overline{X}_1 - \overline{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Huomaa, että varianssia koskeva tulos seuraa siitä, että virheiden lukumäärät oletettiin riippumattomiksi ryhmien *välillä*. Näiden tulosten perusteella *standardoitu muuttuja*

$$\frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

noudattaa standardoitua normaalijakaumaa $\mathrm{N}(0,1)$. Tämä lauseke ei kuitenkaan kelpaa sellaisenaan testisuureeksi, koska se riippuu tuntemattomista ryhmäkohtaisista odotusarvoista μ_1 ja μ_2 varianssеista σ_1^2 ja σ_2^2.

Huomaa kuitenkin, että erotus

$$\mu_1 - \mu_2 = 0,$$

jos nollahypoteesi H_0 *pätee*. Jos lisäksi varianssit σ_1^2 ja σ_2^2 korvataan vastaavilla otosvariansseilla, saadaan testisuure

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

Voidaan osoittaa, että testisuure Z noudattaa suurissa otoksissa *approksimatiivisesti* standardoitua normaalijakaumaa $\mathrm{N}(0,1)$, *jos nollahypoteesi pätee*.

Esimerkissämme otoskeskiarvojen erotus saa arvon

$$\overline{X}_1 - \overline{X}_2 = 14.4 - 16.5 = -2.1$$

ja erotuksen hajonnan estimaatti on

$$\hat{\mathrm{D}}(\overline{X}_1 - \overline{X}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{2.28^2}{16} + \frac{2.45^2}{16}} = 0.837.$$

Testisuureen Z arvoksi saadaan siten

$$Z = \frac{-2.1}{0.837} = -2.51.$$

Poikkeaako tämä testisuureen arvo niin paljon testisuureen normaaliarvosta 0, että nollahypoteesi joudutaan asettamaan kyseenalaiseksi?

Ennen kuin vastaamme tähän kysymykseen, tarkastelemme mitä seurauksia on nollahypoteesin kyseenalaistamisesta.

VAIHTOEHTOINEN HYPOTEESI

Vaihtoehtoinen hypoteesi H_1 on oletus, joka pätee, jos nollahypoteesi *ei ole tosi*.

Esimerkissämme eräs mahdollinen vaihtoehtoinen hypoteesi on se, että työskentely-ympäristön meluisuudella tai meluttomuudella on vaikutusta koehenkilöiden tekemien virheiden määrään. Siten

$$H_1: \mu_1 \neq \mu_2.$$

On syytä huomata, että vaihtoehtoinen hypoteesi voidaan tavallisesti muotoilla usealla eri tavalla; kts. yksi- ja kaksisuuntaisia testejä käsittelevää kohtaa myöhemmin tässä kappaleessa. Huomaa, että sekä nollahypoteesi että vaihtoehtoinen hypoteesi ovat yleisen hypoteesin osia.

On syytä huomata se, että *useissa testaustilanteissa itse asiassa toivotaan, että vaihtoehtoinen hypoteesi tulee hyväksytyksi*. Näin ei tietystikään ole aina asian laita. Verrataanpa vaikka tilannetta, jossa testataan onko uudella lääkkeellä parantavaa vaikutusta, tilanteeseen, jossa testataan onko uudella lääkkeellä sivuvaikutusta. Kummassakin tapauksessa nollahypoteesi on muotoa "*ei vaikutusta*". Edellisessä tapauksessa toivotaan, että nollahypoteesi voidaan hylätä, jälkimmäisessä tapauksessa, että se jää voimaan. Huomaa, että tutkittaessa pitääkö uusi lääke ottaa käyttöön vai ei, joudutaan samanaikaisesti *molempien* testausongelmien eteen. Tämä on totta yleisemminkin: Tutkimuksessa kohdataan tavallisesti samanaikaisesti useita erilaisia testausongelmia. Testeistä saatavat tulokset saattavat vaikuttaa eri suuntiin ja lopullisen johtopäätöksen tekeminen esimerkiksi siitä pitääkö uusi lääke ottaa käyttöön vai ei, saattaa olla hyvinkin vaikea tehtävä.

Nollahypoteesi on siis tavallisesti muotoa "*ei eroa*" tai "*ei vaikutusta*". Siten nollahypoteesi vastaa tasaisen harmaata tai konservatiivista maailmankuvaa. Sen sijaan vaihtoehtoiseen hypoteesiin liittyy rikkaampi maailmankuva, joka kertoo, että "*eroja on*" tai "*vaikutusta on*". Siten vaihtoehtoisen hypoteesin hyväksyminen lisää tavallisesti *informaatiota*.

Palataan nyt kysymykseen siitä, miten päätetään, että testisuureen arvo poikkeaa normaaliarvostaan niin paljon, että nollahypoteesi on syytä hylätä ja siten hyväksyä vaihtoehtoinen hypoteesi.

***P*-ARVO**

Testin *P-arvo* on todennäköisyys saada testisuureen arvo, joka poikkeaa havaitusta arvosta *saman verran tai enemmän* kuin havaittu arvo, *kun nollahypoteesi pätee*.

Koska testisuure mittaa miten paljon havaittu maailmantila poikkeaa nollahypoteesin mukaisesta maailmantilasta, voidaan päätellä seuraavaa: Mitä *pienempi* on testin *P*-arvo, sitä voimakkaammin havaintoaineisto todistaa nollahypoteesia vastaan. On syytä huomata, että *P*-arvon pitää olla kyllin pieni, jotta nollahypoteesi kannattaa hylätä. Ylärajana pidetään usein arvoja 0.01 — 0.05.

Testisuureen arvoa vastaava *P*-arvo määrätään testisuureen jakauman avulla. Oletetaan, että testisuureen normaaliarvo on 0 ja, että jakauma on normaaliarvonsa suhteen symmetrinen. Esimerkissämme on tällainen tilanne. Koska *P*-arvo kuvaa todennäköisyyttä, että testisuure saa havaitun arvon tai vielä poikkeuksellisemman arvon, *kun nollahypoteesi pätee*, *P*-arvo määrätään testisuureen jakauman kertymäfunktion avulla seuraavalla tavalla:

Testisuureen arvoa z vastaava P-arvo on testisuureen jakauman kertymäfunktion arvo pisteessä $-|z|$.

Esimerkissämme testisuureen arvoa –2.51 vastaavaksi *P*-arvoksi saadaan standardoidun normaalijakauman taulukoista 0.006. Jos siis satunnaismuuttuja Z noudattaa standardoitua normaalijakaumaa, niin

$$P(Z \leq -2.51) = 0.006.$$

Siten testisuureen arvoa voidaan pitää varsin poikkeuksellisena, jos nollahypoteesi pätee: Vain 6 kertaa 1000:sta havaitaan testisuureen arvo, joka poikkeaa normaaliarvostaan 0 näin paljon tai vielä enemmän, *jos nollahypoteesi pätee*. Nollahypoteesi voidaan siksi hylätä ja vaihtoehtoinen hypoteesi hyväksyä: Suorituskyvyt meluttomassa ja meluisassa ympäristössä eroavat toisistaan testin mukaan.

MERKITSEVYYSTASOT

Usein on tapana *kiinnittää* eli valita ennen testiä todennäköisyys, jota pienemmäksi *P*-arvon pitää tulla, jotta nollahypoteesi hylättäisiin. Tätä todennäköisyyttä kutsutaan *merkitsevyystasoksi*. Merkitsevyystasoa merkitään tavallisesti kirjaimella α.

Valitaan merkitsevyystasoksi esimerkiksi $\alpha = 0.05$. Tällöin havaintoaineistolta vaaditaan niin voimakkaita todisteita nollahypoteesia vastaan, että vain 5%:ssa tapauksia (1:ssä tapauksessa 20:stä) nollahypoteesi hylätään, *kun se pätee eli on tosi*. Jos testin *P*-arvo on α tai pienempi, sanomme, että *testin tulos on nollahypoteesin kannalta tilastollisesti merkitsevä merkitsevyystasolla* α.

Merkitsevyystasoon liitetään usein ns. *kriittiset arvot*, jotka erottavat ne testisuureen arvot, jotka johtavat nollahypoteesin hylkäämiseen niistä testisuureen arvoista, jotka jättävät nollahypoteesin voimaan. Sitä testisuureen arvojen aluetta, joka johtaa nollahypoteesin hylkäämiseen sanotaan *hylkäysalueeksi* tai *kriittiseksi alueeksi*. Sitä testisuureen arvojen aluetta, joka johtaa nollahypoteesin jäämiseen voimaan sanotaan *hyväksymisalueeksi*. Kriittiset arvot määrätään testisuureen jakauman perusteella. Hylkäysalueen muoto riippuu vaihtoehtoisen hypoteesin muodosta. Tätä kysymystä käsitellään myöhemmin tässä kappaleessa.

Testin tulos voidaan siis ilmoittaa seuraavilla vaihtoehtoisilla tavoilla:

1. Tulos voidaan ilmoittaa käyttäen testisuureen arvoa vastaavaa P-arvoa:

“Testin tulos on merkitsevä, koska sen P-arvo on p.”

Tämä on suositeltava tapa ilmoittaa testin tulos.

2. Tulos voidaan ilmoittaa käyttäen merkitsevyystasoa:

“Testin tulos on merkitsevä merkitsevyystasolla α.”

Tämä on vähemmän suositeltava tapa ilmoittaa testin tulos.

Jos testin tulos on merkitsevä, nollahypoteesi H_0 *hylätään* valitulla merkitsevyystasolla. Tällöin *hyväksytään* vaihtoehtoinen hypoteesi H_1. Huomaa, että emme kuitenkaan sano, että nollahypoteesi hyväksytään — nollahypoteesi *jää voimaan*.

TAVANOMAISET MERKITSEVYYSTASOT

Testauksessa käytetään usein ns. *tavanomaisia merkitsevyystasoja*. Nämä merkitsevyystasot sekä niihin liittyvät puhetavat ja merkinnät käyvät ilmi seuraavasta taulukosta:

Taulukko 1. Tavanomaiset merkitsevyystasot.

Merkitsevyystaso α	Sanonta	Merkintä
0.05	Melkein merkitsevä	*
0.01	Merkitsevä	**
0.001	Erittäin merkitsevä	***

Jos siis testi johtaa nollahypoteesin hylkäämiseen merkitsevyystasolla 0.01, mutta ei merkitsevyystasolla 0.001, sanomme, että testin tulos oli merkitsevä, mutta ei erittäin merkitsevä. Tähtimerkintöjä käytetään etenkin silloin, kun useiden testien tulokset ilmoitetaan taulukon muodossa.

Esimerkissämme voimme sanoa, että testin tulos oli merkitsevä (**), mutta ei erittäin merkitsevä.

Merkitsevyystasojen käyttö on vanhanaikainen tapa ilmaista testin tulos. Nykyaikainen tapa on olla käyttämättä etukäteen valittuja merkitsevyystasoja, vaan rakentaa johtopäätös suoraan testin P-arvon varaan. Tämä suositus perustuu siihen, että P-arvo antaa aina rikkaamman kuvan testin tuloksesta. Suositus saa lisävahvistusta siitä, että tilasto-ohjelmistot tulostavat tavallisesti testisuureen arvon ja siihen liittyvän P-arvon.

Oletetaan, että olemme etukäteen valinneet merkitsevyystasoksi 0.05 ja testin P-arvoksi tulee 0.06. Nollahypoteesi on silloin jätettävä voimaan. Jos tällöin kerromme vain sen, että testi ei johtanut nollahypoteesin hylkäämiseen merkitsevyystasolla 0.05, annamme vajavaisen kuvan testin tuloksesta. Tutkimusselostetta lukevat saavat paljon enemmän informaatiota, jos heille kerrotaan testin P-arvo 0.06, joka kertoo, että näin poikkeuksellisen testisuureen arvon (tai vielä poikkeuksellisemman) voi saada vain 6 kertaa 100:sta (olettaen, että nollahypoteesi on tosi).

On kuitenkin syytä ottaa tavanomaisiin merkitsevyystasoihin liittyvistä puhetavoista opiksi se, että nollahypoteesia ei ole tapana hylätä, ellei testin tulos todista kyllin voimakkaasti nollahypoteesia vastaan. Tämä on *tieteen konservatiivisuusperiaatteen* mukaista: Nollahypoteesista on syytä pitää kiinni, kunnes on esitetty kyllin vahvoja perusteita sen hylkäämiseksi.

YKSI- JA KAKSISUUNTAISET TESTIT

Vaihtoehtoinen hypoteesi voi olla *kaksisuuntainen*, kuten esimerkissämme, jossa vaihtoehtoisena hypoteesina oli se, että työskentely-ympäristön meluisuudella tai meluttomuudella on vaikutusta yksinkertaisissa laskutoimituksissa tehtyjen virheiden lukumäärään, ts.

$$H_1: \quad \mu_1 \neq \mu_2.$$

Jos vaihtoehtoiseksi hypoteesiksi otetaan se, että työskentely meluttomassa ympäristössä tuottaa vähemmän virheitä, ts., jos

$$H_1: \quad \mu_1 < \mu_2,$$

puhutaan *yksisuuntaisesta* vaihtoehtoisesta hypoteesista.

Miten vaihtoehtoisen hypoteesin muoto vaikuttaa testiin? Oletetaan, että testisuure on symmetrinen normaaliarvonsa suhteen. Tällöin vaihtoehtoisen hypoteesin yksi- tai kaksisuuntaisuus vaikuttaa siihen, pidetäänkö testisuureen poikkeamia normaaliarvostaan molempiin suuntiin vai vain toiseen suuntaan todisteina nollahypoteesia vastaan. Tämän mukaan puhutaan *yksi- ja kaksisuuntaisista testeistä.*

Valinta yksi- ja kaksisuuntaisten vaihtoehtoisten hypoteesien välillä on usein vaikea. Karkeasti voidaan sanoa, että yksisuuntainen vaihtoehtoinen hypoteesi on usein luonteva silloin, kun toivotaan (tai pelätään), että nollahypoteesi hylätään ja vaihtoehtoinen hypoteesi hyväksytään. Tällainen on tilanne sellaisissa vertailuissa, joissa vaihtoehdon hyväksymisestä seuraa jonkin asiantilan paraneminen (tai huononominen) entiseen verrattuna.

Esimerkiksi verrattaessa uutta lääkettä vanhaan lääkkeeseen kiinnostuksen kohteena on parantaako uusi lääke *enemmän* potilaita kuin vanha tai onko uudella lääkkeellä *vähemmän* sivuvaikutuksia kuin vanhalla. Nollahypoteesi on kummassakin taoauksessa muotoa:

Uusi ja vanha lääke vaikuttavat potilaisiin samalla tavalla.

Vaihtoehtoiseksi hypoteesiksi (jonka toivotaan tällaisessa tilanteessa tulevan hyväksytyksi) on tällöin luonteva ottaa seuraava yksisuuntainen hypoteesi:

Uuden lääkkeen vaikutus potilaisiin on parempi kuin vanhan.

Jos tällaisessa tilanteessa havaittaisiin, että uusi lääke parantaa vähemmän potilaita kuin vanha, ei ole tarpeen tehdä mitään tilastollista testiä. Tällöin on muutenkin selvää, että vanhaa lääkettä ei kannata korvata uudella lääkkeellä!

Tarkastellaan vielä yksi- ja kaksisuuntaisuuden vaikutusta kriittisen alueen muotoon tilanteessa, jossa testisuureen normaaliarvo on 0 ja sen jakauma on jatkuva ja symmetrinen normaaliarvon suhteen. Seuraavassa oletetaan, että testille on valittu merkitsevyystasoksi α.

Tarkastellaan ensin kaksisuuntaisia testejä. Tällöin suuret poikkeamat testisuureen normaaliarvosta 0 kumpaan suuntaan tahansa todistavat nollahypoteesia vastaan. Tällöin kriittinen alue määrätään seuraavalla tavalla: Merkitään testisuuretta

Z:lla. Määrätään satunnaismuuttujan Z jakauman (kertymäfunktion) avulla pisteet $-z_{\alpha/2}$ ja $+z_{\alpha/2}$ siten, että

$$P(-z_{\alpha/2} \leq Z \leq -z_{\alpha/2}) = 1 - \alpha.$$

Tällöin pisteet $-z_{\alpha/2}$ ja $+z_{\alpha/2}$ ovat testin *kriittiset arvot* ja *hylkäysalue* eli *kriittinen alue* on muotoa

$$Z < -z_{\alpha/2} \text{ tai } Z > +z_{\alpha/2}.$$

Nollahypoteesi siis hylätään merkitsevyystasolla α, jos testisuureen arvo joutuu hylkäysalueelle. Jos näin ei käy, jää nollahypoteesi voimaan.

Tarkastellaan nyt yksisuuntaisia testejä, jolloin ainoastaan poikkeama toiseen suuntaan testisuureen normaaliarvosta 0 todistaa nollahypoteesia vastaan. Tällöin kriittinen alue määrätään seuraavalla tavalla: Merkitään testisuuretta jälleen Z:lla. Määrätään satunnaismuuttujan Z jakauman (kertymäfunktion) avulla joko piste $-z_\alpha$ tai piste $+z_\alpha$ siten, että

$$P(-z_\alpha \leq Z) = 1 - \alpha$$

tai

$$P(+z_\alpha \geq Z) = 1 - \alpha.$$

Kumpi pisteistä määrätään, riippuu vaihtoehtoisen hypoteesin suunnasta. Pisteet $-z_\alpha$ ja $+z_{\alpha/2}$ ovat testin *kriittisiä arvoja* ja *hylkäysalue* eli *kriittinen alue* on joko muotoa

$$Z \leq -z_\alpha$$

tai muotoa

$$Z \geq +z_\alpha.$$

Nollahypoteesi siis hylätään merkitsevyystasolla α, jos testisuureen arvo joutuu hylkäysalueelle. Jos näin ei käy, jää nollahypoteesi voimaan.

Samaan merkitsevyystasoon α liittyvät yksi- ja kaksisuuntaisten testien kriittiset arvot eroavat toisistaan. Yksisuuntaisten testien kriittiset arvot ovat samaa merkitsevyystasoa vastaavan kaksisuuntaisen testin kriittisten arvojen sisäpuolella, ts.

$$-z_{\alpha/2} < -z_\alpha < +z_\alpha < +z_{\alpha/2}.$$

Tämä merkitsee sitä, että yksisuuntainen testi saattaa hylätä tilanteessa, jossa kaksisuuntainen testi ei hylkää.

Oletetaan esimerkiksi, että testisuure noudattaa standardoitua normaalijakaumaa ja valitaan $\alpha = 0.05$. Tällöin

$$z_{\alpha/2} = 1.960$$

ja

$$z_\alpha = 1.645.$$

Esimerkissämme testisuureen arvoksi saatiin -2.51. Sitä vastaava P-arvo on 0.006. Tarkastellaan ensin kaksisuuntaista vaihtoehtoista hypoteesia $\mu_1 \neq \mu_2$. Merkitsevyystasoa 0.01 vastaaviksi kriittisiksi rajoiksi saadaan standardoidun normaalijakauman taulukoista pisteet -2.58 ja $+2.58$. Koska $-2.58 < -2.51$, nollahypoteesia ei voida hylätä kaksisuuntaisessa testissä, jossa merkitsevyystasona on 0.01. Tarkastellaan nyt yksisuuntaista vaihtoehtoista hypoteesia $\mu_1 < \mu_2$. Merkitsevyystasoa 0.01 vastaava kriittinen raja on -2.32. Koska $-2.32 > -2.51$, nollahypoteesi voidaan hylätä yksisuuntaisessa testissä, jossa merkitsevyystasona on 0.01.

Huomaa, että tässä kappaleessa esitettyjä tarkasteluita ei tarvita, jos johtopäätökset perustetaan etukäteen valittujen merkitsevyystasojen sijasta testisuureen arvoa vastaavaan P-arvoon. Tuiemme jatkossa nojaamaan johtopäätöksiä tehtäessä pääsääntöisesti testisuureen arvoa vastaavaan P-arvoon.

VIRHEET TESTAUKSESSA

Jos nollahypoteesi hylätään silloin, kun se pätee, tehdään *hylkäysvirhe*. Testin P-arvo kertoo kuinka suuri voi hylkäysvirheen todennäköisyys olla korkeintaan. Hylkäysvirheen todennäköisyys α voidaan ilmaista *ehdollisena todennäköisyytenä*

$$P(H_0 \text{ hylätään} | H_0 \text{ on tosi}) = \alpha.$$

Tieteen *konservatiivisuusperiaatteen* mukaan hylkäysvirhe halutaan pitää mahdollisimman pienenä. Siksi testisuureen P-arvon on oltava kyllin pieni, ennenkuin nollahypoteesi hylätään. Tämä tulee esille tavanomaisia merkitsevyystasoja käytettäessä siitä, että testiä ei kutsuta kuin *melkein merkitseväksi*, jos nollahypoteesi tulee hylätyksi merkitsevyystasolla 0.05.

Jos nollahypoteesi jää voimaan, kun se ei päde, tehdään *hyväksymisvirhe*. Hyväksymisvirheen todennäköisyys β voidaan ilmaista *ehdollisena todennäköisyytenä*

$$P(H_0 \text{ jää voimaan} | H_0 \text{ ei ole tosi}) = \beta.$$

Todennäköisyyttä

$$P(H_0 \text{ hylätään} | H_0 \text{ ei ole tosi}) = 1 - \beta$$

sanotaan testin *voimakkuudeksi*. Hyvä testi on voimakas, koska voimakkaalla testillä on pieni hyväksymisvirhe.

Hyvä testi tekee sekä hylkäys- että hyväksymisvirheen todennäköisyydet mahdollisimman pieniksi. Sivuutamme tässä virheellisten johtopäätösten todennäköisyyksien tarkemman käsittelyn. Toteamme kuitenkin, että valitettavasti on mahdotonta tehdä hylkäys- ja hyväksymisvirheen todennäköisyyksiä samanaikaisesti mielivaltaisen pieniksi. Käytännössä tyydytään kompromissiin, jossa ensisijaisesti varotaan hylkäysvirhettä vaatimalla, että hylkäyspäätökseen johtavan P-arvon on oltava kyllin pieni.

Koska testejä tehtäessä varotaan ensisijaisesti sitä, että nollahypoteesi tulee hylätyksi, jos se on tosi, kutsutaan hylkäysvirhettä myös *1. lajin virheeksi*. Tällöin hyväksymisvirhettä kutsutaan *2. lajin virheeksi*.

Nollahypoteesi voi kussakin testaustilanteessa olla todellisuudessa joko voimassa tai sitten se ei ole voimassa. Testin tuloksena nollahypoteesi voi joko jäädä voimaan tai sitten se hylätään. Siten testin tulosta ja siihen liittyviä virheitä voidaan kuvata seuraavalla nelikentällä:

Taulukko 3. Todellisuus ja testin tulokset.

	Nollahypoteesi pätee	Nollahypoteesi ei päde
Nollahypoteesi jää voimaan	Tehdään oikea johtopäätös	Tehdään hyväksymisvirhe
Nollahypoteesi hylätään	Tehdään hylkäysvirhe	Tehdään oikea johtopäätös

MERKITSEVYYSTESTI

Seuraavassa on esitetty yhteenveto tilastollisen testin työvaiheista:

1. Formuloidaaan yleisen hypoteesin H puitteissa nollahypoteesi H_0 ja vaihtoehtoinen hypoteesi H_1.

2. Suunnitellaan testi vahvistamaan havaintoaineiston sisältämiä todisteita H_0:aa vastaan. H_0 hylätään ja H_1 hyväksytään, jos todisteet nollahypoteesia vastaan ovat kyllin voimakkaita.

3. Valitaan testisuure, johon testi perustetaan. Testisuure mittaa havaintoaineiston ja nollahypoteesin H_0 yhteensopivuutta.

   Valitaan merkitsevyystaso α (ei ole pakollinen vaihe).

4. Määrätään testisuureen arvoa vastaava *P*-arvo. *P*-arvo on todennäköisyys, että testisuure saa havaitun arvon tai vielä poikkeuksellisemman arvon, kun oletetaan, että H_0 on tosi.

5. Jos *P*-arvo on kyllin pieni, nollahypoteesi hylätään.

   Jos P-arvo on merkitsevyystasoa α pienempi, nollahypoteesi hylätään merkitsevyystasolla α (ei ole pakollinen vaihe).

4.2 TESTEJÄ LAATUEROASTEIKOLLISILLE AINEISTOILLE

4.2.1 JOHDANTO

Mittaus on tehty *nominaali-* eli *laatueroasteikolla*, jos mittaus kertoo mihin luokkaan havainto kuuluu mitattavan ominaisuuden perusteella. Seuraavassa esitellään laatueroasteikollisille tilastollisille muuttujille sopivia testejä. Tarkastelun kohteeksi otetaan seuraavat testaustilanteet:

1. Suhteellisen frekvenssin ja tapahtuman todennäköisyyden vertailu.
2. Todennäköisyyksien vertailu kahdessa ryhmässä.
3. Frekvenssijakauman ja teoreettisen jakauman vertailu: yhteensopivuustesti.
4. Frekvenssijakaumien vertailu: yhteensopivuustesti.
5. Kahden muuttujan riippumattomuuden testaaminen: yhteensopivuustesti.

On syytä huomata, että esitetyt testit sopivat nominaaliasteikollisten muuttujien liäksi myös ordinaali-, väli- ja suhdeasteikollisille muuttujille.

4.2.2 SUHTEELLISEN FREKVENSSIN JA TAPAHTUMAN TODENNÄKÖISYYDEN VERTAILU

Oletetaan, että tapahtuman

A = "Perusjoukon alkiolla on ominaisuus $\mathscr{P}$"

todennäköisyys $P(A) = p$. Jos perusjoukko on äärellinen, p kuvaa niiden perusjoukon alkioiden suhteellista osuutta, joilla on ominaisuus $\mathscr{P}$. Olkoon nollahypoteesina

$$H_0\colon p = p_0$$

ja olkoon

$$\hat{p} = \frac{X}{n}$$

tapahtuman A *havaittu suhteellinen frekvenssi* n:n riippumattoman havainnon joukossa.

Testi nollahypoteesille H_0 voidaan perustaa siihen, että testisuure

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

noudattaa suurissa otoksissa *approksimatiivisesti* standardoitua normaalijakaumaa $N(0,1)$, *jos nollahypoteesi* H_0 *pätee*.

Testisuure on samaa muotoa kuin tavanomainen t-testisuure satunnaismuuttujan odotusarvolle. Osoittajana on otossuureen ja nollahypoteesin mukaisen parametrin arvon erotus. Jos nollahypoteesi on tosi,

$$E(\hat{p}) = p_0,$$

jolloin osoittajan odotusarvo on nolla. Nimittäjässä on suhteellisen frekvenssin otosjakauman hajonta silloin, kun nollahypoteesi pätee, ts.,

$$D(\hat{p}) = \sqrt{\frac{p_0(1-p_0)}{n}},$$

jos H_0 pätee.

Testisuureen voimakkaasti nollasta poikkeavat arvot johtavat nollahypoteesin hylkäämiseen.

ESIMERKKI 1.

Erään astmalääkkeen valmistaja väittää lääkkeen mainoksissa, että lääkettä saaneista koivun siitepölylle allergisista astmaatikoista 90%:lla oireet lievittyvät vähintään 8 tunniksi. Kun lääkettä annettiin 200 astmaatikolle, 160:llä oireet lievittyivät vähintään 8 tunniksi ja 40:llä näin ei tapahtunut. Onko lääkkeen valmistajan väite oikeutettu havaintojen perusteella?

Olkoon p todennäköisyys, että lääke lievittää satunnaisesti valitun astmaatikon oireita vähintään 8 tunniksi. Nollahypoteesi on

$$H_0: \quad p = 0.9.$$

Havainnoista p:lle saadaan estimaatti

$$\hat{p} = \frac{160}{200} = 0.8.$$

Tässä tapauksessa nollahypoteesin testaamisessa on kyse seuraavasta: Kuinka todennäköistä on *havaita*, että jonkin ominaisuuden omaavien alkioiden suhteellinen osuus otoksessa on 0.8, jos ko. ominaisuuden omaavien alkioiden suhteellinen osuus perusjoukossa on 0.9?

Testisuureen arvoksi saadaan

$$Z = \frac{0.8 - 0.9}{\sqrt{\frac{0.9(1-0.9)}{200}}} \approx -4.71,$$

jota vastaava P-arvo on 0.000001.

On siis erittäin epätodennäköistä saada testisuureen arvo −4.71, jos asetettu nollahypoteesi pätee: Näin pieni tai vielä pienempi testisuureen Z arvo saadaan nollahypoteesin pätiessä vain noin 1 kerran 1,000,000:sta. Siten nollahypoteesi voidaan hylätä. Kaikki tavanomaiset merkitsevyystasot johtavat tietysti samaan johtopäätökseen.

Voimme siis suurella varmuudella tehdä johtopäätöksen, että valmistajan väite ei päde. Lisäksi voidaan suurella varmuudella sanoa, että lääkkeen teho on vähäisempi, kuin valmistaja on ilmoittanut. ●

4.2.3 TODENNÄKÖISYYKSIEN VERTAILU

Oletetaan, että tutkimuksen kohteena oleva perusjoukko voidaan jakaa kahteen ryhmään. Oletetaan edelleen, että tapahtuman

A = "Perusjoukon alkiolla on ominaisuus $\mathscr{P}$"

todennäköisyydet mainituissa ryhmissä ovat p_1 ja p_2, ts.,

$\mathrm{P}(A) = p_1$ ryhmässä 1,

$\mathrm{P}(A) = p_2$ ryhmässä 2.

Jos perusjoukko on äärellinen, p_1 ja p_2 kuvaavat niiden perusjoukon alkioiden suhteellista osuutta, joilla on ominaisuus $\mathscr{P}$.

Otetaan testattavaksi nollahypoteesiksi

H_0: $p_1 = p_2$.

Testaustilanteessa vertaillaan siis *saman* tapahtuman todennäköisyyttä kahdessa saman perusjoukon ryhmässä (tai kahdessa perusjoukossa). Jos nollahypoteesi pätee, ryhmäjako ei anna informaatiota tutkittavan tapahtuman todennäköisyydestä. Tällöin ryhmiä ei tarvitse pitää ominaisuutta $\mathscr{P}$ tarkasteltaessa erillään toisistaan.

Oletetaan, että mainituista ryhmistä on poimittu kaksi *toisistaan riippumatonta* yksinkertaista satunnaisotosta, joiden koot ovat n_1 ja n_2. Kummastakin otoksesta määrätään niiden otosyksiköiden frekvenssi, joilla on ominaisuus $\mathscr{P}$. Merkitään tätä frekvenssiä otoksessa 1 X_1:llä ja otoksessa 2 X_2:lla. Siten otoksista saadut tiedot voidaan esittää seuraavana taulukkona:

Otos	Otoskoko	Suhteellinen frekvenssi
1	n_1	$\hat{p}_1 = X_1/n_1$
2	n_2	$\hat{p}_2 = X_2/n_2$

Tapahtuman A todennäköisyyksien p_1 ja p_2 vertailu perustetaan tapahtuman A *havaittuihin suhteellisiin frekvensseihin* $\hat{p}_1$ ja $\hat{p}_2$.

Testi H_0:lle voidaan perustaa siihen, että testisuure

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}},$$

noudattaa suurissa otoksissa *approksimatiivisesti* standardoitua normaalijakaumaa $N(0,1)$, *jos nollahypoteesi* H_0 *pätee*. Testisuureen lausekkeessa

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2}$$

on ominaisuuden $\mathscr{P}$ omaavien perusjoukon alkioiden suhteellinen osuus *yhdistetyssä otoksessa*.

Testisuure on samaa muotoa kuin testisuure kahden riippumattoman satunnaismuuttujan odotusarvojen vertaamiseksi. Testisuureen osoittaja on havaittujen suhteellisten frekvenssien erotus. Jos nollahypoteesi on tosi,

$$E(\hat{p}_1) = E(\hat{p}_2) = p \text{ (merkintä)},$$

jolloin osoittajan odotusarvo on 0. Testisuureen nimittäjä on suhteellisten frekvenssien erotuksen hajonnan estimaattori, kun nollahypoteesi pätee, ts.

$$D(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}},$$

jos H_0 pätee.

Voimakkaasti nollasta poikkeavat testisuureen arvot johtavat nollahypoteesin hylkäämiseen.

ESIMERKKI 1.

Erääseen vakavaan tautiin on kehitetty uusi lääke. Lääkkeen tehon selvittämiseksi järjestetään koe, johon valitaan 200 tautia sairastavaa potilasta. Potilaat jaetaan satunnaisesti kahteen ryhmään A ja B, joissa kummassakin on 100 potilasta. Ryhmälle A annetaan uutta lääkettä ja ryhmälle B vanhaa lääkettä. Kumpaankin ryhmää hoidetaan muuten samalla tavalla. Ryhmä A muodostaa siis *hoitoryhmän* ja ryhmä B *kontrolliryhmän*.

Ryhmässä A parantuneiden potilaiden lukumäärä on 75, kun taas ryhmässä B parantuneiden potilaiden lukumäärä on 65. Ovatko paranemistuloksia uudella lääkkeellä todella parempia kuin vanhalla lääkkeellä, vai voiko paranemistuloksissa havaittu ero olla seurausta sattumasta?

Olkoon p_A todennäköisyys, että satunnaisesti valittu uutta lääkettä saava potilas paranee ja p_B todennäköisyys, että satunnaisesti valittu vanhaa lääkettä saava potilas paranee. Nollahypoteesina on tällöin

$$H_0: \quad p_A = p_B.$$

Parametrien p_A ja p_B estimaateiksi saadaan

$$\hat{p}_A = \frac{75}{100} = 0.75,$$

$$\hat{p}_B = \frac{65}{100} = 0.65.$$

Kun parantumistodennäköisyys estimoidaan yhdistetystä otoksesta, saadaan

$$\hat{p} = \frac{75+65}{100+100} = 0.7.$$

Testisuureen arvoksi saadaan siten

$$Z = \frac{0.75 - 0.65}{\sqrt{0.7 \times 0.3 \times \left(\frac{1}{100} + \frac{1}{100}\right)}} \approx 1.54.$$

Testisuureen arvoa vastaava *P*-arvo on 0.06. Siis noin 6 kertaa sadasta eli hieman useammin kuin 1:n kerran 20:stä saadaan testisuureen *Z* arvo 1.54 tai vielä suurempi *Z*:n arvo, jos nollahypoteesi pätee. Jos testin tulosta verrataan 5%:n merkitsevyystasoon, voidaan todeta, että testin tulos on *lähes* melkein merkitsevä.

Mitä johtopäätöksiä testin tuloksesta voidaan tehdä? Jos tauti on vakava ja lääkkeellä ei ole sivuvaikutuksia, voidaan uutta lääkettä hyvinkin ryhtyä antamaan potilaille, vaikka todisteet nollahypoteesia vastaan eivät ole kovin voimakkaita. ●

4.2.4 FREKVENSSIJAKAUMAN JA TEOREETTISEN JAKAUMAN VERTAILU: YHTEENSOPIVUUSTESTI

Edellä on useita kertoja tehty oletus, että tutkimuksen kohteena olevan ominaisuuden jakauma perusjoukossa on normaalinen tai ainakin approksimatiivisesti normaalinen. Normaalisuusoletus perustuu usein *keskeiseen raja-arvolauseeseen*, jonka mukaan muuttujat, jotka ovat useiden osatekijöiden summia, ovat approksimatiivisesti normaalisia, jos summassa on riittävän paljon tekijöitä ja mikään summan tekijöistä ei ole dominoiva.

Tyypillisiä esimerkkejä normaalisuusoletuksen käytöstä ovat seuraavat: Aritmeettisen keskiarvon käyttöä odotusarvon estimaattorina voidaan perustella sillä, että aritmeettinen keskiarvo on normaalijakauman odotusarvoparametrin suurimman uskottavuuden estimaattori. Odotusarvolle voidaan muodostaa t-jakaumaan perustuva luottamusväli, jos perusjoukko on normaalijakautunut. Myöhemmin tulemme vetoamaan perusjoukon normaalisuuteen esimerkiksi odotusarvoa tai regressiomallin kertoimia koskevien testien yhteydessä.

Perusjoukon normaalisuutta koskevaa oletusta on syytä testata aina, kun käytettävä tilastollinen menetelmä edellyttää normaalisuusoletuksen tekemistä. Alla esitetään yleinen *yhteensopivuustesti*, jolla testataan onko havaintoarvojen jakauma yhteensopiva jonkin teoreettisen jakauman kanssa.

Oletetaan, että tutkimuksen kohteena olevan perusjoukon ominaisuutta kuvaavasta satunnaismuuttujasta on käytettävissä n riippumatonta havaintoa. Oletetaan, että havaintoarvot on luokiteltu toisensa poissulkeviin luokkiin, joiden lukumäärä on k. Tällöin havaintoarvojen luokiteltu frekvenssijakauma voidaan esittää seuraavana taulukkona:

Luokka	1	2	...	k	Summa
Havaittu frekvenssi	O_1	O_2	...	O_k	n

Havaittu frekvenssi O_i on siis niiden havaintojen lukumäärä, jotka kuuluvat luokkaan i. Havaittujen frekvenssien O_i summa yhtyy havaintojen kokonaislukumäärään n:

$$O_1 + O_2 + \ldots + O_k = n.$$

Tehtävänä on testata nollahypoteesia, jonka mukaan perusjoukon jakauma on annettua muotoa:

H_0: Perusjoukon jakauma on $\mathrm{F}(x)$,

jossa $\mathrm{F}(x)$ on perusjoukon jakauman määrittelevä kertymäfunktio. Olkoon p_i todennäköisyys, että tarkastelun kohteena oleva satunnaismuuttuja saa arvon luokasta i, kun oletetaan, että *nollahypoteesi pätee*. Tämä todennäköisyys voidaan määrätä kertymäfunktion $\mathrm{F}(x)$ avulla.

Jos nollahypoteesi H_0 pätee, luokkaan i tulevien havaintojen *odotettu frekvenssi* E_i voidaan määrätä kaavasta

$$E_i = np_i.$$

Koska nollahypoteesin määrittelemän jakauman parametreista joudutaan tavallisesti estimoimaan havaintoaineistosta, todennäköisyydet p_i on korvattava estimaateillaan P_i. Myös odotettujen frekvenssien E_i summa yhtyy havaintojen kokonaislukumäärään n:

$$E_1 + E_2 + \ldots + E_k = n.$$

Testi H_0:lle perustuu havaittujen frekvenssien O_i ja odotettujen frekvenssien E_i vertailuun.

Testi H_0:lle voidaan perustaa siihen, että testisuure

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

noudattaa suurissa otoksissa *approksimatiivisesti* ns. χ^2-jakaumaa, *jos* H_0 *pätee*. Testisuureen lausekkessa

O_i = havaintojen havaittu (engl. *observed*) frekvenssi luokassa i,

E_i = havaintojen odotettu (engl. *expected*) frekvenssi luokassa i,

k = luokkien lukumäärä.

Tähän testisuureeseen perustuvaa testiä nollahypoteesille H_0 sanotaan *χ^2-yhteensopivuustestiksi*.

Testisuure X^2 voidaan kirjoittaa myös seuraavaan yhtäpitävään muotoon:

$$X^2 = n\sum_{i=1}^{k} \frac{(\hat{p}_i - P_i)^2}{P_i},$$

jossa

$$\hat{p}_i = O_i/n$$

on luokkaan i kuuluvien havaintojen *havaittu suhteellinen luokkafrekvenssi* ja P_i on nollahypoteesin mukaan estimoitu todennäköisyys sille, että tarkasteltava satunnaismuuttuja saa arvon, joka kuuluu luokkaan i.

Tarkasteltava satunnaismuuttuja voi olla yhtä hyvin *nominaali-*, *ordinaali-*, *väli-* tai *suhdeasteikollinen* muuttuja. Olennaista on vain se, että havaintoarvot voidaan luokitella. Luokat määräytyvät nominaaliasteikollisen muuttujan tapauksessa luonnollisella tavalla muuttujan asteikosta itsestään. Väli- tai suhdeasteikollisten muuttujien tapauksessa, mutta usein myös ordinaaliasteikollisten muuttujien tapauksessa, testin käyttäjä (tai käytettävä ohjelmisto) valitsee sopivan luokituksen.

Testisuure noudattaa siis *approksimatiivisesti* χ^2-jakaumaa, jos nollahypoteesi pätee. χ^2-jakauma riippuu ns. *vapausasteiden* lukumäärästä. Vapausasteiden lukumäärä on tässä tapauksessa

$$k - 1 - m,$$

jossa m on odotettujen frekvenssien E_i määräämiseksi estimoitujen parametrien lukumäärä ko. satunnaismuuttujan jakaumassa. Normaalijakauman tapauksessa joudutaan aineistosta estimoimaan tavallisesti sekä odotusarvo μ että varianssi σ^2, jolloin $m = 2$. Binomijakauman tapauksessa joudutaan aineistosta estimoimaan tavallisesti todennäköisyys p, jolloin $m = 1$.

Jakaumatulos approksimatiivinen eli se pätee suurissa otoksissa. Approksimaatio on riittävän hyvä, jos kaikki odotetut frekvenssit ovat suurempia kuin 5. Jos tämä ehto ei toteudu, saadaan tämän ehdon toteuttava luokitus aikaan yhdistämällä luokkia.

Testisuure mittaa sopivasti skaalattuna havaittujen frekvenssien poikkeamia odotetuista frekvensseistä kussakin luokassa. Jos nollahypoteesi pätee,

$$\mathrm{E}(O_i) = \mathrm{E}(E_i) = p_i,\ i = 1,2,\ldots,k,$$

ts. havaittujen frekvenssien odotetut arvot yhtyvät odotettuihin frekvenssiin. Siten nollahypoteesin pätiessä testisuureen lausekkeen summatermien osoittajissa olevat erotukset ovat keskimäärin pieniä. Jos nämä erotukset ovat suuria, on todennäköistä, että nollahypoteesi ei päde. Voidaan osoittaa, että testisuureen odotettavissa oleva arvo on nollahypoteesin pätiessä $k - 1 - m$, ts.

$$\mathrm{E}(X^2) = k - 1 - m,$$

jos H_0 pätee. Testisuureen odotettavissa olevaa arvoa selvästi *suuremmat* arvot johtavat nollahypoteesin hylkäämiseen.

χ^2-JAKAUMA

χ^2-jakauma on eräs jatkuvista todennäköisyysjakaumista. Jakauma on muodoltaan epäsymmetrinen ja se on määritelty ainoastaan x-akselin positiivisella puoliskolla. χ^2-jakauman muoto riippuu vapausasteiden lukumäärästä f.

Jakaumaan liittyviä todennäköisyyksiä on taulukoitu Kirjassa 1. Taulukko liittyy merkitsevyystasojen käyttöön. Taulukko antaa χ^2-jakauman alapuolisen alueen pinta-aloja eli todennäköisyyksiä pisteestä χ_p^2 *oikealle*:

$$\mathrm{P}(X^2 > \chi_p^2) = p.$$

Esimerkiksi, jos $f = 30$ niin pisteestä 31.410 oikealle oleva todennäköisyysmassa on suuruudeltaan 0.05. On hyvä huomata, että χ_p^2 on merkitsevyystasoa p vastaava *kriittinen arvo*.

Taulukko antaa karkeudestaan huolimatta mahdollisuuden arvioida χ^2-testin P-arvon suuruusluokan. Oletetaan esimerkiksi, että $f = 10$ ja $\chi^2 = 30.1$. Pisteestä 29.588 oikealle jää todennäköisyysmassa, joka on suuruudeltaan 0.001, jos $f = 10$. Siten testin P-arvo on pienempi kuin 0.001.

Miksi taulukossa on annettu myös sellaisia kriittisiä arvoja, joista *oikealle* oleva todennäköisyysmassa on niinkin suuri kuin 0.99? Tällöin ko. kriittisestä arvosta *vasemmalle* oleva todennäköisyysmassa on 0.01. Tämä merkitsee sitä, että χ^2-testisuure saa ko. kriittistä arvoa *pienempiä* arvoja melko pienellä todennäköisyydellä, jos nollahypoteesi pätee. Miten tällainen tilanne on tulkittava? Tulkinta seuraa siitä, että χ^2-testisuure reagoi myös siihen, että havaittu frekvenssijakauma sopii *liian hyvin* yhteen nollahypoteesin mukaisen teoreettisen jakauman kanssa. On syytä muistaa, että satunnaisvaihtelun takia ei ole todennäköistä, että havainnot ja teoria sopivat *täydellisesti* toisiinsa. Vaikka nollahypoteesin mukaan on odotettavissa, että havaitut frekvenssit vaihtelevat odotettujen frekvenssien ympärillä, ei ole todennäköistä, että havaitut frekvenssit yhtyvät (tai ovat edes kovin lähellä) odotettuja frekvenssejä. Siten χ^2-testisuureen poikkeuksellisen pienet arvot viittaavat havaintojen ja nollahypoteesin *liian* hyvään yhteensopivuuteen.

Jos tällainen tilanne havaitaan, saattaa syynä olla se, että havaintoaineisto on väärennetty siksi, että se saataisiin sopimaan paremmin teoriaan. Satunnaisvaihtelun merkitys aliarvioidaan tavallisesti, kun aineistoa väärennetään. Tähän viitattiin eräänä niistä väärinkäsityksistä, joka liittyy suurten lukujen lakiin.

ESIMERKKI 1.

Erään yrityksen johto epäili, että työntekijöiden keskuuteen oli levinnyt tapa venyttää viikonlopun viettoa ilmoittautumalla sairaaksi perjantaisin ja maanantaisin. Asiaa tutkittiin 4:n viikon aikana rekisteröimällä poissaolojen lukumäärät työpäivinä. Keskiarvotiedot ko. ajanjaksolta on ilmoitettu seuraavassa taulukossa:

Taulukko 1: Havaittu poissaolojen jakauma.

Viikonpäivä	ma	ti	ke	to	pe	Summa
Poissaolojen lkm (ka)	49	35	32	39	45	200

Testattavaksi nollahypoteesiksi asetetaan siis se, että poissaolot jakaantuvat *tasaisesti* työpäiville. Tämän oletuksen mukaan poissaolojen jakauman pitäisi muistuttaa seuraavaa jakaumaa:

Taulukko 2: Odotettu poissaolojen jakauma, jos nollahypoteesi pätee.

Viikonpäivä	ma	ti	ke	to	pe	Summa
Poissaolojen lkm	40	40	40	40	40	200

Tehtävänä on verrata havaittuja ja nollahypoteesin mukaan odotettuja frekvenssejä toisiinsa.

χ^2-testisuureen arvo nollahypoteesille on

$$X^2 = \frac{(49-40)^2}{40} + \cdots + \frac{(45-40)^2}{40} = 4.9.$$

Nollahypoteesin pätiessä χ^2-testisuure X^2 on jakaantunut kuten χ^2-jakauma vapausastein

$$k-1-m = 5-1-0 = 4,$$

koska odotettavissa olevia frekvenssejä muodostettaessa ei ole tarvinnut estimoida yhtään nollahypoteesin määrittelemän teoreettisen jakauman (diskreetti tasainen jakauma) parametria.

Koska taulukoiden mukaan

$$\chi^2_{0.05}(4) = 9.49 > 4.9,$$

testisuureen arvoa vastaava *P*-arvo on selvästi suurempi kuin 0.05. Nollahypoteesia ei siis kannata hylätä tässä tilanteessa: Poissaolojen jakautumisessa eri työpäiville ei voida havaita "viikonloppuefektiä". ●

4.2.5 FREKVENSSIJAKAUMIEN VERTAILU: YHTEENSOPIVUUSTESTI

Oletetaan, että tutkimuksen kohteena oleva perusjoukko voidaan jakaa kahteen tai useampaan ryhmään. Tehtävänä on verrata jonkin ominaisuuden jakautumista ko. ryhmissä. Tarkoituksena on muodostaa testi nollahypoteesille, jonka mukaan tutkittava ominaisuus jakautuu ko. ryhmissä *samalla tavalla*. Vertailun tekemiseksi ko. ryhmistä poimitaan *toisistaan riippumattomat* satunnaisotokset, joista määrätään ryhmäkohtaiset frekvenssijakaumat *samaa* luokitusta käyttäen. Testi nollahypoteesille perustuu näiden frekvenssijakaumien vertailuun. Huomaa, että nollahypoteesin pätiessä ryhmäjako ei anna informaatiota tutkittavan ominaisuuden jakautumisesta. Tällöin ryhmiä ei tarvitse erottaa toisistaan tutkittavaa ominaisuutta tarkasteltaessa, vaan ryhmät voidaan yhdistää.

Oletetaan siis, että luokkafrekvenssit eri ryhmissä voidaan esittää seuraavana taulukkona:

Luokka → Ryhmä ↓	1	2	...	c	Summa
1	O_{11}	O_{12}	...	O_{1c}	n_1
2	O_{21}	O_{22}	...	O_{2c}	n_2
...	...	...	...	...	...
r	O_{r1}	O_{r2}	...	O_{rc}	n_r
Summa	C_1	C_2	...	C_c	n

Taulukossa

r = ryhmien lukumäärä,

c = luokkien lukumäärä,

O_{ij} = havaittu frekvenssi ryhmän i luokassa j,

n_i = otoskoko ryhmässä i,

C_j = havaittu frekvenssi yhdistetyn otoksen luokassa j,

n = havaintojen kokonaislukumäärä.

Frekvenssejä O_{ij} on tapana kutsua *solufrekvensseiksi* ja frekvenssejä C_i *sarakefrekvensseiksi*.

Tällaista taulukkoa kutsutaan tavallisesti *kaksiulotteiseksi frekvenssitaulukoksi*. Se määrätään luokittelemalla havainnot *ristiin* ryhmien ja luokkien suhteen: Soluun, jonka määrittelee i. rivi ja j. sarake kuuluvat ne havainnot, jotka kuuluvat ryhmään i ja luokkaan j.

Huomaa, että solufrekvenssien *rivisummat* yhtyvät ryhmäkohtaisiin otoskokoihin n_i:

$$O_{i1} + O_{i2} + \ldots + O_{ic} = n_i,\ i = 1,2,\ldots,r$$

ja *sarakefrekvenssit* C_j saadaan solufrekvenssien sarakkeittain määrättyinä summina:

$$C_j = O_{1j} + O_{2j} + \ldots + O_{rj},\ j = 1,2,\ldots,c.$$

Sarakefrekvenssi C_j ilmaisee luokkaan j kuuluvien havaintojen lukumäärän siinä yhdistetyssä otoksessa, joka saadaan yhdistämällä ryhmistä $1,2,\ldots,r$ poimitut havainnot. Havaintojen kokonaislukumäärä n saadaan sekä ryhmäkohtaisten otoskokojen että sarakefrekvenssien summana:

$$n = n_1 + n_2 + \ldots + n_r = C_1 + C_2 + \ldots + C_c.$$

Huomaa, että ym. taulukko voidaan muuntaa *prosenttitaulukoksi* niin, että jokaisen *rivin* frekvenssit ilmaistaan prosentteina vastaavasta rivisummasta. Nämä prosentit muodostavat *suhteellisten* frekvenssien jakauman. Jatkon kannalta on hyvä muistaa, että suhteelliset frekvenssit voidaan tulkita empiirisiksi todennäköisyyksiksi.

Nollahypoteesin mukaan tutkittava ominaisuus jakautuu luokkiin $1,2,\ldots,c$ eri ryhmissä samalla tavalla. Tämä voidaan ilmaista myös sanomalla, että ko. ominaisuuden jakautuminen luokkiin perusjoukon eri ryhmissä on *riippumatonta* ryhmäjaosta.

Olkoon $p_{j|i}$ *ehdollinen* todennäköisyys, että tutkimuksen kohteena olevan perusjoukon alkio kuuluu luokkaan j, kun alkio kuuluu ryhmään i, ts.

$$p_{j|i} = \mathrm{P}(\text{Alkio kuuluu luokkaan } j \mid \text{Alkio kuuluu ryhmään } i).$$

Olkoon $p_{\cdot j}$ todennäköisyys, että alkio kuuluu luokkaan j, kun ryhmäjakoa ei oteta huomioon. Tällöin nollahypoteesi voidaan ilmaista muodossa

$$p_{j|i} = p_{\cdot j},\ j = 1,2,\ldots,c,\ i = 1,2,\ldots,r.$$

Tällöin todennäköisyys, että tarkasteltavaa ominaisuutta vastaava muuttuja saa arvon luokasta j ei riipu ryhmästä i. Nollahypoteesi merkitsee sitä, että luokitukseen liittyvät diskreetit todennäköisyysjakaumat ovat jokaisessa ryhmässä samat. Jos nollahypoteesi pätee, havaittujen frekvenssijakaumien eri ryhmissä (rivien yo. taulukossa) pitäisi muistuttaa *muodoltaan* toisiaan. Jos havaittujen frekvenssijakaumat eri ryhmissä eroavat toisistaan muodoltaan kyllin paljon, nollahypoteesi on syytä asettaa kyseenalaiseksi.

Edellä kuvattua nollahypoteesia voidaan testata *χ^2-yhteensopivuustestillä*, jossa *havaittuja solufrekvenssejä* verrataan nollahypoteesin mukaan *odotettujen solufrekvenssien estimaatteihin*.

Miten nollahypoteesin mukaan odotettujen solufrekvenssien estimaatit määrätään? Jos nollahypoteesi pätee, havaitut frekvenssijakaumat eri ryhmissä (rivit yo. taulukossa) eivät saa siis erota *muodoltaan* toisistaan kovin paljon. Tällöin ryhmäkohtaisten havaittujen frekvenssijakaumien $O_{i1}, O_{i2}, \ldots, O_{ic}$ pitää siis muistuttaa *muodoltaan* sarakesummien $C_1, C_2, \ldots, C_c$ muodostamaa frekvenssijakaumaa. Siten *nollahypoteesin pätiessä odotetut frekvenssit* E_{ij} voidaan määrätä yhtälöstä

$$E_{ij} = \frac{C_j}{n} n_i, j = 1,2,\ldots,c, i = 1,2,\ldots,r.$$

Tällöin odotetut frekvenssijakaumat ovat eri ryhmissä saman muotoisia. Tämä nähdään siitä, että tällöin *odotetut suhteelliset frekvenssit* yhtyvät kullakin rivillä:

$$\frac{E_{ij}}{n_i} = \frac{C_j}{n}, j = 1,2,\ldots,c, i = 1,2,\ldots,r.$$

χ^2-testisuure saa tässä tapauksessa muodon

$$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

jossa E_{ij} on määritelty edellä. Testisuure X^2 on jakautunut suurissa otoksissa *approksimatiivisesti* kuten χ^2-jakauma $(r-1)(c-1)$:llä vapausteella, *jos nollahypoteesi pätee*.

Testisuureen jakaumaa koskeva tulos on siis approksimatiivinen eli se pätee, jos otoskoko n on kyllin iso. Approksimaatio on yleensä riittävän hyvä, jos kaikki *keskimääräisten* odotettujen frekvenssien estimaatit C_j/r ovat suurempia kuin 5 ja kaikki odotettujen solufrekvenssien estimaatit E_{ij} ovat suurempia kuin 1. Jos alkuperäinen taulukko ei toteuta näitä ehtoja, saadaan ehdot toteutumaan yhdistämällä luokkia. 2×2-taulukossa vaaditaan kuitenkin enemmän: Kaikkien odotettujen solufrekvenssien estimaattien E_{ij} on oltava suurempia kuin 5.

Jos nollahypoteesi pätee, testisuureen X^2 odotettavissa oleva arvo on

$$E(X^2) = (r-1)(c-1).$$

Tätä arvoa selvästi suuremmat arvot johtavat nollahypoteesin hylkäämiseen.

χ^2-testisuure voidaan kirjoittaa myös seuraavaan edellisen kanssa yhtäpitävään muotoon:

$$X^2 = n\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(\hat{p}_{ij} - P_{ij})^2}{P_{ij}},$$

jossa

$$\hat{p}_{ij} = \frac{O_{ij}}{n}, j = 1,2,\ldots,c, i = 1,2,\ldots,r,$$

$$P_{ij} = \frac{E_{ij}}{n} = \frac{C_j}{n}\cdot\frac{n_i}{n}, j = 1,2,\ldots,c, i = 1,2,\ldots,r.$$

Huomaa, että

$$\frac{C_j}{n}$$

on *empiirinen todennäköisyys* sille, että satunnaisesti valittu havaintoyksikkö kuuluu luokkaan j, kun eri ryhmiin kuuluvat havainnot on yhdistetty yhdeksi otokseksi. Voidaan osoittaa, että tämä on suurimman uskottavuuden estimaattori todennäköisyydelle $p_{\cdot j}$, että havainto kuuluu luokkaan j. On hyvä muistaa, että eri ryhmistä poimittujen otosten yhdistäminen on luvallista, jos nollahypoteesi pätee. Myös

$$\frac{n_i}{n}$$

voidaan tulkita empiiriseksi todennäköisyydeksi: Se on empiirinen todennäköisyys, että satunnaisesti valittu havaintoyksikkö kuuluu ryhmään i. Siten yhtälö

$$p_{ij} = \frac{C_j}{n} \cdot \frac{n_i}{n}$$

vastaa nollahypoteesia, jonka mukaan *tarkasteltavan ominaisuuden jakautuminen eri luokkiin on riippumatonta ryhmäjaosta*. Tämä ehto on yhtäpitävää sen kanssa, että suhteellisten frekvenssien jakaumat eivät riipu ryhmästä.

Myös tässä tilanteessa ainoa tarkasteltavalle ominaisuudelle asettava vaatimus on se, että sitä vastaavan muuttujan arvot voidaan luokitella. Siten esitetty testi sopii sekä nominaali-, ordinaali-, väli- että suhdeasteikoillisille muuttujille.

Kuten frekvenssijakauman ja teoreettisen jakauman yhteensopivuutta testattaessa, nytkin on mahdollista, että χ^2-testisuure saa liian pieniä arvoja. Tällöin havaitut frekvenssijakaumat muistuttavat eri ryhmissä muodoltaan "liian paljon" toisiaan. On syytä muistaa, että vaikka nollahypoteesin mukaan havaittujen frekvenssijakaumien pitääkin muistuttaa muodoltaan toisiaan, satunnaisvaihtelun takia jakaumat eivät tavallisesti ole aivan samanmuotoisia.

ESIMERKKI 1.

Tarkastellaan naisten ja miesten itselleen urheilussa asettamia päämääriä.

Urheilussa voidaan päämääränä pitää kilpailemista (= K) tai suoritusta itseään (= S). Urheilua harrastavien naisten ja miesten suhtautumista näihin päämääriin tutkittiin toisistaan riippumattomien satunnaisotosten avulla. Kummankin otoksen koko oli 67. Otoksiin poimitut naiset ja miehet vastasivat kyselyyn, jolla tutkittiin suhtautumista em. päämääriin. Sekä naiset että miehet jaettiin kyselyn perusteella 4:ään luokkaan sen mukaan pitivätkö he kilpailemista tärkeänä (= KT) vai vähemmän tärkeänä (= KV) vai pitivätkö he suoritusta itseään tärkeänä (= ST) vai vähemmän tärkeänä (= SV).

Kyselyn tulokset voidaan ilmaista seuraavana taulukkona:

Taulukko 1: Havaitut frekvenssit

Luokka→ Ryhmä↓	KT	KV	ST	SV	Summa
Miehet	31	18	5	13	67
Naiset	14	7	21	25	67
Summa	45	25	26	38	134

Taulukkoa katselemalla saa vahvan mielikuvan siitä, että naiset eivät pidä kilpailemista urheilun motiivina yhtä tärkeänä kuin miehet ja miehet eivät puolestaan ole kovin kiinnostuneita urheilusuorituksista itsestään. Seuraavassa tutkitaan, saako tämä mielikuva vahvistusta myös tilastollisesta analyysista.

Nollahypoteesiksi otetaan

H_0: Suhtautuminen urheilun päämääriin on riippumatonta sukupuolesta.

Huomaa, että nollahypoteesin pätiessä tieto henkilön sukupuolesta ei siis sisällä informaatiota henkilön suhtautumista urheilun päämääriin.

Muodostetaan yo taulukkoa vastaava nollahypoteesin mukainen odotettujen frekvenssien taulukko:

Taulukko 2: Odotetut frekvenssit

Luokka→ Ryhmä↓	KT	KV	ST	SV	Summa
Miehet	22.5	12.5	13	19	67
Naiset	22.5	12.5	13	19	67
Summa	45	25	26	38	134

Esimerkiksi odotettu *miesten* frekvenssi luokassa KT on nollahypoteesin mukaan

$$67 \times 45/134 = 22.5$$

ja odotettu *naisten* frekvenssi luokassa ST on nollahypoteesin mukaan

$$67 \times 26/134 = 13.$$

Huomaa, että tässä tapauksessa nollahypoteesin mukaan odotettavissa olevat frekvenssijakaumat eivät ole vain samanmuotoisia, vaan täsmälleen samoja. Tämä johtuu siitä, että otokset olivat kummassakin ryhmässä yhtä suuria.

χ^2-testisuure saa arvon

$$X^2 = \frac{(31-22.5)^2}{22.5} + \cdots + \frac{(25-19)^2}{19} = 24.9.$$

Testisuure noudattaa *nollahypoteesin pätiessä* suurissa otoksissa approksimatiivisesti χ^2-jakaumaa vapausastein

$$(r-1)(c-1) = (2-1)(4-1) = 3.$$

Testisuureen arvoa vastaava P-arvo on pienempi kuin 0.000, joten nollahypoteesi voidaan hylätä: Miehet ja naiset suhtautuvat erilailla urheilun päämääriin. ●

4.2.6 RIIPPUMATTOMUUDEN TESTAAMINEN

Oletetaan, että tutkimuksen kohteena on *kahden* ominaisuuden $\mathcal{A}$ ja $\mathcal{B}$ *riippuvuus* toisistaan. Riippuvuus on perusjoukon ominaisuus, joka auttaa arvioimaan perusjoukon alkioiden ominaisuuksia toistensa perusteella. Tehtävänä on muodostaa testi nollahypoteesille, jonka mukaan ominaisuudet eivät riipu toisistaan. Tätä varten perusjoukosta poimitaan riippumaton satunnaisotos ja havainnot luokitellaan samanaikaisesti molempien ominaisuuksien suhteen. Testi nollahypoteesille perustetaan tämän *ristiintaulukoinnin* tuloksena syntyvään *kaksiulotteiseen* frekvenssijakaumaan. Huomaa, että nollahypoteesin pätiessä havaintojen jakautuminen toisen ominaisuuden suhteen ei anna informaatiota havaintojen jakautumisesta toisen ominaisuuden suhteen. Tällöin ominaisuuksia voidaan tarkastella perusjoukossa toisistaan riippumatta.

Oletetaan siis, että havainnot voidaan luokitella ominaisuuksien $\mathcal{A}$ ja $\mathcal{B}$ suhteen seuraavaksi kaksiulotteiseksi frekvenssitaulukoksi:

$\mathcal{A}$-luokka → $\mathcal{B}$-luokka ↓	1	2	...	c	Summa
1	O_{11}	O_{12}	...	O_{1c}	R_1
2	O_{21}	O_{22}	...	O_{2c}	R_2
...	...	...	...	...	...
r	O_{r1}	O_{r2}	...	O_{rc}	R_r
Summa	C_1	C_2	...	C_c	n

Taulukossa

c = $\mathcal{A}$-luokkien lukumäärä,

r = $\mathcal{B}$-luokkien lukumäärä,

O_{ij} = havaittu frekvenssi $\mathcal{B}$-ominaisuuden luokassa i ja $\mathcal{A}$-ominaisuuden luokassa j,

C_j = havaittu frekvenssi $\mathcal{A}$-ominaisuuden luokassa j,

R_i = havaittu frekvenssi $\mathcal{B}$-ominaisuuden luokassa i,

n = havaintojen kokonaislukumäärä.

Frekvenssejä O_{ij} on tapana kutsua *solufrekvensseiksi*, frekvenssejä C_j *sarakefrekvensseiksi* ja frekvenssejä R_i *rivifrekvensseiksi*. Yhteisellä nimellä sarake- ja rivifrekvenssejä kutsutaan *reunafrekvensseiksi*.

Tällaista taulukkoa kutsutaan tavallisesti *kaksiulotteiseksi frekvenssitaulukoksi*. Se määrätään luokittelemalla havainnot *ristiin* kummankin luokituksen suhteen:

Soluun, jonka määrittelee i. rivi ja j. sarake kuuluvat ne havainnot, jotka kuuluvat $\mathcal{B}$-ominaisuuden luokkaan i ja $\mathcal{A}$-ominaisuuden luokkaan j.

Huomaa, että solufrekvenssien *rivisummat* yhtyvät rivifrekvensseihin:

$$O_{i1} + O_{i2} + \ldots + O_{ic} = R_i,\ i = 1,2,\ldots,r$$

ja solufrekvenssien *sarakesummat* yhtyvät sarakefrekvensseihin:

$$O_{1j} + O_{2j} + \ldots + O_{rj} = C_j,\ j = 1,2,\ldots,c.$$

Lisäksi havaintojen kokonaislukumäärä n saadaan molempien reunafrekvenssien summana:

$$n = R_1 + R_2 + \ldots + R_r = C_1 + C_2 + \ldots + C_c.$$

Huomaa, että kaksiulotteinen frekvenssitaulukko voidaan muuntaa *prosenttitaulukoksi* kolmella eri tavalla:

1. Jokaisen *rivin* solufrekvenssit ilmaistaan prosentteina vastaavasta rivifrekvenssistä.
2. Jokaisen *sarakkeen* solufrekvenssit ilmaistaan prosentteina vastaavasta sarakefrekvenssistä.
3. Jokainen solufrekvenssi ilmaistaan prosentteina havaintojen kokonaislukumäärästä.

Kaikissa tapauksissa prosentit muodostavat erään *suhteellisten* frekvenssien jakauman. On hyvä muistaa jatkossa, että suhteelliset frekvenssit voidaan tulkita empiirisiksi todennäköisyyksiksi.

Nollahypoteesin mukaan perusjoukon alkiot jakautuvat ominaisuuksien $\mathcal{A}$ ja $\mathcal{B}$ suhteen luokkiin *riippumattomalla tavalla*. Jos nollahypoteesi pätee, pitäisi havaittujen frekvenssijakaumien muistuttaa muodoltaan toisiaan sekä rivien että sarakkeiden suunnassa. Jos havaitut frekvenssijakaumat eroavat muodoltaan toisistaan kyllin paljon, nollahypoteesi joudutaan asettamaan kyseenalaiseksi.

Olkoon p_{ij} todennäköisyys, että tutkimuksen kohteena olevan perusjoukon alkio kuuluu samanaikaisesti $\mathcal{A}$-ominaisuuden luokkaan j ja $\mathcal{B}$-ominaisuuden luokkaan i. Olkoon $p_{\cdot j}$ todennäköisyys, että tutkimuksen kohteena olevan perusjoukon alkio kuuluu $\mathcal{A}$-ominaisuuden luokkaan j, kun $\mathcal{B}$-ominaisuutta ei oteta huomioon ja olkoon $p_{i\cdot}$ todennäköisyys, että tutkimuksen kohteena oleva perusjoukon alkio kuuluu $\mathcal{B}$-ominaisuuden luokkaan i, kun $\mathcal{A}$-ominaisuutta ei oteta huomioon. Jos

$$p_{ij} = p_{i\cdot}p_{\cdot j},\ i = 1,2,\ldots,r,\ j = 1,2,\ldots,c,$$

nollahypoteesi pätee. Tällöin todennäköisyys, että $\mathcal{A}$-ominaisuutta vastaava muuttuja saa arvon $\mathcal{A}$-ominaisuuden luokassa i on riippumaton siitä, minkä arvon $\mathcal{B}$-ominaisuutta vastaava muuttuja saa $\mathcal{B}$-ominaisuuden luokassa j. Nollahypoteesi merkitsee siis $\mathcal{A}$- ja $\mathcal{B}$-ominaisuuksia vastaavien muuttujien riippumattomuutta. Jos nollahypoteesi pätee, pitäisi myös havaittujen frekvenssijakaumien kuvastaa tätä riippumattomuutta.

Nollahypoteesia voidaan jälleen testata *χ^2-yhteensopivuustestillä*, jossa *havaittuja solufrekvenssejä* verrataan nollahypoteesin mukaan *odotettujen solufrekvenssien estimaatteihin*.

Miten nollahypoteesin mukaan odotetut solufrekvenssit määrätään? Jos nollahypoteesi pätee, havaitut frekvenssijakaumat eivät voi erota *muodoltaan* kovin paljon toisistaan kummassakaan suunnassa. Tällöin jokaisen yo. taulukon riveistä pitää siis muistuttaa *muodoltaan* sarakesummien $C_1, C_2, \ldots, C_c$ muodostamaa frekvenssijakaumaa ja jokaisen yo. taulukon sarakkeista pitää taas muistuttaa *muodoltaan* rivisummien $R_1, R_2, \ldots, R_r$ muodostamaa frekvenssijakaumaa. Siten *nollahypoteesin pätiessä odotetut frekvenssit* E_{ij} voidaan määrätä yhtälöstä

$$E_{ij} = \frac{R_i C_j}{n}, \; i = 1,2,\ldots,r, j = 1,2,\ldots,c.$$

χ^2-testisuure saa myös tässä tapauksessa muodon

$$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

jossa E_{ij} on määritelty edellä. Testisuure X^2 on jakautunut suurissa otoksissa *approksimatiivisesti* kuten χ^2-jakauma $(r-1)(c-1)$:llä vapausteella, *jos nollahypoteesi pätee*.

Testisuureen jakaumaa koskeva tulos on approksimatiivinen eli se pätee, jos otoskoko n on kyllin iso. Approksimaatio on riittävän hyvä, jos kaikki *keskimääräisten* odotettujen frekvenssien estimaatit R_i/c ja C_j/r ovat suurempia kuin 5 ja kaikki odotettujen solufrekvenssien estimaatit E_{ij} ovat suurempia kuin 1. Jos alkuperäinen taulukko ei toteuta näitä ehtoja, saadaan ehdot toteutumaan yhdistämällä luokkia. 2×2-taulukossa vaaditaan kuitenkin enemmän: Kaikkien odotettujen solufrekvenssien estimaattien E_{ij} on oltava suurempia kuin 5.

Jos nollahypoteesi pätee, testisuureen X^2 odotettavissa oleva arvo on

$$\mathrm{E}(X^2) = (r-1)(c-1).$$

Tätä arvoa selvästi suuremmat arvot johtavat nollahypoteesin hylkäämiseen.

χ^2-testisuure voidaan kirjoittaa myös edellisen kanssa yhtäpitävään muotoon

$$X^2 = n\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(\hat{p}_{ij} - P_{ij})^2}{P_{ij}},$$

jossa

$$\hat{p}_{ij} = \frac{O_{ij}}{n}, \; i = 1,2,\ldots,r, j = 1,2,\ldots,c,$$

$$P_{ij} = \frac{E_{ij}}{n} = \frac{C_j}{n}\cdot\frac{R_i}{n}, \; i = 1,2,\ldots,r, j = 1,2,\ldots,c.$$

Huomaa, että

$$\frac{C_j}{n}$$

on empiirinen todennäköisyys sille, että satunnaisesti valittu havaintoyksikkö kuuluu $\mathcal{A}$-ominaisuuden luokkaan j. Voidaan osoittaa, että tämä suhteellinen frekvenssi on todennäköisyyden $p_{\cdot j}$ suurimman uskottavuuden estimaattori. Samoin

$$\frac{R_i}{n}$$

on empiirinen todennäköisyys sille, että satunnaisesti valittu havaintoyksikkö kuuluu $\mathcal{B}$-ominaisuuden luokkaan i. Voidaan osoittaa, että tämä suhteellinen frekvenssi on todennäköisyyden $p_{i\cdot}$ suurimman uskottavuuden estimaattori. Siten yhtälö

$$P_{ij} = \frac{C_j}{n} \cdot \frac{R_i}{n}$$

vastaa nollahypoteesia, jonka mukaan *tarkasteltavat ominaisuudet ovat riippumattomia*.

Myös tässä tilanteessa ainoa tarkasteltaville ominaisuuksille asettava vaatimus on se, että sitä vastaavien muuttujien arvot voidaan luokitella. Siten esitetty testi sopii sekä nominaali-, ordinaali-, väli- että suhdeasteikoillisille muuttujille.

Kuten yhteensopivuustestienkin tapauksessa, nytkin on mahdollista, että χ^2-testisuure saa liian pieniä arvoja. Tällöin havaittu solufrekvenssien jakauma toteuttaa liian hyvin riippumattomuusoletuksen. On syytä muistaa, että vaikka nollahypoteesin mukaan havaittujen frekvenssijakaumien pitää muistuttaa muodoltaan toisiaan sekä rivien että sarakkeiden suunnissa, satunnaisvaihtelun takia jakaumat eivät tavallisesti ole aivan samanmuotoisia.

JAKAUMIEN VERTAILUTESTI JA RIIPPUMATTOMUUSTESTI: YHTEYDET JA EROT

On syytä huomata, että sekä frekvenssijakaumien yhteensopivuutta että riippumattomuutta testattaessa syntyvä kaksiulotteinen frekvenssitaulukko on rivi- ja sarakesummineen aivan samaa muotoa. Frekvenssitaulukosta ei voi siis sellaisenaan päätellä kummasta tilanteesta on kyse. Samoin testisuure on kummassakin tapauksessa sama ja myös testisuureen jakauma on nollahypoteesin pätiessä sama.

Kuitenkin testaustilanteet eroavat toisistaan selvästi: Jakaumien yhteensopivuutta testattaessa kyseessä on *saman ominaisuuden* jakautumisen vertailu kahdessa tai useammassa ryhmässä. Tällöin havaintoaineisto koostuu riippumattomista *ryhmäkohtaisista* satunnaisotoksista. Riippumattomuustestissä kyseessä on *kahden ominaisuuden* riippumattomuuden tutkiminen. Tällöin havaintoaineisto koostuu *yhdestä* riippumattomasta satunnaisotoksesta. Huomaa kuitenkin, että ryhmäjakoa voidaan aina asettaa vastaamaan sopivasti määritelty laatueroasteikollinen muuttuja. Siten ero näiden testaustilanteiden välillä ei ole aina selvä.

Vaikka frekvenssitaulukosta ei sellaisenaan voi päätellä kummasta tilanteesta on kyse, taulukoissa on kuitenkin seuraava ero: Jos kyseessä on jakaumien vertailuun liittyvä tilanne, voidaan rivisummia eli ryhmäkohtaisia otoskokoja pitää *kiinteinä*, kun taas sarakesummat eli ryhmät yhdistämällä saadut otoksen luokkafrekvenssit ovat *satunnaisia*. Sen sijaan riippumattomuustestin yhteydessä sekä rivi- että sarakesummat ovat *satunnaisia* ja vain otoskoko n on *kiinteä*.[1]

---

[1] Muuttuja on *kiinteä* (ei-satunnainen), jos sen arvot *valitaan*.

ESIMERKKI 1.

Erään yhtiön johto harkitsee sulautumista toiseen yhtiöön. Yhtiön johto haluaa saada selville osakkaiden mielipiteet jo ennen yhtiökokousta. Johto on erityisen kiinnostunut siitä riippuuko mielipide omistettujen osakkeiden lukumäärästä.

Asian selvittämiseksi osakkaiden joukosta poimitaan 250 osakkaan satunnaisotos. Tulokset ilmenevät seuraavasta taulukosta, josta näkyvät myös käytetyt luokitukset:

Taulukko 1. Havaitut frekvenssit.

Mielipide → Osakkeiden lkm ↓	Puolesta	Ei mieli-pidettä	Vastaan	Summa
Alle 200	38	9	29	76
200 — 1000	30	7	42	79
Yli 1000	32	4	59	95
Summa	100	20	130	250

Taulukon perusteella näyttää siltä, että suuret osakkaat vastustavat fuusiota pieniä enemmän. Tämä nähdään esimerkiksi siitä, että otoksessa on alle 200 osaketta omistavien joukossa fuusion puolesta olevia osakkaita *enemmän* kuin fuusiota vastaan olevia osakkaita. Sen sijaan otoksessa on yli 1000 osaketta omistavien joukossa fuusion puolesta olevia osakkaita *vähemmän* kuin fuusiota vastaan olevia osakkaita.

Riippuuko vastustus todellakin omistettujen osakkeiden lukumäärästä, voidaan selvittää testaamalla nollahypoteesia

H_0: Mielipide ei riipu omistettujen osakkeiden lukumäärästä.

Onko taulukosta silmämääräisesti tehty havainto ristiriidassa nollahypoteesin kanssa nähdään rakentamalla nollahypoteesille χ^2-testi.

Yllä esitettyä taulukkoa vastaavat *nollahypoteesin mukaan odotetut frekvenssit* ilmenevät seuraavasta taulukosta:

Taulukko 2. Odotettavissa olevat frekvenssit.

Mielipide → Osakkeiden lkm ↓	Puolesta	Ei mieli-pidettä	Vastaan	Summa
Alle 200	30.40	6.08	39.52	76
200 — 1000	31.60	6.32	41.08	79
Yli 1000	38.00	7.60	49.40	95
Summa	100	20	130	250

Esimerkiksi siihen soluun, jonka määräävät yli 1000 osaketta omistavat fuusion puolesta olevat osakkaat, odotettu frekvenssi on määrätty seuraavalla laskutoimituksella:

$$\frac{95 \cdot 100}{250} = 38.$$

χ^2-testisuureen arvoksi saadaan

$$X^2 = \frac{(38-30.40)^2}{30.40} + \cdots + \frac{(59-49.40)^2}{49.40} = 10.80.$$

Testisuure on jakautunut *nollahypoteesin pätiessä* kuten χ^2-jakauma vapausastein

$$(r-1)(c-1) = (3-1)(3-1) = 4.$$

Todennäköisyyttä 0.05 vastaava kriittinen arvo χ^2-jakaumalle on 11.070, ts.

$$\mathrm{P}(X^2 > 11.070) = 0.05,$$

jos nollahypoteesi pätee. Siten nollahypoteesia ei voi hylätä ainakaan 5%:n merkitsevyystasoa käyttäen. Sen sijaan nollahypoteesi olisi kyllä mahdollista hylätä 10% merkitsevyystasoa käyttäen.

Johtopäätöksenä on siis se, että osakkaiden mielipide fuusiosta ei näytä riippuvan omistettujen osakkaiden lukumäärästä ainakaan kovin voimakkaasti.

4.3 TESTEJÄ JÄRJESTYSASTEIKOLLISILLE MUUTTUJILLE

4.3.1 JOHDANTO

Mittaus on tehty *ordinaali-* eli *järjestysasteikolla*, jos mittaus kertoo onko havaintoyksiköllä ennemmän mitattavaa ominaisuutta kuin jollakin toisella havaintoyksiköllä. Seuraavassa esitellään muutamia järjestysasteikollisille muuttujille sopivia testejä. On syytä huomata, että myös kaikki laatueroasteikollisille muuttujille esitetyt testit sopivat järjestysasteikollisille muuttujille. Toisaalta järjestysasteikollisille muuttujille esitetyt testit sopivat myös väli- ja suhdeasteikollisille muuttujille.

Tarkastelun kohteeksi otetaan seuraavat testaustilanteet:

1. Kahden otoksen vertailu: Mannin ja Whitneyn testi.
2. Riippumattomuuden testaus: Järjestyskorrelaatiot.

4.3.2 KAHDEN OTOKSEN VERTAILU: MANNIN JA WHITNEYN TESTI

Oletetaan, että tutkimuksen kohteena oleva perusjoukko voidaan jakaa kahteen ryhmään ja tarkastelun kohteena on *saman* järjestysasteikollisen muuttujan arvojen jakautuminen näissä ryhmissä. Merkitään ko. muuttujan arvoja ryhmässä 1 kirjaimella X ja ryhmässä 2 kirjaimella Y. Mielenkiinnon kohteena on se ovatko jakaumat niin samanlaisia, että ryhmät voidaan yhdistää ko. muuttujaa tarkasteltaessa yhdeksi perusjoukoksi.

Poimitaan mainituista ryhmistä toisistaan riippumattomat yksinkertaiset satunnaisotokset. Olkoot muuttujan X havaitut arvot otoksessa 1

$$X_1, X_2, \ldots, X_n$$

järjestettynä suuruusjärjestykseen pienimmästä suurimpaan ja olkoot muuttujan Y havaitut arvot otoksessa 2

$$Y_1, Y_2, \ldots, Y_m$$

järjestettynä suuruusjärjetykseen pienimmästä suurimpaan.

Nollahypoteesina on se, että tarkastelun kohteena olevan muuttujan jakautumat ovat samoja, ts. nollahypoteesi on muotoa

$$H_0: \quad F_X(x) = F_Y(y),$$

jossa

$$F_X(x) = X\text{-muuttujan kertymäfunktio},$$

$$F_Y(y) = Y\text{-muuttujan kertymäfunktio}.$$

Miten otosinformaatiota voidaan käyttää hyväksi tämän nollahypoteesin testaamisessa?

Sopivaksi menettelyksi on osoittautunut seuraava: Yhdistetään otokset yhdeksi otokseksi ja tutkitaan miten X- ja Y-havainnot seuraavat toisiaan. Ajatellaan tilanteita jossa *kaikki* X-havainnot ovat pienempiä kuin Y-havainnot tai, jossa *kaikki* Y-havainnot ovat pienempiä kuin X-havainnot. Tuntuu ilmeiseltä, että nollahypoteesi ei voi päteä tällaisissa tilanteissa. Jos perujoukot noudattavat samaa jakaumaa, on X- ja Y-havaintoarvojen sekoituttava sopivasti toisiinsa. *Mannin ja Whitneyn testisuure* on kehitetty mittaamaan tällaista sekoittumista.

Mannin ja Whitneyn testisuureesta on kaksi erilaista, mutta yhtäpitävää muotoa:

1. Määritellään satunnaismuuttuja

$$D_{ij}^{(1)} = \begin{cases} 1, \text{ jos } X_i < Y_j \\ 0, \text{ jos } X_i > Y_j \end{cases}$$

Mannin ja Whitneyn testisuure U_1 voidaan määritellä muuttujien $D_{ij}^{(1)}$ avulla seuraavalla tavalla:

MANNIN JA WHITNEYN TESTISUURE

$$U_1 = \sum_{i=1}^{n} \sum_{j=1}^{m} D_{ij}^{(1)}.$$

Testisuure U_1 on saa sitä suurempia arvoja mitä enemmän on X-havaintoarvoja, jotka ovat Y-havaintoarvoja pienempiä. Jos *jokainen* X-havaintoarvot on jokaista Y-havaintoarvoa pienempi, $U_1 = nm$. Tämä seuraa siitä, että testisuureen lausekkeen summassa termien $D_{ij}^{(1)}$ lukumäärä on nm ja, että $D_{ij}^{(1)} = 1$ kaikille i ja j, jos kaikki X-havaintoarvot ovat pienempiä kuin kaikki Y-havaintoarvot. Sen sijaan, jos *jokainen* X-havaintoarvo on jokaista Y-havaintoarvoa suurempi, $U_1 = 0$, koska tällöin $D_{ij}^{(1)} = 0$ kaikille i ja j. Muissa tapauksissa testisuureen U_1 arvo on arvojen 0 ja nm välissä.

2. Määritellään satunnaismuuttuja

$\mathrm{R}(X_i)$ = havainnon X_i järjestysnumero *yhdistetyssä* otoksessa.

Mannin ja Whitneyn testisuure U_1 voidaan määritellä muuttujien $\mathrm{R}(X_i)$ avulla seuraavasti:

MANNIN JA WHITNEYN TESTISUURE

$$U_1 = nm + \frac{1}{2}n(n+1) - \sum_{i=1}^{n} \mathrm{R}(X_i).$$

Huomaa, että U_1 saa sitä suurempia arvoja mitä enemmän on X-havaintoarvoja, jotka ovat Y-havaintoarvoja pienempiä. Jos *jokainen* X-havaintoarvo on jokaista Y-havaintoarvoa pienempi, $U_1 = nm$. Tämä seuraa siitä, että tällöin X-havaintoarvot sijoittuvat yhdistetyssä otoksessa järjestysnumeroille $1, 2, \ldots, n$ ja

$$1+2+\ldots+n=\frac{1}{2}n(n+1).$$ [1]

Sen sijaan, jos *jokainen* X-havaintoarvo on jokaista Y-havaintoarvoa suurempi, $U_1 = 0$. Tämä seuraa siitä, että tällöin X-havaintoarvot sijoittuvat yhdistetyssä otoksessa järjestysnumeroille $m+1, m+2, \ldots, m+n$ ja

$$\begin{aligned} & m+1+m+2+\ldots+m+n \\ & \quad = mn+1+2+\ldots+n \\ & \quad = mn+\frac{1}{2}n(n+1). \end{aligned}$$

Muissa tapauksissa testisuureen U_1 arvo on arvojen 0 ja nm välissä.

Kuten edellä esitetyistä määritelmistä nähdään, Mannin ja Whitneyn testisuure ei riipu lainkaan X- ja Y-havaintoarvojen suuruudesta, vaan ainoastaan niiden keskinäisestä järjestyksestä. Niissä tapauksissa, joissa tarkastelun kohteena oleva muuttuja on väli- tai suhdeasteikoillinen, havaintoarvoja tarvitaan havaintoarvojen järjestyksen muodostamisessa. Mutta sen jälkeen, kun havainnot on järjestetty, havaintoarvoja ei tarvita; ainoat Mannin ja Whitneyn testisuureen määräämiseen tarvittavat tiedot sisältyvät havaintojen järjestysnumeroihin yhdistetyssä otoksessa ja tunnistimeen, joka kertoo kummasta otoksesta havainto on peräisin.

Edellä sanotusta tulee selvästi esille se, että Mannin ja Whitneyn testisuure sopii muuttujille, joiden mitta-asteikko on järjestys-, väli- tai suhdeasteikollinen. Sen sijaan Mannin ja Whitneyn testisuuretta ei voida mielekkäällä tavalla määritellä laatueroasteikollisille muuttujille.

[1] Tulos on *aritmeettisen sarjan summan kaavan* sovellus.

Vaihtamalla X- ja Y-havaintojen roolit voidaan määritellä Mannin ja Whitneyn testisuure U_2 seuraavissa yhtäpitävissä muodoissa:

1. Määritellään satunnaismuuttuja

$$D_{ji}^{(2)} = \begin{cases} 1, \text{ jos } Y_j < X_i \\ 0, \text{ jos } Y_j < X_i \end{cases}.$$

Mannin ja Whitneyn testisuure U_2 voidaan määritellä muuttujien $D_{ji}^{(2)}$ avulla seuraavasti:

MANNIN JA WHITNEYN TESTISUURE

$$U_2 = \sum_{j=1}^{m} \sum_{i=1}^{n} D_{ji}^{(2)}.$$

Huomaa, että U_2 on saa sitä suurempia arvoja mitä enemmän on Y-havaintoarvoja, jotka ovat X-havaintoarvoja pienempiä. Jos *jokainen* Y-havaintoarvo on jokaista X-havaintoarvoa pienempi, $U_2 = nm$. Sen sijaan, jos *jokainen* Y-havaintoarvo on jokaista X-havaintoarvoa suurempi, $U_2 = 0$. Muissa tapauksissa testisuureen U_2 arvo on näiden arvojen 0 ja nm välissä. Perustelu on samantapainen kuin testisuureen U_1 tapauksessa.

2. Määritellään satunnaismuuttuja

R(Y_j) = havainnon Y_j järjestysnumero *yhdistetyssä* otoksessa.

Mannin ja Whitneyn testisuure U_2 voidaan määritellä muuttujien R(Y_j) avulla seuraavasti:

MANNIN JA WHITNEYN TESTISUURE

$$U_2 = nm + \frac{1}{2}m(m+1) - \sum_{j=1}^{m} \mathrm{R}(Y_j).$$

Testisuure U_2 ei myöskään riipu X- ja Y-havaintoarvojen suuruudesta, vaan ainoastaan niiden keskinäisestä järjestyksestä. Lienee selvää, että testisuureet U_1 ja U_2 ovat toisiinsa nähden komplementaarisessa suhteessa: Jos testisuureen U_1 arvo on lähellä arvoa 0, niin testisuureen U_2 arvo on lähellä arvoa nm ja kääntäen, jos testisuureen U_1 arvo on lähellä arvoa nm, niin testisuureen U_2 arvo on lähellä arvoa 0.

Seuraavassa on lueteltu joitakin Mannin ja Whitneyn testisuureiden keskeiset ominaisuudet:

MANNIN JA WHITNEYN TESTISUUREEN OMINAISUUKSIA

1. $U_1 + U_2 = nm.$

2. $\mathrm{E}(U_1) = \mathrm{E}(U_2) = \frac{1}{2}nm$, *jos* H_0 *pätee*.

3. $$D^2(U_1) = D^2(U_2) = \frac{1}{12}nm(n+m+1), \textit{jos } H_0 \textit{ pätee}.$$

4. *Jos* H_0 *pätee*, niin standardoitu satunnaismuuttuja

$$Z_1 = \frac{U_1 - E(U_1)}{D(U_1)}$$

noudattaa *suurissa otoksissa approksimatiivisesti* standardoitua normaalijakaumaa N(0,1).

Jos H_0 *pätee*, niin standardoitu satunnaismuuttuja

$$Z_2 = \frac{U_2 - E(U_2)}{D(U_2)}$$

noudattaa *suurissa otoksissa approksimatiivisesti* standardoitua normaalijakaumaa N(0,1). Kohdassa 4 mainitut approksimaatiot ovat käytännössä riittävän hyviä, jos

$$n > 20 \textit{ ja } m > 20.$$

Jos nollahypoteesi pätee, ts., jos otokset ovat samasta jakaumasta, kohdista (1) ja (2) seuraa, että

$$U_1 \approx U_2.$$

Mitä kaumpana testisuureiden arvot ovat toisistaan, sitä todennäköisempää on se, että nollahypoteesi ei päde.

Testi nollahypoteesille voidaan perustaa suurissa otoksissa ($n > 20$ *ja* $m > 20$) standardoituun muuttujaan Z_1 tai Z_2. Jos niiden arvot poikkeavat voimakkaasti nollasta, on nollahypoteesi hylättävä. Testisuureiden arvoja vastaavat P-arvot saadaan normaalijakauman taulukoista. Pienille otoksille ($n \leq 20$ *ja* $m \leq 20$) on rakennettu erillisiä taulukoita. Sivuutamme kuitenkin tässä niiden käytön. On syytä huomata, että testisuureet Z_1 ja Z_2 ovat aina vastakkaismerkkisiä:

$$Z_1 = -Z_2,$$

joten riittää tarkastella vain toista mainituista testisuureista.

Edellä esitetyt Mannin ja Whitneyn testisuureiden ominaisuudet perustuvat seuraavaan tosiseikkaan: Jos nollahypoteesin pätee, niin jokaisen X-havaintoarvon *odotettevissa oleva sijoitus* yhdistetyssä otoksessa on

$$E(R(X_i)) = \frac{1}{2}(n+m+1).$$

Siten summan $\sum_{i=1}^{n} R(X_i)$ odotusarvo on

$$E\left(\sum_{i=1}^{n} R(X_i)\right) = \frac{1}{2}n(n+m+1).$$

Koska satunnaismuuttujat $R(X_i)$ ovat riippumattomia, summan $\sum_{i=1}^{n} R(X_i)$ varianssi on

$$D^2\left(\sum_{i=1}^{n} R(X_i)\right) = \frac{1}{12}nm(n+m+1).$$

Voidaan osoittaa, että Mannin ja Whitneyn testi on testi myös seuraavalle nollahypoteesille:

$$H_0': \quad Md_X = Md_Y,$$

jossa Md_X ja Md_Y ovat satunnaismuuttujien X ja Y *mediaanit*[1]. Nollahypoteesin H_0' pätiessä X- ja Y-havaintoarvojen otosmediaanit eivät ole todennäköisesti kovin kaukana toisistaan. Jos otosmediaanit yhtyvät, sekä X- että Y-havaintoarvoja on oltava saman verran kummallakin puolella niiden yhteistä mediaania. Tällöin testisuureet U_1 ja U_2 yhtyvät niiden nollahypoteesin mukaan odottettavissa olevaan arvoon $nm/2$.

Koska mediaanin määrääminen vaatii, että tarkasteltava muuttuja on joko järjestys-, väli- tai suhdeasteikollinen, nähdään edellä esitetystä jälleen se, että Mannin ja Whitneyn testi ei sovi laatueroasteikollisille muuttujille.

ESIMERKKI 1.

Tarkastellaan Mannin ja Whitneyn testisuureen ja siihen liittyvien laskutoimitusten havainnollistamiseksi tilannetta, jossa on X-havaintoja on 2 ja Y-havaintoja on 3.

Koska havaintoarvoilla itsellään ei ole Mannin ja Whitneyn testisuuretta laskettaessa merkitystä, vaan ainoastaan havaintoarvojen suuruusjärjestyksellä yhdistetyssä otoksessa, kaikkien mahdollisten X- ja Y-havaintojen keskinäisten järjestysten lukumäärän antaa multinomikerroin

$$\begin{pmatrix} & 5 & \\ 2 & & 3 \end{pmatrix} = \frac{5!}{2!3!} = 10.$$

Seuraavassa taulukossa on lueteltu X- ja Y-havaintojen kaikki mahdolliset yhdistetyt suuruusjärjestykset sekä vastaavat testisuureiden U_1 ja U_2 arvot:

---

[1] Mediaani jakaa perusjoukon kahteen saman kokoiseen osaan, joista toisessa kaikki tarkasteltavan muuttujan arvot ovat mediaania pienempiä, toisessa suurempia.

Tapaus	Järjestys	U_1	U_2
1	$X_1X_2Y_1Y_2Y_3$	6	0
2	$X_1Y_1X_2Y_2Y_3$	5	1
3	$X_1Y_1Y_2X_2Y_3$	4	2
4	$X_1Y_1Y_2Y_3X_2$	3	3
5	$Y_1X_1X_2Y_2Y_3$	4	2
6	$Y_1X_1Y_2X_2Y_3$	3	3
7	$Y_1X_1Y_2Y_3X_2$	2	4
8	$Y_1Y_2X_1X_2Y_3$	2	4
9	$Y_1Y_2X_1Y_3X_2$	1	5
10	$Y_1Y_2Y_3X_1X_2$	0	6

Huomaa, että olemme olettaneet, että X- ja Y-havainnot on asetettu suuruusjärjestykseen ennen yhdistetyn järjestyksen määräämistä.

Otetaan lähempään tarkasteluun tapaus 2:

Liittämällä havaintoarvojen yhteiseen järjestykseen järjestysnumerot saadaan seuraava aputaulukko:

Havainto	X_1	Y_1	X_2	Y_2	Y_3
Järjestys-numero	1	2	3	4	5

Lasketaan testisuure U_1 ensin satunnaismuuttujiin D^1_{ij} perustuvalla kaavalla:

Koska

$$D^1_{ij} = \begin{cases} 1, \text{ jos } X_i < Y_j \\ 0, \text{ jos } X_i > Y_j \end{cases},$$

niin

$$D^1_{11} = 1, \quad D^1_{12} = 1, \quad D^1_{13} = 1,$$
$$D^1_{21} = 0, \quad D^1_{22} = 1, \quad D^1_{23} = 1.$$

Siten

$$\sum_{i=1}^{2}\sum_{j=1}^{3} D^1_{ij} = (1+1+1)+(0+1+1) = 6.$$

Satunnaismuuttujiin $R(X_i)$ perustuvaa kaavaa käytetäessä tarvittavat laskutoimitukset ovat seuraavat:

Koska

$$R(X_i) = \text{havainnon } X_i \text{ järjestysnumoro yhdistetyssä otoksessa,}$$

niin

$$R(X_1) = 1, \qquad R(X_2) = 3.$$

Siten

$$\begin{aligned} U_1 &= nm + \frac{1}{2}n(n+1) - \sum_{i=1}^{n} R(X_i) \\ &= 2\cdot 3 + \frac{1}{2}2\cdot 3 - (1+3) \\ &= 5. \end{aligned}$$

Määrätään vielä testisuure U_2 satunnaismuuttujiin $R(Y_j)$ perustuvalla kaavalla:

$$\begin{aligned} U_2 &= nm + \frac{1}{2}m(m+1) - \sum_{i=1}^{m} R(Y_j) \\ &= 2\cdot 3 + \frac{1}{2}3\cdot 4 - (2+4+5) \\ &= 1. \end{aligned}$$

Näistä tuloksista saa vahvistuksen se, että

$$U_1 + U_2 = 5 + 1 = 6 = 2\cdot 3 = nm.$$

Huomaa, että nopein tapa määrätä U-testisuureen arvo satunnaismuuttujiin D_{ij} perustuvaa kaavaa käytettäessä, on laskea ns. *voittoja*:

Y_j *voittaa* X_i:n, jos $Y_j > X_i$.

Testisuureen U_2 määräämistä tapauksessa 7:

$Y_1X_1Y_2Y_3X_2$.

Voittojen lukumäärät ovat tällöin seuraavat:

Y_1 voittaa X:t	0 kertaa,
Y_2 voittaa X:t	1 kerran,
Y_3 voittaa X:t	1 kerran.

Yhteensä Y:t voittavat X:t 2 kertaa ja siten

$$U_2 = 2.$$

Määrätään lopuksi Z-testisuureen arvo laskutoimituksien havainnollistamiseksi tapauksessa 9. Tällöin siis

$$U_2 = 5.$$

Jos nollahypoteesi pätee,

$$E(U_2) = \frac{1}{2}nm = 3,$$

$$D^2(U_2) = \frac{1}{12}nm(n+m+1) = 3.$$

Siten

$$Z_2 = \frac{U_2 - E(U_2)}{D(U_2)} = \frac{5-3}{\sqrt{3}} \approx 1.16.$$

Testisuure Z_2 noudattaa suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa.[1] Koska

$$P(Z \geq 1.16) = 0.123,$$

jossa Z noudattaa standardoitua normaalijakaumaa., nollahypoteesi voidaan jättää voimaan: X- ja Y-havainnot ovat samasta jakaumasta.

Verrataan normaaliapproksimaatioosta saatua tulosta oikeaan testisuureen U_2 arvoa 5 vastaavaan P-arvoon: Oikea P-arvo on

$$P(U_2 \geq 5) = 0.2.$$

Tässä tapauksessa johtopäätös on siis oikeaa ja approksimatiivista P-arvoa käytettäessä sama. Kannattaa kuitenkin huomata, että oikea ja approksimatiivinen P-arvo eroavat toisistaan niin paljon, että pienissä otoksissa on yleensä syytä käyttää Mannin ja Whitneyn testille erikoisesti määrättyjä taulukoita. ●

Ennen toisen esimerkin esittämistä kiinnitettäköön huomiota Mannin ja Whitneyn testin siihen piirteeseen, että sitä tehtäessä perusjoukon jakaumasta ei tehdä mitään oletuksia. Siksi Mannin ja Whitneyn testiä kutsutaan *jakaumista vapaaksi* tai *ei-parametriseksi*. Myös edellisessä kappaleessa tarkastellut χ^2-yhteensopivuustestit ovat jakaumista vapaita eli ei-parametrisia. Tyypillinen esimerkki parametrisista testeistä on taas tämän luvun johdannossa käsitelty kahden perusjoukon odotusarvojen vertailuun tarkoitettu t-testi.

Jakaumista vapaat testejä on syytä käyttää erityisesti sellaisissa tilanteissa, joissa jakaumaoletuksia ei kyetä tekemään tai otoskoko on niin pieni, että jakaumaoletuksen tekeminen ei ole järkevää. Tällaisia tilanteita kohdataan varsinkin yhteiskunta-, käyttäytymis- ja biotieteellisissä tutkimusasetelmissa.

On hyvä huomata se, että jakaumista vapaat testit sopivat erityisen hyvin sellaisille aineistoille, jotka eivät ole väli- tai suhdeasteikollisia, vaikka niitä saakin soveltaa myös sellaisille aineistoille. Jakaumista vapaat menetelmät ovat usein parametrisia menetelmiä *robustimpia* eli ne ovat vastustuskykyisempiä poikkeaville havainnoille. Tämä johtuu siitä, että jakaumista vapaat menetelmät vaativat tarkasteltavien muuttujien asteikoilta vähemmän kuin vastaavat parametriset menetelmät. Tämä nähdään esimerkiksi verrattaessa kahden perusjoukon odotusarvojen vertailemiseen tarkoitettua t-testiä Mannin ja Whitneyn testiin. Mannin ja Whitneyn testin tulos riippuu vain havaintojen keskinäisestä suuruusjärjestyksestä toisin kuin t-testin tulos.

[1] Muista, että approksimaatio on hyvä vasta, kun $n > 20$ ja $m > 20$.

ESIMERKKI 2.

Poikkeavatko poikien ja tyttöjen syntymäpainot toisistaan? Ovatko pojat jo syntyessään painavampia kuin tytöt?

Asian tutkimiseksi erään sairaalan synnytysosastolla poimittiin vastasyntyneiden joukosta toisistaan riippumattomat 7 pojan ja 4 tytön muodostamat yksinkertaiset satunnaisotokset. Poikien ja tyttöjen syntymäpainot otoksissa olivat seuraavat:

Pojat (otos 1):	3.50	3.39	3.27	4.20	3.04	4.32	3.20
Tytöt (otos 2):	3.06	3.15	3.45	2.49			

Käytetään Mannin ja Whitneyn testiä selvittämään onko otoksien perusteella järkevää olettaa, että poikien ja tyttöjen syntymäpainot noudattavat samaa jakaumaa.

Muodostetaan otoksista 1 ja 2 yhdistetty järjestetty otos:

Havainto	Y_1	X_1	Y_2	Y_3	X_2	X_3	X_4	Y_4	X_5	X_6	X_7
Havainto-arvo	2.49	3.04	3.06	3.15	3.20	3.27	3.39	3.45	3.50	4.20	4.32
Järjestys-nro	1	2	3	4	5	6	7	8	9	10	11

Sovelletaan Mannin ja Whitneyn testisuureen määräämiseksi ensin voittojen laskemiseen perustuvaa menettelyä:

Y_1 voittaa X:t: 0 kertaa,

Y_2 voittaa X:t: 1 kerran,

Y_3 voittaa X:t: 1 kerran,

Y_4 voittaa X:t: 4 kertaa.

Y:t voittavat X:t yhteensä 6 kertaa ja siten

$$U_1 = 6.$$

Satunnaismuuttujiin $R(X_i)$ perustuvaa kaavaa käytettäessä laskutoimitukset muodostuvat seuraaviksi:

$$U_1 = nm + \frac{1}{2}n(n+1) - \sum_{i=1}^{n} \mathrm{R}(X_i)$$

$$= 7 \cdot 4 + \frac{1}{2} 7 \cdot 8 - (2+5+6+7+9+10+11)$$

$$= 6.$$

Koska $nm = 7 \cdot 4 = 28$ ja $U_1 + U_2 = nm$ niin

$$U_2 = 22.$$

Jos nollahypoteesi (poikien ja tyttöjen syntymäpainot noudattavat samaa jakaumaa) pätee,

$$\mathrm{E}(U_1) = 14,$$

$$\mathrm{D}^2(U_1) = 28.$$

Sovelletaan laskutoimitusten havainnollistamiseksi näin pieniin otoksiin soveltumatonta Z-testisuuretta. Standardoitu muuttuja

$$Z_1 = \frac{6-14}{\sqrt{28}} \approx -1.51$$

noudattaa suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa.

Koska

$$\mathrm{P}(Z \leq -1.51) = 0.0655,$$

jossa Z noudattaa standardoitua normaalijakaumaa, nollahypoteesi voidaan hylätä 10%:n merkitsevyystasolla, mutta ei 5%:n merkitsevyystasolla.

Oikea P-arvo on tässä tapauksessa

$$\mathrm{P}(U_1 \leq 6) = 0.082.$$

Siten oikean P-arvon käyttäminen ei johda tässäkään tapauksessa erilaiseen johtopäätökseen. Tosin ero approksimatiivisen ja oikean P-arvon välillä on nytkin melko suuri. ●

4.3.3 RIIPPUMATTOMUUDEN TESTAAMINEN JÄRJESTYSASTEIKOLLISILLA MUUTTUJILLA

Tarkastellaan aluksi seuraavaa esimerkkiä:

ESIMERKKI 1.

Kahdelle viinintuntijalle annettiin sokkotestissä maistettavaksi 10 viiniä. He asettivat viinit seuraaviin järjestyksiin (1 = paras, 10 = huonoin):

Viini	Maistaja A	Maistaja B
1	6	3
2	2	4
3	8	7
4	9	6
5	10	10
6	7	9
7	1	2
8	5	5
9	3	8
10	4	1

Tehtävänä on tutkia ovatko viinintuntijoiden maut samanlaiset, ts. käyttävätkö he samanlaista *arvosteluskaalaa* eli *-asteikkoa* arvioidessaan viinejä. Jos tuntijoiden arvosteluasteikot ovat samanlaisia, merkitsee se sitä, että tuntijoiden maut *riippuvat toisistaan* niin, että toisen tuntijan jollekin viinille antaman sijaluvun perusteella voidaan päätellä miten hyvin tai huonosti ko. viini sijoittuu toisen tuntijan arvosteluasteikolla.

Nollahypoteesiksi otetaan

H_0: Viinintuntijoiden maut ovat toisistaan riippumattomia

ja vaihtoehtoiseksi hypoteesiksi on

H_1: Viinintuntijoiden maut ovat samanlaiset.

Taulukkoa tarkastelemalla havaitaan, että esimerkiksi viinit 5 ja 8 sijoittuvat viinintuntijoiden arvosteluasteikolla samalle sijaluvulle ja muidenkin viinien kohdalla sijoitukset ovat lähellä toisiaan. Tällaiset vertailut viittaavat selvästi viinintuntijoiden makujen samankaltaisuuteen.

Miten maistajien arvosteluasteikoiden samankaltaisuutta tai erilaisuutta voidaan tutkia? ●

Esimerkin lopussa esitettyä ongelmaa voidaan tutkia käyttämällä sopivaa mittaa viinintuntijoiden arvosteluasteikoiden riippuvuudelle (samanaisuudelle). Seuraavassa tarkastellaan kahta järjestyasteikollisille muuttujille tarkoitettua *korrelaation* eli riippuvuuden mittaa sekä esitetään näitä mittoja koskevat testit.

SPEARMANIN JÄRJESTYSKORRELAATIOKERROIN

Oletetaan, että tutkimuksen kohteena on kahden muuttujan X ja Y riippuvuus ja olkoot X ja Y vähintään järjestysasteikollisia muuttujia. Oletetaan, että muuttujista X ja Y on kerätty n toisistaan riippumatonta havaintoparia. Annetaan kummankin muuttujan arvoille järjestysnumerot niiden suuruusjärjestyksen mukaan. Olkoot

X_i = X-muuttujan arvon järjestysnumero parissa i,

Y_i = Y-muuttujan arvon järjestysnumero parissa i.

Siten havaintopari (X_i, Y_i) liittyvät samaan *tapaukseen*.

Määritellään *Spearmanin järjestyskorrelaatiokerroin* ρ kaavalla

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n},$$

jossa

$d_i = X_i - Y_i$ = järjestysnumeroiden erotus parissa i,

n = havaintoparien lukumäärä.

Spearmanin järjestyskorrelaatiokerrointa kutsutaan usein Spearmanin ρ:ksi (ρ on kreikan kielen r-kirjain, luet. *roo*). Jos X-muuttujien ja Y-muuttujien arvojen järjestysnumerot ovat jokaisessa havaintoparissa samat, $\Sigma d_i = 0$ ja siten $\rho = 1$. Tällöin X-muuttujan ja Y-muuttujan arvojen järjestysnumeroiden välillä vallitsee täydellinen riippuvuus.

Voidaan osoittaa, että Spearmanin ρ:lla on kaikki korrelaatiokertoimelta vaadittavat ominaisuudet:

1. $-1 \le \rho \le +1$
2. Jos suuret X-muuttujan arvojen järjestysnumerot liittyvät suuriin Y-muuttujan arvojen järjestysnumeroihin ja pienet X-muuttujan arvojen järjestysnumerot liittyvät pieniin Y-muuttujan arvojen järjestysnumeroihin, ρ saa positiivisia arvoja. ρ:n arvo on sitä lähempänä arvoa +1, mitä voimakkaampi on mainittu yhteys.
3. Jos suuret X-muuttujan arvojen järjestysnumerot liittyvät pieniin Y-muuttujan arvojen järjestysnumeroihin ja pienet X-muuttujan arvojen järjestysnumerot liittyvät suuriin Y-muuttujan arvojen järjestysnumeroihin, ρ saa negatiivisia arvoja. ρ:n arvot ovat sitä lähempänä arvoa −1, mitä voimakkaampi on mainittu yhteys.
4. Jos X-muuttujan arvojen järjestysnumerot liittyvät satunnaisesti Y-muuttujan arvojen järjestysnumeroihin, ρ:n arvo on lähellä arvoa 0. Jos $\rho = 0$, sanotaan, että X- ja Y-muuttuja ovat *korreloimattomia*.

Spearmanin korrelaatiokertoimen käyttäytymistä voidaan havainnollistaa seuraavalla esimerkillä:

ESIMERKKI 2.

Seuraavassa on esitetty kaikki mahdolliset kahden muuttujan arvojen järjestykset sekä vastaavat Spearmanin korrelaatiokertoimen arvot, kun havaintoja on 2 ja 3:

2 havaintoa:

1	1	$\rho = +1$
2	2	

1	2	$\rho = -1$
2	1	

3 havaintoa:

1	1	$\rho = +1$
2	2	
3	3	

1	3	$\rho = -1$
2	2	
3	1	

1	2	1	1	$\rho = +0.5$
2	1	2	3	
3	3	3	2	

1	3	1	2	$\rho = -0.5$
2	1	2	3	
3	2	3	1	

Esimerkiksi viimeisessä 3:n havainnon tapauksessa saadaan

$$\rho = 1 - \frac{6\left[(1-3)^2 + (2-1)^2 + (3-2)^2\right]}{27-3} = -\frac{1}{2}.$$

On hyvä huomata, että toisen muuttujan havaintoarvojen järjestyksen vaihtaminen päivastaiseksi vaihtaa korrelaatiokertoimen merkin. ●

Oletetaan, että haluamme testata nollahypoteesia

$$H_0: \quad \rho = 0$$

eli, että X- ja Y-muuttujat ovat korreloimattomia. Testi tälle nollahypoteesille voidaan perustaa siihen, että testisuure

$$Z = \frac{\rho\sqrt{n-2}}{\sqrt{1-\rho^2}}$$

jakautuu suurissa otoksissa approksimatiivisesti kuten standardoitu normaalijakauma N(0,1), *jos nollahypoteesi pätee*. Siten testisuureen kyllin voimakkaasti nollasta poikkeavat arvot johtavat nollahypoteesin hylkäämiseen.

Normaaliapproksimaatio on melko hyvä, jos $n > 10$ ja käytännössä täysin riittävä, jos $n > 30$. Pienille havaintojen lukumäärille on konstruoitu erillisiä taulukoita.

ESIMERKKI 1 (JATKOA).

Spearmanin järjestyskorrelaatiokerroin saa esimerkin 1 aineistossa arvon

$$\rho = 1 - \frac{6\left[(6-3)^2 + \cdots + (4-1)^2\right]}{1000-10} \approx 0.62.$$

Vastaava Z-testisuureen arvo on

$$Z = \frac{0.62\sqrt{10-2}}{\sqrt{1-0.62^2}} = 2.40$$

Koska testisuureen arvoa vastaava P-arvo on normaalijakauman taulukoiden mukaan

$$P(Z \geq 2.40) = 0.0082,$$

voidaan nollahypoteesi hylätä esimerkiksi 1%:n merkitsevyystasolla.

Johtopäätöksenä on se, että viinintuntijoiden maut eivät ole riippumattomia, vaan he käyttävät samanlaista arvosteluasteikkoa viineille. ●

KENDALLIN JÄRJESTYSKORRELAATIOKERROIN

Oletetaan, että tutkimuksen kohteena on kahden muuttujan X ja Y riippuvuus ja olkoot X ja Y vähintään järjestysasteikollisia muuttujia. Oletetaan, että muuttujista X ja Y on kerätty n toisistaan riippumatonta havaintoparia. Annetaan kummankin muuttujan arvoille järjestysnumerot niiden suuruusjärjestyksen mukaan. Olkoot

X_i = X-muuttujan arvon järjestysnumero parissa i,

Y_i = Y-muuttujan arvon järjestysnumero parissa i.

Siten havaintopari (X_i, Y_i) liittyy samaan *tapaukseen* i.

Kendallin järjestyskorrelaatiokerroin τ voidaan määritellä kaavalla

$$\tau = \frac{S}{\frac{1}{2}n(n-1)},$$

jossa

S = epäjärjestystä toisen muuttujan järjestysluvuissa kuvaava tunnusluku.

Epäjärjestyksen mitalla S kuvataan seuraavaa seikkaa: Oletetaan, että havaintoparit (X_i, Y_i) on järjestetty (esimerkiksi) muuttujan X arvojen mukaan järjestykseen. Otetaan mitkä tahansa kaksi tapausta tarkastelun kohteeksi. Jos Y-muuttujan arvojen suuruusjärjestys on tässä parissa *sama* kuin X-muuttujan arvojen suuruusjärjestys, sanotaan, että Y-muuttujan arvot ovat *järjestyksessä* (*X-muuttujan arvojen suhteen*). Päivastaisessa tapauksessa sanotaan, että Y-muuttujan arvot ovat *epäjärjestyksessä* (*X-muuttujan arvojen suhteen*). Tarkastellaan samalla tavalla kaikkia tapauspareja. Lienee selvää, että mitä useammissa tapauspareissa Y-muuttujan arvot ovat järjestyksessä X-muuttujan arvojen suhteen, sitä voimakkaammin X-muuttujan ja Y-muuttujan arvot korreloivat keskenään.

Tunnusluku S määrätään seuraavalla tavalla: Järjestetään siis havaintoparit (esimerkiksi) X-muuttujan arvojen mukaan järjestykseen. Liitetään *jokaiseen* Y-muuttujan arvojen järjestysluvun *jälkeen* tuleviin järjestyslukuihin pisteitä "oikeista" ja "vääristä" järjestyksistä seuraavalla tavalla:

oikea järjestys = +1,

väärä järjestys = −1.

Käydään tällä tavalla läpi kaikki Y-muuttujan arvojen järjestysluvut ja määritellään epäjärjestystä kuvaava tunnusluku S seuraavalla tavalla:

S = pisteiden summa.

Huomaa, että τ:n lausekkeen nimittäjä saa selityksensä siitä, että summassa S on

$$(n-1) + (n-2) + \ldots + 2 + 1 = \frac{1}{2}n(n-1)$$

termiä ja, että summan S maksimiarvo on $n(n-1)/2$ ja minimiarvo on $-n(n-1)/2$.

Kendallin järjestyskorrelaatiokerrointa kutsutaan usein Kendallin τ:ksi (τ on kreikan kielen t-kirjain, luet. *tau*). Myös Kendallin τ toteuttaa korrelaatiokertoimelle Spearmanin ρ:ta käsitelleessä kappaleessa asetetut ehdot ja on siis järkevä mitta muuttujien riippuvuudelle.

Olkoon nollahypoteesina muuttujien X ja Y korreloimattomuus, ts.

$$H_0: \quad \tau = 0.$$

Testi H_0:lle voidaan perustaa siihen, että satunnaismuuttuja

$$Z = \frac{\tau}{\sqrt{\dfrac{2(2n+5)}{9n(n-1)}}}$$

jakautuu suurissa otoksissa approksimatiivisesti kuten standardoitu normaalijakauma $N(0,1)$, *jos nollahypoteesi* H_0 *pätee*. Kyllin voimakkaasti nollasta poikkeavat testisuureen Z arvot johtavat nollahypoteesin hylkäämiseen.

Normaaliapproksimaatio on melko hyvä, jos $n > 10$ ja käytännössä täysin riittävä, jos $n > 30$. Pienille havaintojen lukumäärille on konstruoitu erillisiä taulukoita.

ESIMERKKI 1 (JATKOA).

Järjestetään viinit viinintuntija A maun mukaiseen järjestykseen. Tarkastelukulmaksi otetaan siis A:n arvosteluasteikko viineille: A:n antamaa järjestystä pidetään viinien "oikeana" järjestyksenä. Viinit voidaan tietysti yhtä hyvin järjestää B:n maun mukaan, jolloin B:n antama järjestystä pidetään "oikeana".

Viini	Maistaja A	Maistaja B
7	1	2
2	2	4
9	3	8
10	4	1
8	5	5
1	6	3
6	7	9
3	8	7
4	9	6
5	10	10

Taulukosta nähdään, että viinit 7 ja 2 ovat A:lla sijaluvuilla 1 ja 2 ja B:llä sijaluvuilla 2 ja 4, ts. niiden keskinäinen järjestys on A:lla ja B:llä sama, vaikka ne sijoittuvat A:n ja B:n arvoasteikoilla eri paikkaan. Sen sijaan viinit 9 ja 10 ovat A:lla sijaluvuilla 3 ja 4, mutta B:llä sijaluvuilla 8 ja 1, ts. sen lisäksi, että ne sijoittuvat arvoasteikoilla erilailla, niiden keskinäinen järjestys on A:lla ja B:llä vastakkainen. Lienee selvää, että mitä useammin B on asettanut kaksi viiniä samaan järjestykseen kuin A, sitä enemmän B:n maku muistuttaa A:n makua.

Epäjärjestyksen mitan S määräämiseksi tarvittavien "epäjärjestyspisteiden" laskeminen voidaan järjestää seuraavan taulukon muotoon:

Viini	A	B	1	2	3	4	5	6	7	8	9
7	1	2	*	*	*	*	*	*	*	*	*
2	2	4	+1	*	*	*	*	*	*	*	*
9	3	8	+1	+1	*	*	*	*	*	*	*
10	4	1	–1	–1	–1	*	*	*	*	*	*
8	5	5	+1	+1	–1	+1	*	*	*	*	*
1	6	3	+1	–1	–1	+1	–1	*	*	*	*
6	7	9	+1	+1	+1	+1	+1	+1	*	*	*
3	8	7	+1	+1	–1	+1	+1	+1	–1	*	*
4	9	6	+1	+1	–1	+1	+1	+1	–1	–1	*
5	10	10	+1	+1	+1	+1	+1	+1	+1	+1	+1
		Yht.	+7	+4	–3	+6	+3	+4	–1	0	+1

Esimerkiksi 3. sarakkeen luvut on saatu seuraavalla tavalla:

Verrataan A:n asteikossa 3. viinin (viini 9) järjestyslukua B:n asteikossa sitä B-sarakkeessa seuraaviin järjestyslukuihin. Viini 9 on B:n asteikossa sijalla 8. Sitä seuraavat B-sarakkeen järjestysluvut ovat 1, 5, 3, 9, 7, 6 ja 10. Niistä 1, 5, 3, 7 ja 6 ovat "väärässä" järjestyksessä järjestyslukuun 8 nähden, kun taas 9 ja 10 ovat "oikeassa" järjestyksessä. Siten "oikeita" järjestyksiä on 5 ja "vääriä" järjestyksiä on 2 ja vastaava sarakesumma on

$$-5+2=-3.$$

Epäjärjestysmitta S saadaan taulukosta sarakesummien summana:

$$S=7+4-3+6+3+4-1+0+1=21.$$

Kendallin τ saa siten arvon

$$\tau=\frac{21}{\frac{1}{2}\cdot 10\cdot 9}\approx 0.47.$$

Siten testiruure Z saa arvon

$$Z = \frac{0.47}{\sqrt{\frac{2\cdot(2\cdot10+5)}{9\cdot10\cdot(10-1)}}} \approx 1.88.$$

Koska vastaava P-arvo on

$$P(Z \geq 1.88) = 0.0301,$$

voidaan nollahypoteesi hylätä merkitsevyystasolla 0.05. ●

Huomaa, että Kendallin τ:n arvo ei ole yleensä sama kuin Spearmanin ρ:n arvo. Myöskään korreloimattomuushypoteesin H_0 testaamiseen liittyvät P-arvot eivät ole yleensä samoja.

TASAPELIT JA NIIDEN MERKITYS

Esimerkissä 1 ajateltiin, että viinintunitjat A ja B pystyivät järjestämään viinit siten, että viineille ei tullut samoja sijalukuja. Näin ei tietystikään tarvitse aina olla asian laita. Miten toimitaan sellaisessa tilanteessa, jossa kahteen tai useampaan tapaukseen liittyy sama sijaluku? Tarkastellaan tätä seuraavan esimerkin avulla.

ESIMERKKI 3.

Oletetaan, että viinintuntijat A ja B ovat asettaneet 10 viiniä seuraavalla tavalla paremmuusjärjestykseen:

Viini	Maistaja A	Maistaja B
7	1	2
2	2	4
9	2	8
10	4	1
8	5	4
1	6	3
6	7	9
3	8	7
4	9	4
5	10	10

Viinintuntijat ovat sijalukuja jakaessaan käyttäneet sitä urheilukilpailuissa yleisesti sovellettua menetelmää, jossa kahden tai useamman urheilijan saadessa saman tuloksen, he saavat saman sijaluvun ja vastaava määrä sijalukuja heidän jälkeensä jätetään tyhjiksi. Esimerkissä A on sijoittanut viinit 2 ja 9 sijaluvulle 2, jolloin sijaluku 3 on jäänyt tyhjäksi. B on taas asettanut viinit 2, 4 ja 8 sijaluvulle 4, jolloin sijaluvut 5 ja 6 ovat jääneet tyhjiksi.

Järjestysasteikollisiin muuttujiin liittyvissä testeissä on kuitenkin parempi antaa yhteiseksi sijaluvuksi sijalukujen keskiarvo. Siten A:n pitää antaa viineille 2 ja 9 sijaluku

$$\frac{2+3}{2} = 2.5$$

ja B:n viineille 2, 4 ja 8 sijaluku

$$\frac{4+5+6}{3} = 5.$$

Tätä menettelyä käyttäen saadaan seuraava taulukko:

Viini	Maistaja A	Maistaja B
7	1	2
2	2.5	5
9	2.5	8
10	4	1
8	5	5
1	6	3
6	7	9
3	8	7
4	9	5
5	10	10

Laskutoimitukset Kendallin τ:n määräämiseksi pitää perustaa tähän taulukkoon.

●

Esimerkissä 3 esitettyä (tai jotakin vastaavaa) menetelmää on syytä käyttää aina, kun järjestysasteikollisten muuttujien arvoissa esiintyy tasapelejä. Tilastolliset ohjelmistot osaavat yleensä *ratkaista tasapelit* itsenäisesti.

4.4 TESTEJÄ VÄLI- JA SUHDEASTEIKOLLISILLE MUUTTUJILLE

4.4.1 JOHDANTO

Väli- ja suhdeasteikollisten satunnaismuuttujien jakaumien parametreja koskevia hypoteeseja testataan tavallisesti *parametrisilla* testeillä. Esimerkin tällaisista testeistä antaa tavallinen t-testi perusjoukon odotusarvolle. Tässä kappaleessa tarkastellaan testejä seuraaville testausasetelmille:

- Onko perusjoukon odotusarvo annettu vakio?
- Ovatko kahden perusjoukon odotusarvot samoja?
- Onko perusjoukon varianssi annettu vakio?
- Ovatko kahden perusjoukon varianssit samoja?
- Ovatko kaksi satunnaismuuttujaa korreloimattomia?

Kaikissa tässä käsiteltävissä testeissä oletetaan, että tutkittavat muuttujat ovat *väli-* tai *suhdeasteikollisia*. Lisäksi muuttujien arvojen jakaumasta perusjoukossa tehdään oletukset, joihin sisältyy tieto muuttujien *odotusarvoista* ja *variansseista*. Odotusarvoja koskevat testit ovat muodoltaan t-testejä. Sen sijaan variansseja koskevat testit ovat muodoltaan χ^2- tai F-testejä. Myös korrelaatioon liittyvä testi on muodoltaan t-testi.[1]

Parametristen testien yhteydessä tutkimuksen kohteena olevien muuttujien oletetaan usein noudattavan perusjoukossa jotakin tiettyä jakaumaa. Tällaiset oletukset ovat osa testausasetelmassa muotoiltua *yleistä hypoteesia*. Tyypillinen tällainen oletus on *normaalisuusoletus*. Tästä oletuksesta seuraa se, että tavanomainen t-testisuure noudattaa Studentin t-jakaumaa. Vaikka perusjoukko ei olisikaan jakautunut normaalijakauman mukaan, t-testisuure on suurissa otoksissa yleensä kuitenkin approksimatiivisesti normaalinen. Tämä tulos perustuu keskeiseen raja-arvolauseeseen.

On hyvä tietää, että t-testin tapaiset parametriset testit toimivat yleensä sitä huonommin mitä enemmän perusjoukon jakauma poikkeaa yleisessä hypoteesissa oletetusta jakaumasta. Tämä johtuu siitä, että parametriset testit ovat tällaisissa tilanteissa tavallisesti *voimattomia*, ts. *hyväksymisvirhe* eli 2. *lajin virhe* on tällaisissa tilanteissa suuri. Tällöin on usein edullisempaa käyttää sellaisia *ei-parametrisia testejä*, joita esiteltiin laatuero- ja järjestysasteikollisille muuttujille, vaikka tutkittavat muuttujat olisivatkin väli- tai suhdeasteikollisia. On hyvä muistaa, että yleiseen hypoteesiin sisältyviä jakaumaoletusta voidaan testata esimerkiksi χ^2-yhteensopivuustestillä.

[1] t- ja χ^2-jakaumia on käsitelty edellä; F-jakauma esitellään alla.

4.4.2 TESTI PERUSJOUKON ODOTUSARVOLLE.

Olkoon

$$X_1, X_2, \ldots, X_n$$

riippumaton satunnaisotos jakaumasta, jonka odotusarvo on μ ja varianssi on σ^2, ts.

$$E(X_i) = \mu,$$

$$D^2(X_i) = \sigma^2$$

kaikille $i = 1,2,\ldots,n$. Olkoot

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

ja

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

tavanomaiset harhattomat estimaattorit parametreille μ ja σ^2.

Olkoon nollahypoteesina

$$H_0: \quad \mu = \mu_0.$$

Määritellään testisuure

$$t = \frac{\overline{X} - \mu_0}{s/\sqrt{n}}.$$

Jos H_0 pätee,

$$E(\overline{X}) = \mu_0.$$

Tällöin t-testisuureen erotuksen $\overline{X} - \mu_0$ itseisarvo $|\overline{X} - \mu_0|$ saa keskimäärin pieniä arvoja. Jos taas H_0 ei päde, $|\overline{X} - \mu_0|$ saa keskimäärin suuria arvoja. Sopivan *mitta-yksikön* sille, koska $|\overline{X} - \mu_0|$ on pieni tai suuri tarjoaa otoskeskiarvon $\overline{X}$ hajonnan

$$D(\overline{X}) = \frac{\sigma}{\sqrt{n}}$$

estimaattori

$$\hat{D}(\overline{X}) = \frac{s}{\sqrt{n}}.$$

Testi nollahypoteesille H_0 perustuu siten testisuureeseen t. Itseisarvoltaan suuret testisuureen arvot johtavat nollahypoteesin hylkäämiseen. t-testisuureen jakauma riippuu perusjoukosta tehtävistä oletuksista:

1. Jos perusjoukko on normaalisti jakautunut, testiruure t jakautuu kuten Studentin t-jakauma $(n-1)$:llä vapausasteella, *jos nollahypoteesi pätee*. Huomaa, että perusjoukon normaalisuutta voidaan testata χ^2-testillä.

2. Jos perusjoukko ei ole normaalisti jakautunut, voidaan testiä tehtäessä yleensä nojata siihen, että testisuure t jakautuu suurissa otoksissa approksimatiivisesti kuten standardoitu normaalijakauma N(0,1), *jos nollahypoteesi pätee*.

ESIMERKKI 1.

Oletetaan, että kuulalaakereiden kuulia valmistava kone on säädetty niin, että satunnaisesti valitun kuulan odotettavissa oleva paino

$$\mu = 5.00 \text{ g.}$$

Laakereiden painon valvomiseksi poimitaan 100 kuulan yksinkertainen satunnaisotos ja punnitaan kuulat. Otoksesta saadaan seuraavat tunnusluvut:

$$\bar{X} = 5.09 \text{ g},$$

$$s^2 = (0.30 \text{ g})^2.$$

Onko otoksesta saatu tieto sopusoinussa kuulien odotusarvosta tehdyn oletuksen kanssa, vai onko kuulia valmistava kone alkanut tehdä liian painavia kuulia?

Nollahypoteesina on

$$H_0: \quad \mu = 5.00 \text{ g.}$$

t-testisuure tälle nollahypoteesille on

$$t = \frac{5.09 - 5.00}{0.3 / \sqrt{100}} = 3.$$

Normaalijakauman taulukoista (tai 99.7-säännöstä) nähdään, että vastaava approksimatiivinen P-arvo on 0.003 ja nollahypoteesi voidaan hylätä esimerkiksi 1%:n merkitsevyystasolla: Koneen asetukset ovat todennäköisesti siirtyneet pois paikaltaan ja koneen toiminta on syytä tarkistaa.

Jos kuulien painon voidaan olettaa olevan normaalijakautuneen, voidaan testisuureen arvoa vastaava P-arvo katsoa t-jakauman taulukoista. Koska vapausasteiden luku on kuitenkin niinkin suuri kuin 99, ei tulos eroa olennaisesti normaalijakaumaan perustuvasta tuloksesta. ●

4.4.3 KAHDEN PERUSJOUKON ODOTUSARVOJEN VERTAAMINEN

Tarkastelu jaetaan kahteen osaan:

1. Otokset ovat riippumattomia.
2. Otokset riippuvat toisistaan.

RIIPPUMATTOMAT OTOKSET

Oletetaan, että tarkastelun kohteena on satunnaismuuttujan X odotusarvojen vertaaminen kahdessa perusjoukossa (tai kahdessa saman perusjoukon ryhmässä). Oletetaan, että perusjoukoista on tehty kaksi *riippumatonta* satunnaisotosta ja olkoon

X_{1i} = muuttujan X arvo 1. perusjoukossa havaintoyksikölle $i = 1,2,\ldots,n_1$,

X_{2j} = muuttujan X arvo 2. perusjoukossa havaintoyksikölle $j = 1,2,\ldots,n_2$,

Oletetaan lisäksi, että

$$\mathrm{E}(X_{1i}) = \mu_1, \quad \mathrm{D}^2(X_{1i}) = \sigma_1^2,$$

$$\mathrm{E}(X_{2j}) = \mu_2, \quad \mathrm{D}^2(X_{1j}) = \sigma_2^2.$$

Olkoon nollahypoteesi on muotoa

$$\mathrm{H}_0: \quad \mu_1 = \mu_2 = \mu,$$

jossa odotusarvojen yhteistä arvoa on merkitty μ:llä.

Testisuure perustetaan otoskeskiarvojen erotukseen

$$\overline{X}_1 - \overline{X}_2$$

jossa

$$\overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i},$$

$$\overline{X}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} X_{2j}.$$

Jos nollahypoteesi H_0 pätee,

$$\mathrm{E}(\overline{X}_1) = \mathrm{E}(\overline{X}_2) = \mu.$$

Tällöin erotuksen $\overline{X}_1 - \overline{X}_2$ itseisarvo $|\overline{X}_1 - \overline{X}_2|$ saa keskimäärin pieniä arvoja. Jos taas H_0 ei päde, itseisarvo $|\overline{X}_1 - \overline{X}_2|$ saa keskimäärin suuria arvoja. Sopivan mittayksikön sille, koska $|\overline{X}_1 - \overline{X}_2|$ on pieni tai suuri tarjoaa otoskeskiarvojen erotuksen $\overline{X}_1 - \overline{X}_2$ hajonnan

$$D(\bar{X}_1 - \bar{X}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

estimaattori

$$\hat{D}(\bar{X}_1 - \bar{X}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

jossa s_1^2 ja s_2^2 ovat varianssien σ_1^2 ja σ_2^2 tavanomaiset harhattomat estimaattorit. Huomaa, että erotuksen $\bar{X}_1 - \bar{X}_2$ hajontaa koskeva tulos perustuu siihen, että otokset 1 ja 2 oletettiin riippumattomiksi.

Määritellään testisuure

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

Voidaan osoittaa, että testisuure Z noudattaa suurilla otoksilla approksimatiivisesti standardoitua normaalijakaumaa $N(0,1)$, *jos nollahypoteesi pätee.*

Tätä tilannetta tarkasteltiin jo tämän luvun johdannossa. Oletetaan nyt, että perusjoukoista voidaan olettaa enemmän: Oletetaan, että perusjoukot ovat kumpikin normaalijakautuneita ja oletetaan lisäksi, että

$$\sigma_1^2 = \sigma_2^2 = \sigma^2,$$

jossa varianssien yhteistä arvoa on merkitty σ^2:lla.

Huomaa, että tässä tapauksessa nollahypoteesista H_0 seuraa, että perusjoukot noudattavat täsmälleen *samaa* normaalijakaumaa. Tällöin perusjoukot voidaan yhdistää yhdeksi perusjoukoksi satunnaismuuttujaa X tarkasteltaessa. Huomaa, että oletusta varianssien samuudesta voidaan testata. Tähän tarkoitukseen sopivaa testiä käsitellään myöhemmin tässä luvussa.

Testisuureeksi nollahypoteesille H_0 voidaan em. lisäoletusten pätiessä ottaa

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

jossa s on yhdistettyjen perusjoukkojen standardipoikkeaman *harhaton* estimaattori, *jos nollahypoteesi* H_0 *pätee.* Estimaattori s saadaan kaavasta

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2},$$

Jos H_0 ja em. lisäoletukset pätevät, testisuure t on jakautunut kuten Studentin t-jakauma (n_1+n_2-2):lla vapausasteella. Kahden perusjoukon odotusarvojen vertailutesti esitetään usein tässä muodossa myös sellaisissa tilanteissa, joissa oletukset varianssien samuudesta ja perusjoukkojen normaalisuudesra eivät ole perusteltuja.

ESIMERKKI 1.

Tarkastellaan uudelleen tämän luvun johdannossa esitettyä esimerkkiä psykologisesta testistä, jolla tutkittiin suorituskykyä meluisassa ja meluttomassa ympäristössä.

Oletetaan, että laskutoimituksissa tehtyjen virheiden varianssi on sama kummassakin ympäristössä. Palaamme jatkossa siihen onko tämä hypoteesi oikeutettu.

Yhdistettyjen perusjoukkojen varianssin harhaton estimaattori on nollahypoteesin pätiessä

$$s^2 = \frac{15 \cdot 2.28^2 + 15 \cdot 2.45^2}{16+16-2} \approx 5.40.$$

Siten

$$s = 2.32.$$

t-testisuure saa siten arvon

$$t = \frac{14.4 - 16.5}{2.32\sqrt{\frac{1}{16} + \frac{1}{16}}} \approx -2.56.$$

Testisuure on jakautunut kuten Studentin t-jakauma $(16+16-2) = 30$:lla vapausasteella, jos nollahypoteesi pätee. Siten testisuureetta vastaava P-arvo on 0.008. Nollahypoteesi voidaan siis hylätä esimerkiksi 1%:n merkitsevyystasolla. ●

RIIPPUVAT OTOKSET

Oletetaan, että haluamme verrata satunnaismuuttujan X odotusarvoa kahdessa perusjoukossa kuten edelläkin. Oletetaan, että asetelma on muuten samanlainen kuin edellä, mutta että perusjoukoista poimitut otokset riippuvat toisistaan sillä erityisellä tavalla, joka liittyy *sovitettuihin pareihin* koeasetelmiessa. Tarkastellaan seuraavaa esimerkkiä:

ESIMERKKI 1.

Oletetaan, että tarkoituksena on verrata kahta menetelmää A ja B, joita käytetään perunoiden tärkkelyspitoisuuden mittaamiseen. A saattaa olla hidas sekä kallis ja B nopea sekä halpa. Tutkijoita kiinnostaa voisiko menetelmän A korvata menetelmällä B. Tämä on mahdollista, jos A ja B antavat keskimäärin saman tuloksen. Miten menetelmiä pitää verrata toisiinsa?

Ensinnäkin on syytä huomata, että menetelmien vertailua ei ole järkevää toteuttaa seuraavalla tavalla: Valitaan toisistaan riippumattomasti kaksi yksinkertaista satunnaisotosta perunoita ja sovelletaan menetelmää A toisen otokseen ja menetelmää B toiseen otokseen. Menetelmien vertailua ei voida nimittäin perustaa otoksista määrättyihin tärkkelyspitoisuuksien otoskeskiarvoihin. Tämä johtuu siitä, että tärkkelyspitoisuuksien vaihtelu perunasta toiseen tuhoaa mahdollisuuden päätellä, että mahdolliset erot otoskeskiarvoissa johtuvat mittausmenetelmistä.

Menetelmien vertailu on syytä toteuttaa seuraavalla tavalla: Valitaan satunnaisesti joukko perunoita, jotka halkaistaan ja sovelletaan menetelmää A *toiseen puoliskoon* ja B menetelmää *toiseen puoliskoon*. Näin saatujen pitoisuuksien keskiarvoja ei voida kuitenkaan vertailla edellä kuvatulla kahden perusjoukon odotusarvojen vertailuun tarkoitetulla t-testillä, koska samaan perunaan liittyvät mittaukset korreloivat voimakkaasti keskenään ja edellä käsitellyssä t-testissä oletettiin, että otokset olivat riippumattomia.

Miten testi on tehtävä tällaisessa vertailutilanteessa, jossa samaan tapaukseen liittyvät muuttujan arvot korreloivat voimakkaasti toistensa kanssa?

Testi voidaan perustaa seuraavaan menettelyyn: Muodostetaan *samaan* perunaan liittyvien tärkkelyspitoisuusmittausten erotukset ja tutkitaan tavallisella yhden otoksen t-testillä onko erotusten odotusarvo 0. ●

Esimerkin 1 tilanne on tyypillinen esimerkki *sovitettujen parien asetelmasta* koesuunnittelussa. Tällaisessa tilanteessa havainnot muodostuvat *saman* muuttujan arvojen pareista (X_{1i}, X_{2i}), $i = 1,2,\ldots,n$, joissa muuttujan arvot kytkeytyvät toisiinsa. Määritellään havaintoarvojen erotukset kaavalla

$$D_i = X_{1i} - X_{2i}.$$

Asetetaan nollahypoteesiksi se, että kaikkien erotusten odotusarvo häviää:

$$\mathrm{H_0:}\quad \mathrm{E}(D_i) = 0, \qquad i = 1,2,\ldots,n.$$

Määritellään testisuure

$$t = \frac{\overline{D}}{s_D / \sqrt{n}},$$

jossa

$$\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$$

on erotusten D_i aritmeettinen keskiarvo ja

$$s_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \overline{D})^2$$

on erotusten D_i otosvarianssi.

Jos nollahypoteesi H_0 *pätee,*

$$E(\overline{D}) = 0.$$

Tällöin t-testisuureen osoittajan $\overline{D}$ itseisarvo $|\overline{D}|$ saa keskimäärin pieniä arvoja. Jos taas H_0 ei päde, itseisarvo $|\overline{D}|$ saa keskimäärin suuria arvoja. Sopivan mittayksikön sille, koska $|\overline{D}|$ on suuri tai pieni, tarjoaa erotuksen $\overline{D}$ hajonnan estimaattori s_D. Nollasta voimakkaasti poikkeavat testisuureen arvot johtavat nollahypoteesin hylkäämiseen.

Testisuureen t jakauma riippuu perusjoukon jakaumasta:

1. Jos perusjoukot ovat normaalisti jakautuneita, testisuure t on jakautunut kuten Studentin t jakauma $(n-1)$:llä vapausasteella, *jos nollahypoteesin pätee.*
2. Jos perusjoukot eivät ole normaalisti jakautuneita, voidaan yleensä nojata siihen, että testisuure t noudattaa suurilla otoksilla approksimatiivisesti standardoitua normaalijakaumaa $N(0,1)$, *jos nollahypoteesi pätee.*

ESIMERKKI 1 (JATKOA).

Oletetaan, että mittaukset tehtiin 16 perunalle ja

$$\overline{D} = 0.075, \qquad s_D = 0.169.$$

Tällöin t-testisuure saa arvon

$$t = \frac{0.075}{0.1693 / \sqrt{16}} \approx 1.77.$$

Oletetaan, että mittausten erotukset ovat normaalijakautuneita, jolloin voidaan nojata Studentin t-jakauman taulukoihin. 5%:n merkitsevyystasoa vastaava kriittinen arvo on 1.71, ts.

$$P(t \geq 1.71) = 0.05.$$

Siten testisuureen arvoa vastaava P-arvo on hieman pienempi kuin 0.05 ja nollahypoteesi voidaan hylätä 5%:n merkitsevyystasolla. Tällaisessa tilanteessa menetelmää A ei pidä korvata menetelmällä B, jos on erityisen tärkeää, että tärkkelyspitoisuuden mittaus on tehty oikein. ●

4.4.4 TESTI PERUSJOUKON VARIANSSILLE

Edellä on tarkasteltu lukuisia testejä, jotka liittyvät tutkimuksen kohteena olevan satunnaismuuttujan keskimääräisten arvojen *paikkaan* perusjoukossa. Tyypillinen esimerkki sellaisesta testistä on satunnaismuuttujan jakauman odotusarvolle eli painopisteelle tarkoittu t-testi. t-testillä voidaan tutkia onko satunnaismuuttujan odotusarvo muuttunut.

Satunnaismuuttujan arvojen jakaumassa perusjoukossa voi tietysti tapahtua myös muunlaisia muutoksia. Useissa tutkimustilanteissa on tärkeätä saada selville onko satunnaismuuttujan jakauman *keskittyneisyydessä* (tai hajaantuneisuudessa) jakauman paikkaa kuvaavan parametrin ympärille tapahtunut jokin muutos. Seuraavassa tarkastellaan normaalijakautuneen perusjoukon *varianssiin* liittyvää testiä.

Oletetaan, että

$$X_1, X_2, \ldots, X_n$$

on riippumaton satunnaisotos normaalijakaumasta, jonka odotusarvo on μ ja varianssi on σ^2. Oletetaan, että haluamme testata nollahypoteesia

$$H_0: \quad \sigma^2 = \sigma_0^2.$$

Nollahypoteesin mukaan perusjoukon varianssilla on arvo σ_0^2.

Olkoon

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

varianssin σ^2 harhaton estimaattori eli tavanomainen *otosvarianssi*, jossa

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

on odotusarvon μ harhaton estimaattori eli tavanomainen otoskeskiarvo.

Muodostetaan testisuure

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}.$$

Jos nollahypoteesi H_0 *pätee,*

$$E(s^2) = \sigma_0^2.$$

Siten testisuureen χ^2 odotusarvo on

$$E(\chi^2) = n - 1,$$

jos nollahypoteesi pätee. Siten testisuure χ^2 saa nollahypoteesin pätiessä keskimäärin arvoja, jotka ovat lähellä arvoa $n-1$. Jos taas H_0 ei päde, testisuure χ^2 saa keskimäärin kaukana arvosta $n-1$ olevia arvoja. Voimakkaasti testisuureen χ^2 odotettavissa olevasta arvosta *poikkeavat* testisuureen arvot johtavat nollahypoteesin H_0 hylkäämiseen.

Voidaan osoittaa, että testisuure χ^2 noudattaa χ^2-jakaumaa $(n-1)$:llä vapausasteella, *jos nollahypoteesi* H_0 *pätee.*

ESIMERKKI 1.

Tarkastellaan jälleen kuulalaakereiden kuulia valmistavan koneen toiminnan valvomista.

Oletetaan, että laakereissa voidaan käyttää sellaisia kuulia, joiden halkaisija on

$$5 \text{ mm} \pm 0.03 \text{ mm}.$$

Siten sellaisia kuulia, joiden halkaisija on välin [4.97, 5.03] ulkopuolella ei voi käyttää laakereissa.

Jotta mahdollisimman suuri osa valmistetuista kuulista olisi käyttökelpoisia, kuulia valmistava kone on säädetty siten, että valmistettujen kuulien halkaisijan odotusarvo (kuulien halkaisijoiden tavoitearvo) on

$$\mu = 5.00 \text{ mm}$$

ja varianssi (kuulien halkaisijoiden vaihtelun tavoitearvo) on

$$\sigma^2 = (0.01 \text{ mm})^2.$$

Lisäksi tiedetään, että kuulien halkaisijat jakautuvat tavallisesti normaalijakauman mukaan.

Tehdyistä oletuksista voidaan päätellä, että kuulia valmistavan koneen toimiessa oikein vain 0.3% kuulista on käyttökelvottomia. Tämä johtuu siitä, että em. oletusten ja ns. 99.7-säännön perusteella

$$P(5 - 0.03 \leq X \leq 5 + 0.03) = 0.997,$$

jossa X noudattaa normaalijakaumaa $N(5 \text{ mm}, (0.01 \text{ mm})^2)$.

Kuulien laadun kannalta on tärkeää, että sekä valmistettujen kuulien halkaisijoiden odotusarvolla että varianssilla on niille säädetyt arvot, jotka takaavat sen, että mahdollisimman suurta osaa kuulista voidaan käyttää laakereissa.

Jos kuulien valmistava kone alkaa tuottaa kuulia, joiden halkaisija on liian suuri (tai pieni) kasvaa käyttökelvottomien kuulien osuus. Tätä voidaan testata odotusarvoon kohdistuvalla t-testillä. Vähintään yhtä epätoivottavaa kuulien laadun kannalta on sellainen tilanne, jossa kuulien halkaisijoiden varianssi alkaa kasvaa. Tällöin kuulien halkaisijat alkavat heittelehtiä.

Oletetaan, että 100:n kuulan satunnaisotoksessa havaitaan, että otoshajonta on

$$s = 0.012 \text{ mm}.$$

Onko tämä havainto sopusoinnussa nollahypoteesin

$$H_0: \quad \sigma^2 = (0.01 \text{ mm})^2.$$

kanssa? χ^2-testisuure saa arvon

$$\chi^2 = \frac{(100-1)\cdot 0.012^2}{0.01^2} \approx 142.6.$$

Testisuure χ^2 noudattaa χ^2-jakaumaa vapausastein 99, jos nollahypoteesi pätee. Testisuureen arvoa vastaava P-arvo on

$$P(\chi^2 \geq 142.6) = 0.003,$$

Testisuureen arvo 142.6 tai tätä suurempi arvo saadaan siis vain noin 3 kertaa 1000:sta, jos nollahypoteesi pätee.

Edellä todettiin, että jos kuulien halkaisija on jakautunut kuten normaalijakauma ja hypoteesit, jotka koskevat halkaisijan odotusarvoa ja varianssia pätevät, noin 99.7% kuulista on käyttökelpoisia. Jos siis kuulia valmistava kone toimii oikein, kuulien hukkaprosentti on noin 0.3%.

Jos kuulien halkaisijan X hajonta olisi todella kasvanut arvoon 0.012 mm, niin ym. välillä on enää 98.8% valmistetuista kuulista, koska

$$P(5 - 0.03 \leq X \leq 5 + 0.03) = 0.988,$$

jossa X noudattaa nyt normaalijakaumaa N(5 mm, (0.012 mm)^2). Tällöin hukkaprosentti olisi siis kasvanut 1.2%:iin, mikä merkitsee sitä, että käyttökelvottomien kuulien suhteellinen osuus olisi kasvanut nelinkertaiseksi!

On siis hyvät syyt hylätä nollahypoteesi ja pysäyttää kuulia valmistava kone tarkistuksia varten. ●

4.4.5 KAHDEN PERUSJOUKON VARIANSSIEN VERTAAMINEN

Otetaan tarkastelun kohteeksi tutkimuksen kohteena olevan satunnaismuuttujan varianssien vertaaminen *kahdessa eri perusjoukossa* (tai saman perusjoukon kahdessa eri ryhmässä). Oletetaan, että tutkimuksen kohteena oleva satunnaismuuttuja on väli- tai suhdeasteikollinen. Tällainen vertailu on kiinnostava sellaisissa tilanteessa, joissa halutaan tietää ovatko ko. muuttujan arvot jakautuneet samalla tavalla kahdessa eri perusjoukossa. Jakaumien vertailu voidaan keskittää seuraaviin osatehtäviin silloin, kun ko. muuttujan jakauma on kummassakin perusjoukossa normaalinen (tai tarpeeksi lähellä normaalijakaumaa):

1. Jakaumien paikan vertailu.
2. Jakaumien keskittyneisyyden eli hajaantuneisuuden vertailu.

Edellinen osatehtävä voidaan ratkaista odotusarvojen vertailuun tarkoitetulla t-testillä. Jakaumien hajaantuneisuuden vertailu voidaan suorittaa seuraavassa kuvattavalla tavalla.

Oletetaan, että perusjoukoista on tehty kaksi *riippumatonta* satunnaisotosta ja olkoon

X_{1i} = muuttujan X arvo 1. perusjoukossa havaintoyksikölle $i = 1,2,\ldots,n_1$,

X_{2j} = muuttujan X arvo 2. perusjoukossa havaintoyksikölle $j = 1,2,\ldots,n_2$,

Oletetaan, että havainnot X_{1i} ja X_{2j} ovat jakautuneet normaalijakauman mukaan niin, että

$$E(X_{1i}) = \mu_1, \quad D^2(X_{1i}) = \sigma_1^2,$$
$$E(X_{2j}) = \mu_2, \quad D^2(X_{1j}) = \sigma_2^2.$$

Olkoon kiinnostuksen kohteena oleva nollahypoteesi muotoa

$$H_0: \quad \sigma_1^2 = \sigma_2^2 = \sigma^2,$$

jossa σ^2 on perusjoukkojen yhteinen varianssi, jos nollahypoteesi pätee. Nollahypoteesin mukaan perusjoukoilla on siis sama varianssi.

Olkoot s_1^2 ja s_2^2 varianssien σ_1^2 ja σ_2^2 harhattomat estimaattorit eli tavanomaiset otosvarianssit:

$$s_1^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(X_{1i} - \bar{X}_1)^2,$$

$$s_2^2 = \frac{1}{n_2}\sum_{j=1}^{n_2}(X_{2j} - \bar{X}_2)^2,$$

jossa

$$\bar{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i},$$

$$\overline{X}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} X_{2j},$$

ovat odotusarvojen μ_1 ja μ_2 harhattomat estimaattorit eli tavanomaiset otoskeskiarvot.

Muodostetaan testisuure

$$F = \frac{s_1^2}{s_2^2}.$$

Jos nollahypoteesi H_0 *pätee,*

$$E(s_1^2) = E(s_2^2) = \sigma^2.$$

Tällöin testisuure F saa keskimäärin arvoa 1 lähellä olevia arvoja. Jos taas nollahypoteesi ei päde, testisuure F saa keskimäärin arvosta 1 kaukana olevia arvoja.

Jos nollahypoteesi pätee, testisuure F on jakautunut kuten *Fisherin* F-*jakauma* vapausastein n_1-1 ja n_2-1. Huomaa, että olemme olettaneet, että perusjoukot ovat jakautuneet normaalisti. Tämä on oletus, jota voidaan testata edellä kuvatulla χ^2-yhteensopivuustestillä.

Nollahypoteesin pätiessä testisuureen F odotusarvo on

$$E(F) = \frac{n_2-1}{n_2-3} \approx 1.$$

Voimakkaasti odotettavissa olevasta arvosta poikkeavat testisuureen arvot johtavat siis nollahypoteesin hylkäämiseen.

F-JAKAUMA

Fisherin F-jakauma on eräs jatkuvista todennäköisyysjakaumista. Se saa positiivisia arvoja vain vaaka-akselin positiivisella puolella ja se on muodoltaan epäsymmetrinen. F-jakauman muoto riippuu *vapausasteiden* lukumääristä f_1 ja f_2. Huomaa, että f_1 liittyy aina F-testisuureen osoittajaan ja f_2 sen nimittäjään.

F-jakautuneen satunnaismuuttujan odotusarvo on

$$\frac{f_2}{f_2-2}$$

ja riippuu siten vain testisuureen nimittäjästä.

F-jakauman kertymäfunktion arvoja on taulukoitu ja useimmat tilasto-ohjelmistot osaavat laskea F-jakaumaan liittyviä todennäköisyyksiä. F-jakaumaa sovelletaan yleisesti sellaisissa testaustilanteissa, joissa vertailtavana on kaksi varianssia.

F-jakauman taulukot on laadittu tavallisesti niin, että taulukoista voidaan lukea vain *suuriin* F-testisuureen arvoihin liittyviä kriittisiä arvoja. Esimerkiksi liitteen taulukossa on taulukoitu yhtälön

$$P(F \geq F) = p$$

määräämiä kriittisiä arvoja F, kun merkitsevyystaso p saa arvot 0.05 ja 0.01. Esimerkiksi, jos $f_1 = 12, f_2 = 20$ ja $p = 0.05$, niin F = 2.28. Siten

$$P(F \geq 2.28) = 0.05.$$

Edellä esitetyssä sovelluksessa nollahypoteesin hylkäämiseen johtavat myös *pienet* F-testisuureen arvot. Miten näitä arvoja vastaavat todennäköisyydet määrätään em. taulukosta? Ongelma ratkeaa käyttämällä apuna seuraavaa tulosta: Jos F noudattaa F-jakaumaan vapausastein f_1 ja f_2, niin $1/F$ noudattaa F-jakaumaa vapausastein f_2 ja f_1. Oletetaan esimerkiksi, että haluamme määrätä kriittisen arvon G siten, että

$$P(F \leq G) = 0.01,$$

kun $f_1 = 6$, $f_2 = 12$. Tällöin todennäköisyyteen 0.01 liittyvästä taulukosta haetaan vapausasteisiin $f_1 = 12, f_2 = 6$ liittyvä kriittinen arvo F siten, että

$$P(F \geq F) = 0.01.$$

Taulukoista nähdään, että F = 7.72. Kysytty kriittinen arvo on siten

$$G = \frac{1}{F} = \frac{1}{7.72} \approx 0.13.$$

Kuvattu ongelma voidaan tietysti ratkaista myös siten, että F-testisuure varianssien vertailemiseksi pyritään muodostamaan siten, että *suurempi* otosvariansseista on osoittajassa.

ESIMERKKI 1.

Tarkastellaan jälleen tämän luvun johdannossa esitettyä esimerkkiä psykologisesta testistä, jossa tutkittiin suorituskykyä meluisassa ja meluttomassa ympäristössä.

Kahden perusjoukon odotusarvojen vertaamiseen tarkoitettua t-testiä käsitelleessä luvussa testi tehtiin olettaen, että varianssit olivat perusjoukoissa samat. Oliko tämä oletus oikeutettu?

Otosvarianssit olivat seuraavat:

Meluton ympäristö: $s_1^2 = 2.28^2$,

Meluisa ympäristö: $s_2^2 = 2.45^2$.

F-testisuureen arvoksi tulee siten

$$F = \frac{2.45^2}{2.28^2} \approx 1.15.$$

Vastaava P-arvo on 0.39, ts.

$$P(F \geq 1.15) = 0.39.$$

Nollahypoteesia varianssien samuudesta ei siis ole syytä hylätä ja oli oikeutettua suorittaa t-testi odotusarvojen samuudelle olettamalla, että varianssit olivat yhtä suuria. ●

4.4.6 RIIPPUMATTOMUUDEN TESTAUS

Haluamme tutkia kahden väli- tai suhdeasteikollisen satunnaismuuttujan riippumattomuutta. Oletetaan, että perusjoukosta on tehty riippumaton satunnaisotos ja olkoon

$$X_i = \text{muuttujan } X \text{ arvo havaintoyksikölle } i = 1,2,\ldots,n,$$

$$Y_i = \text{muuttujan } Y \text{ arvo havaintoyksikölle } i = 1,2,\ldots,n.$$

Oletetaan, että havainnot X_i ja Y_i ovat jakautuneet siten, että

$$E(X_i) = \mu_X, \qquad D^2(X_i) = \sigma_X^2,$$

$$E(Y_i) = \mu_Y, \qquad D^2(Y_i) = \sigma_Y^2.$$

ja, että muuttujien X ja Y välinen korrelaatio perusjoukossa on ρ.[1]

Voidaan osoittaa, että korrelaation ρ harhaton estimaattori on ns. Pearsonin tulomomenttikorrelaatiokerroin eli *otoskorrelaatiokerroin*

$$r = \frac{s_{XY}}{s_X s_Y},$$

jossa

$$s_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

on ns. *otoskovarianssi* ja s_X ja s_Y ovat tavanomaiset harhattomat otoshajonnat.

Oletetaan, että nollahypoteesi on muotoa

$$H_0: \quad \rho = 0.$$

Nollahypoteesin mukaan satunnaismuuttujat X ja Y ovat *korreloimattomia*. Jos satunnaismuuttujat X ja Y noudattavat ns. *kaksiulotteista normaalijakaumaa*[2], tästä seuraa myös muuttujien X ja Y riippumattomuus. Tässä yhteydessä on syytä muistaa, että korreloimattomuus merkitsee korreloimattomuutta *lineaarisessa* mielessä: Muuttujien välillä saattaa olla jopa tarkka *epälineaarinen* riippuvuus, vaikka ne ovat korreloimattomia.

Muodostetaan testisuure

$$Z = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}.$$

Jos nollahypoteesi H_0 pätee,

---

[1] Perusjoukon (teoreettisen) korrelaation käsitettä ei ole määritely. Tässä riittää tieto siitä, että perusjoukon korrelaatiolla tarkoitetaan kahden satunnaismuuttujan odotettavissa olevaa korrelaatiota, kun odotetettavissa oleva korrelaatio määritellään ns. *tulomomenttikorrelaatiokertoimena*.

[2] Normaalijakauma voidaan yleistää useampiulotteisiin tilanteesiin. Ns. *moniulotteisella normaalijakaumalla* on keskeinen rooli esimerkiksi ns. *monimuuttujamenetelmien* yhteydessä. Sivuutamme kuitenkin tämän tärkeän jakauman käsittelyn tällä kurssilla.

$$E(r) = 0.$$

Tällöin testisuureen Z itseisarvo $|Z|$ saa keskimäärin pieniä arvoja. Jos taas H_0 ei päde, $|Z|$ saa keskimäärin suuria arvoja.

Jos nollahypoteesi H_0 *pätee*, niin testisuure Z on jakautunut suurissa otoksissa approksimatiivisesti kuten standardoitu normaalijakauma N(0,1). Nollahypoteesin pätiessä testisuureen Z odotusarvo on

$$E(Z) = 0.$$

Voimakkaasti nollasta poikkeavat testisuureen arvot johtavat nollahypoteesin hylkäämiseen. Jos satunnaismuuttujien X ja Y jakauma on kaksiulotteinen normaalijakauma, Z jakautuu kuten Studentin t-jakauma $(n-2)$:lla vapausasteella.

Todettakoon vielä, että *Spearmanin järjestyskorrelaatiokerroin* ρ voidaan määrätä soveltamalla otoskorrelaatiokertoimen kaavaa *järjestysnumeroihin*. Tämä selittää sen, että testisuure korreloimattomuudelle on otoskorrelaatiokertoimen r ja Spearmanin ρ:n yhteydessä samaa muotoa.

ESIMERKKI 1.

Eräälle tilastollisen tietojenkäsittelyn kurssille osallistui 10 opiskelijaa. Kurssin suorittaminen koostui kahdesta osasuorituksesta: teoriakuulustelusta ja harjoitustyöstä. Opiskelijat asettiin paremmuusjärjestykseen molempien osasuoritusten perusteella. Seuraava taulukko kuvaa näitä paremmuusjärjestyksiä (1 = paras, 10 = huonoin). Tehtävänä on tutkia ovatko opiskelijoiden paremmuusjärjestykset eri osasuorituksissa riippumattomia. Jos riippumattomuushypoteesi kumoutuu, opiskelijoiden osasuoritukset korreloivat keskenään. Tällöin opiskelijan suoritustason tunteminen toisessa osasuorituksessa auttaa ennustamaan hänen suoritustasoaan toisessa osasuorituksessa: Hyvän harjoitustyön tehneet opiskelijat ovat menestyneet hyvin myös teoriakuulustelussa ja kääntäen, huonon harjoitustyön tehneet opiskelijat ovat menestyneet huonosti teoriakuulustelussa.

Opiskelijat on asetettu taulukkoon aakkosjärjestykseen. Koska sijaluvut ovat järjestysasteikollisia muuttujia, sovelletaan sijalukujen korrelaation laskemiseen Spearmanin korrelaatiokerrointa. Taulukon kahdessa viimeisessä sarakkeessa on esitetty Spearmanin ρ:n määräämiseksi tarvittavat aputulokset.

Opiskelija i	Sijaluku harjoitus-työssä X_i	Sijaluku-teoria-kuuluste-lussa Y_i	Sija-lukujen erotus d_i	Sija-lukujen erotusten neliö d_i^2
A	8	9	−1	1
B	3	5	−2	4
C	9	10	−1	1
D	2	1	1	1
E	7	8	−1	1
F	10	7	3	9
G	4	3	1	1
H	6	4	2	4
I	1	2	−1	1
J	5	6	−1	1

Spearmanin järjestyskorrelaatiokertoimen kaavan mukaan

$$\begin{aligned}\rho &= 1-\frac{6\sum_{i=1}^{n} d_i^2}{n^3-n}\\ &= 1-\frac{6\cdot 24}{1000-10}\\ &= 0.855.\end{aligned}$$

Sama tulos saadaan myös soveltamalla otoskorrelaatiokertoimen kaavaa, mutta monimutkaisemmin laskutoimituksin.

Testisuureen Z arvoksi saadaan

$$Z=\frac{0.855\sqrt{10-2}}{\sqrt{1-0.855^2}}\approx 4.65.$$

Vastaava P-arvo on 0.0000016, joten nollahypoteesi osasuoritusten riippumattomuudesta voidaan hylätä. ●

5. REGRESSIOANALYYSI

5.1 YHDEN SELITTÄJÄN REGRESSIOMALLIT

Yhden selittäjän regressiomallia on käsitelty Kirjassa 1. Tähän malliin liittyvät kaavat on esitetty myös tämän luvun lopussa.

5.2 USEAN SELITTÄJÄN REGRESSIOMALLIT

5.2.1 JOHDANTO

Riippuvuuksien tutkiminen on tieteen tärkeimpiä tehtäviä. Tilastotieteessä riippuvuuksia analysoidaan mm. *regressiomallien* avulla. Regressioanalyysi käsittää laajan joukon menetelmiä, joissa pyritään mallittamaan jonkin muuttujan, ns. *selitettävän muuttujan* eli *vastemuuttujan* riippuvuutta toisista muuttujista, ns. *selittävistä muuttujista* eli *selittäjistä* sopivilla selittäjien funktioilla. On syytä huomata, että selitettävän ja selittävän muuttujan välinen ero on regressiomalleja sovellettaessa perustavaa laatua. Regressioanalyysin päämääränä on selvittää riippuvuuden muoto ja voimakkuus muuttujista kerätyn havaintoaineiston perusteella. Analyysin eräs osatehtävistä on selvittää riippuuko selitettävä muuttuja todellakin niistä selittäjistä, joista sen on ajateltu riippuvan.

Seuraavassa tarkastellaan sellaisia tilanteita, joissa selitettävän muuttujan ja selittäjien välistä riippuvuutta kuvataan *lineaarisella funktiolla*.

Oletetaan, että tehtävänä on selvittää miten selitettävän muuttujan Y (satunnaismuuttuja) keskimääräinen arvo riippuu selittäjien $x_1, x_2, \ldots, x_p$ arvoista. Oletamme, että selitettävän muuttujan Y *odotusarvo* $\mathrm{E}(Y) = \mu_Y$ on selittävien muuttujien $x_1, x_2, \ldots, x_p$ arvojen *lineaarinen funktio*:

$$\mathrm{E}(Y) = \mu_Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p,$$

jossa kertoimia $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ sanotaan *regressiokertoimiksi*. Koska regressiokertoimien arvoja ei tavallisesti tunneta, regressioanalyysin eräs osatehtävistä on kertoimien estimointi eli arviointi havaintojen perusteella.

Jos yo. yhtälö pätee, vastemuuttujan arvon eli *vasteen* Y ja selittäjien arvojen $x_1, x_2, \ldots, x_p$ välillä vallitsee *lineaarinen tilastollinen riippuvuus*. Nimitys lineaarinen johtuu siitä, että yo. lausekkeen oikea puoli määrittelee *tason* yhtälön p-ulotteisessa avaruudessa ja taso on esimerkki lineaarisesta funktiosta.[1]

[1] *Lineaarinen* = suoraviivainen tai viivallinen.

Yllä esitetty yhtälö on vastemuuttujan ja selittäjien välinen *regressioyhtälö* perusjoukossa. Yhtälöä ei voida havaita suoraan, koska Y-muuttujan havaitut arvot eli vasteet vaihtelevat niiden odotusarvon μ_Y ympärillä. On syytä huomata, että vastemuuttujan Y odotusarvoa tarkastellaan sillä ehdolla, että selittäjät ovat saaneet arvot $x_1, x_2, \ldots, x_p$. Jokainen selittäjien arvojen yhdistelmä määrittelee Y:n mahdollisten arvojen muodostaman perusjoukon osajoukon. Vaste Y vaihtelee näin määritellyissä osajoukoissa siten, että Y:n odotusarvona on yllä määritelty selittäjien arvojen lineaarinen funktio.

Lineaarisen mallin formulaatioossa on mahdollista ottaa huomioon myös sellainen tilanne, jossa selittäjät ovat *satunnaismuuttujia*, mikä onkin tavanomainen tilanne. Selitettävän muuttujan odotusarvoa tarkastellaan tällöin *ehdollisesti* selittäjien havaittujen arvojen suhteen, mikä merkitsee selittäjien arvojen *kiinnittämistä* havaittuihin arvoihinsa. Kiinnittäminen johtaa sihen, että selittäjien arvot voidaan tulkita *ei-satunnaisiksi*. Joissakin tilanteissa selittäjien arvot ovat aidosti ei-satunnaisia. Tämä merkitsee yleensä sitä, että selittäjien arvot on voitu *kiinnittää* tai *valita* halutuiksi. Tämä on tavallisesti mahdollista vain *koeasetelmien* yhteydessä ja tällöinkään ei ole yleensä mahdollista valita kaikkia selitettävään muuttujaan vaikuttavien tekijöiden arvoja.

Kertoimia $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ kutsutaan siis *regressiokertoimiksi*. Erityisesti kerrointa β_0 kutsutaan regressioyhtälön *vakioksi*. Vakio kuvaa selitettävän muuttujan odotettavissa olevaa arvoa, jos kaikilla muilla selittäjillä on arvo 0. Kertoimilla β_i, $i = 1,2,\ldots,p$ on seuraava tulkinta: Tarkastellaan esimerkiksi kerrointa β_1. Oletetaan, että selittäjillä on kiinteät arvot $x_1, x_2, \ldots, x_p$. Oletetaan, että kerrointa β_1 vastaavan selittäjän arvo x_1 kasvaa yhdellä yksiköllä, ts.

$$x_1 \to x_1 + 1$$

kaikkien muiden selittäjien arvojen pysyessä ennallaan. Tällöin selitettävän muuttujan Y odotusarvo muuttuu β_1 yksiköllä, koska

$$\begin{aligned}\mu_Y &\to \beta_0 + \beta_1(x_1+1) + \beta_2 x_2 + \ldots + \beta_p x_p \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \beta_1 \\ &= \mu_Y + \beta_1.\end{aligned}$$

Regressiokerroin kuvaa siis sitä, miten selitettävän muuttujan odotettavissa oleva arvo muuttuu, jos kerrointa vastaavan selittäjän arvo kasvaa yhdellä yksiköllä ja kaikkien muiden selittäjien arvot pysyvät samaan aikaan muuttumattomina. Jos regressiokerroin $\beta_i = 0$, vaste ei riipu vastaavasta selittäjästä.

Seuraavassa tarkastellaan miten tämä matemaattinen malli voidaan liittää *havaintoihin* tilastolliseksi malliksi.

Oletetaan, että selitettävästä muuttujasta Y ja selittäjistä $x_1, x_2, \ldots, x_p$ on käytettävissä havaintoyksiköitä $i = 1,2,\ldots,n$ koskevat havaintoarvot:

Havaintoyksikkö 1: $Y_1, x_{11}, x_{12}, \ldots, x_{1p}$,

Havaintoyksikkö 2: $Y_2, x_{21}, x_{22}, \ldots, x_{2p}$,

...

Havaintoyksikkö n: $Y_n, x_{n1}, x_{n2}, \ldots, x_{np}$.

Siten

Y_j = selitettävän muuttujan Y arvo havaintoyksikössä j,

x_{ji} = selittäjän x_i arvo havaintoyksikössä j,

n = havaintojen lukumäärä,

p = selittäjien lukumäärä.

Havainnot toteuttavat vain poikkeuksellisesti yllä esitetyn perusjoukon regressioyhtälön. Poikkeamat tästä yhtälöstä vaihtelevat satunnaisesti havainnosta toiseen. Tekemällä poikkeamista sopivia oletuksia voidaan havainnoille muodostaa tilastollinen malli, jota kutsutaan *lineaariseksi regressiomalliksi*.

LINEAARISTA REGRESSIOMALLIA KOSKEVAT STANDARDIOLETUKSET

Oletetaan, että havaintojen välillä vallitsee yhtälö

$$Y_j = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \beta_p x_{jp} + \varepsilon_j, \qquad j = 1,2,\ldots,n,$$

jossa

ε_j = *poikkeama* eli *virhe*- eli *jäännöstermi* havainnossa j.

Vastemuuttujan Y arvot Y_j voidaan jakaa tämän yhtälön mukaan *rakenneosaan* ja *jäännökseen* niin, että

$$Y_j = \textit{rakenne} + \textit{jäännös},$$

jossa

$$\textit{rakenne} = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \beta_p x_{jp}$$

on selittäjien arvon lineaarinen funktio ja kuvaa *systemaattista osaa* havaintoarvossa Y_j ja

$$\textit{jäännös} = \varepsilon_j$$

kuvaa *satunnaista osaa* havaintoarvossa Y_j.

Selittäjistä tehdään seuraavat oletukset:

1. Selittäjien arvot ovat kiinteitä eli ei-satunnaisia.

   Tästä oletuksesta on mahdollista luopua, kunhan jäännöstermin ominaisuuksia koskevat oletukset 3 ja 4 muunnetaan sopivaan muotoon. Näitä muunnettuja oletuksia tarvitaan erityisesti aikasarja-aineistojen yhteydessä, mutta myös monissa muissa sovelluksissa selittäjien arvot ovat satunnaisia. Sivuutamme tässä näiden muunnettujen oletusten tarkemman käsittelyn.

2. Minkään selittäjän arvot eivät riipu lineaarisesti toisten selittäjien arvoista.

   Oletus 2 on tekninen ehto, joka takaa sen, että regressiokertoimille voidaan määrätä pienimmän neliösumman estimaattorit tavanomaiseen tapaan. Ehdosta seuraa mm. se, että sama selittäjä ei saa esiintyä kahdesti selittäjien joukossa ja se, että minkään selittäjän arvoja ei voida lausua kahden muun selittäjän arvojen (painotettuna) summana. Jos ehto ei jostakin syystä päde, joutuvat kertoimien estimoinnissa käytetyt tilasto-ohjelmat tavallisesti vaikeuksiin. Missään tavanomaisessa tapauksessa ei ole pelkoa siitä, ettei ehto pätisi.

Virhetermistä ε_j tehdään seuraavat oletukset:

3. Virhetermin ε_j odotusarvo on nolla:

$$E(\varepsilon_j) = 0 \text{ kaikille } j = 1,2,\ldots,n.$$

Ehto takaa sen, että selitettävän muuttujan arvojen eli vasteiden odotusarvot ovat muotoa

$$E(Y_j) = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \beta_p x_{jp} \text{ kaikille } j = 1,2,\ldots,n.$$

Tämä merkitsee sitä, että havaintoarvon Y_j odotusarvo yhtyy mallin *systemaattiseen osaan*.

4. Virhetermin ε_j varianssi eli *jäännösvarianssi* on vakio:

$$D^2(\varepsilon_j) = \sigma^2 \text{ kaikille } j = 1,2,\ldots,n.$$

Ehdosta seuraa se, että

$$D^2(Y_j) = \sigma^2.$$

Tämä merkitsee sitä, että havaintoarvojen Y_j keskittyneisyys oman odotusarvonsa $E(Y_j)$ ympärille on kaikille havainnoille j sama.

Oletusta 4 kutsutaan usein *homoskedastisuusoletukseksi*. Jos se ei päde sanotaan, että virhetermit ovat *heteroskedastisia*. Homoskedastisuusoletuksen mukaan virhetermien jakauma keskittyy

Koska σ^2 on tavallisesti tuntematon, regressioanalyysin eräs osatehtävistä on jäännösvarianssin estimointi havaintojen perusteella.

5. Virhetermit ε_j ovat keskenään korreloimattomia.

Oletus 5 on merkityksellinen lähinnä aikasarja-aineistojen yhteydessä. Sivuutamme ehdon tarkemman käsittelyn tässä yhteydessä.

Oletuksia 1.—5. kutsutaan lineaariseen regressiomalliin liittyviksi *standardioletuksiksi*.

Oletuksiin 3 ja 4 liitetään usein normaalisuusoletus:

6. Virhetermi ε_j on jakautunut normaalijakauman mukaan:

$$\varepsilon_j \sim N(0, \sigma^2).$$

On syytä huomata, että jos oletukset 1.—6. eivät päde saattaa olla syytä käyttää toisenlaisia tilastollisia menetelmiä, kuin niitä, joita käsitellään alla.

Lineaarisen regressiomallin *parametreja* ovat regressiokertoimet $\beta_0, \beta_1, \beta_2,\ldots,\beta_p$ ja jäännösvarianssi σ^2. Regressioanalyysin päätehtäviä on näiden parametrien estimointi ja niitä koskevien hypoteesien testaaminen. Lisäksi analyysiin pitää aina liittää mallista tehtyjen oletusten 1.—6. tarkistaminen. On hyvä tietää, että standardioletuksia voidaan tutkia graafisesti tai tekemällä sopivia tilastollisia testejä. Standardioletusten tarkistamista kutsutaan usein *diagnostiikaksi*. Sivuutamme diagnostisten menetelmien kuvailun tässä esityksessä. Jos malli toteuttaa siitä tehdyt oletukset ja malli on muutenkin järkevä, estimoitua mallia voidaan käyttää paitsi selitettävän muuttujan ja selittäjien välisen riippuvuuden *kuvaamiseen*, selitettävän muuttujan arvojen *ennustamiseen* ja jopa selitettävän muuttujan arvojen *kontrollointiin*.

ESIMERKKI 1.

Alkoholin kulutusta tutkittaessa formuloidaan usein regressiomalli, jolla pyritään selittämään alkoholin vuosikulutusta absoluuttialkoholilitroina *per capita*[1], kun selittäjinä ovat alkoholin *reaalihinta* ja *reaalinen vuositulo per capita*.[2]

Eräs sellainen malli on muodoltaan seuraava:

$$Y_j = \beta_0 + \beta_1 p_j + \beta_2 Q_j + \varepsilon_j, \qquad j = 1,2,\ldots,n,$$

jossa

Y_j = alkoholin kulutus absoluuttialkoholina (l) per capita vuonna j,

p_j = alkoholin reaalihinta (mk) vuonna j,

Q_j = reaalitulo per capita (mk) vuonna j.

Oletetaan, että malli toteuttaa standardioletukset sopivasti muunnettuina: Alkoholin *nimellishinta* on ALKO:n hallintoneuvoston valittavissa ja on siten ei-satunnainen. Sen sijaan reaalihinta on satunnaismuuttuja, koska elinkustannusindeksi on satunnaismuuttuja. Samoin reaalitulo on satunnaismuuttuja.

Tällaista mallia voidaan käyttää sekä alkoholin kulutuksen *ennustamiseen* että *kontrollointiin*: Jos reaalitulojen ennustetaan pienenevän seuraavana vuonna, voidaan mallin perusteella ennustaa paljonko kulutus tulee muuttumaan. Jos alkoholin kulutusta halutaan vähentää tietyn määrän, voidaan mallin perusteella määrätä paljonko alkoholin hintaa on muutettava, jotta tavoite saavutetaan.

Regressiokertoimilla on seuraavat tulkinnat:

β_0: Kerroin ilmaisee paljonko alkoholia kulutettaisiin, jos alkoholin reaalihinta ja reaalitulot saisivat arvon 0.

β_1: Kerroin ilmaisee paljonko alkoholin kulutus muuttuu, jos alkoholin reaalihinta kasvaa yhdellä markalla.

β_2: Kerroin ilmaisee paljonko alkoholin kulutus muuttuu, jos reaalitulot kasvavat yhdellä markalla.

---

[1] *Per capita* tarkoittaa henkilöä kohti.

[2] *Reaalihinnalla* tarkoitetaan hintaa, jossa on otettu huomioon inflaatio eli rahan arvon huononeminen. Reaalihinta saadaan *nimellishinnan* ja elinkustannusindeksin suhteena. Vastaavasti reaalitulolla tarkoitetaan tuloa, jossa on otettu huomioon inflaatio eli rahan arvon huononeminen. Reaalitulo saadaan *nimellistulon* ja elinkustannusindeksin suhteena.

Huomaa, että regressiokertoimien merkeistä on käytettävissä *priori-* eli *ennakkotietoa*:

$\beta_1 \leq 0$:

Kerroin β_1 on *negatiivinen*, koska minkä tahansa tuotteen reaalihinnan nousu vähentää tuotteen kulutusta.

$\beta_2 \geq 0$:

Kerroin β_2 on luultavasti *positiivinen*, koska reaalitulojen nousu lisää sellaisten tuotteiden kulutusta kuin alkoholi.

Regressioanalyysin tarkoituksena on mm. estimoida regressiokertoimien arvot ja tarkistaa ovatko kertoimien merkit ennakkokäsitysten mukaisia. Tällainen järkevyystarkistus on keskeisessä asemassa regressioanalyysissa. Jos hintamuuttujaa vastaava regeressiokertoimen β_1 estimaatti saa positiivisen merkin, ollaan pahoissa vaikeuksissa, vaikka malli olisi *tilastollisilta ominaisuuksiltaan* kuinka hyväksyttävä tahansa. ●

5.2.2 LINEAARISEN REGRESSIOMALLIN ESTIMOINTI

Lineaarisen regressiomallin

$$Y_j = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \ldots + \beta_p x_{jp} + \varepsilon_j, \qquad j = 1,2,\ldots,n,$$

parametrit estimoidaan tavallisesti *pienimmän neliösumman keinolla*, lyh. PNS-keinolla. Regressiokertoimien $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ PNS-estimaattorit saadaan minimoimalla neliösumma

$$\sum_{j=1}^{n} \varepsilon_j^2 = \sum_{j=1}^{n} (Y_j - \beta_0 - \beta_1 x_{j1} - \beta_j x_{j2} - \cdots - \beta_p x_{jp})^2$$

parametrien $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ suhteen. Jos virhetermit ε_j ovat normaalijakautuneita, *suurimman uskottavuuden* menetelmä tuottaa parametreille $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ samat estimaattorit kuin PNS-menetelmä. Sivuutamme tässä kokonaan regressiokertoimien estimaattien laskukaavat. Alla tosin esitetään yhden selittäjän malliin liittyvät kaavat.

Oletetaan, että PNS-estimaattorit kertoimille $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ ovat $b_0, b_1, b_2, \ldots, b_p$. Niiden avulla voidaan määritellä seuraavat käsitteet:

Sovite:

$$\hat{Y}_j = b_0 + b_1 x_{j1} + b_2 x_{j2} + \ldots + b_p x_{jp}, \qquad j = 1,2,\ldots,n.$$

Sovite on estimoidun mallin selitettävälle muuttujalle Y antama, selittäjien arvoja $x_{j1}, x_{j2}, \ldots, x_{jp}$ vastaava arvo. $\hat{Y}_j$ luetaan y-j-hattu.

Residuaali:

$$e_j = Y_j - \hat{Y}_j = Y_j - b_0 - b_1 x_{j1} - b_2 x_{j2} - \ldots - b_p x_{jp}, \qquad j = 1,2,\ldots,n.$$

Residuaali on selitettävän muuttujan havaitun arvon Y_j ja estimoidun mallin selitettävälle muuttujalle antaman arvon $\hat{Y}_j$ erotus. Residuaalia voidaan pitää jäännöstermin ε_j havaittuna vastineena, jos malli on oikein muodostettu.

Lienee ilmeistä, että malli *selittää* selitettävän muuttujan arvojen vaihtelun sitä paremmin, mitä lähempänä sovite on selitettävän muuttujan havaittuja arvoja eli mitä pienempiä ovat residuaalit.

Virhetermin varianssi eli jäännösvarianssi σ^2 voidaan estimoida residuaalien avulla: Voidaan osoittaa, että

$$s^2 = \frac{1}{n-p-1} \sum_{j=1}^{n} e_j^2$$

on jäännösvarianssin σ^2 harhaton estimaattori. s^2 on residuaalien varianssi ja kuvaa siten residuaalien vaihtelua niiden keskiarvon ympärillä. Tulkinnan kannalta on tärkeätä huomata se, että s^2 kuvaa selitettävän muuttujan havaittujen arvojen vaihtelua *regressiotason*

$$y = b_0 - b_1 x_1 - b_2 x_2 - \ldots - b_p x_p$$

ympärillä.

Se, että edellä esitetty kaava s^2:lle on todellakin residuaalien *varianssi*, seuraa siitä, että residuaalien aritmeettinen keskiarvo

$$\bar{e} = \frac{1}{n}\sum_{j=1}^{n} e_j = 0$$

aina, kun mallissa on vakio.

Kuten tunnettua selitettävän muuttujan arvojen vaihtelun mittaaminen voidaan perustaa neliösummaan

$$SST = \sum_{j=1}^{n} (Y_j - \bar{Y})^2,$$

jossa

$$\bar{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j$$

on selitettävän muuttujan arvojen otoskeskiarvo. Neliösummaa SST kutsutaan *kokonaisneliösummaksi*, koska se kuvaa vasteiden vaihtelua niiden keskiarvon ympärillä. Selitettävän muuttujan Y havaittujen arvojen $Y_1, Y_2, \ldots, Y_n$ otosvarianssi s_Y^2 voidaan määritellä kokonaisneliösumman SST avulla seuraavalla kaavalla:

$$s_Y^2 = \frac{SST}{n-1}.$$

Residuaalien vaihtelun mittaaminen taas voidaan perustaa neliösummaan

$$SSE = \sum_{j=1}^{n} e_j^2.$$

Neliösummaa SSE kutsutaan *jäännösneliösummaksi*. Jäännösvarianssin estimaattori s^2 voidaan määritellä jäännösneliösumman SSE avulla seuraavalla kaavalla:

$$s^2 = \frac{SSE}{n-p-1}.$$

Regressioanalyysin keskeisiä tuloksia on se, että jäännösneliösumma on aina korkeintaan yhtä suuri kuin kokonaisneliösumma:

$$SSE \le SST.$$

Erotusta

$$SSM = SST - SSE$$

kutsutaan *regressioneliösummaksi* tai *mallineliösummaksi*. Voidaan osoittaa, että

$$SSM = \sum_{j=1}^{n} (\hat{Y}_j - \bar{Y})^2.$$

Jäännösneliösumman ja kokonaisneliösumman välinen epäyhtälö mahdollistaa sen, että suuretta

$$R^2 = 1 - \frac{SSE}{SST}$$

voidaan käyttää mallin *selitysasteen* mittaamiseen. Selitysasteella R^2 on seuraavat ominaisuudet:

1. $0 \leq R^2 \leq 1$.

2. Jos kaikki residuaalit ovat nollia, jolloin malli sopii havaintoihin täydellisesti,

$$R^2 = 1.$$

3. Jos kaikki residuaalit ovat muotoa $Y_i - \overline{Y}$, jäännösneliösumma yhtyy kokonaisneliösummaan ja

$$R^2 = 0.$$

Tällöin malli ei selitä ollenkaan selitettävän muuttujan arvojen vaihtelua.

Selitysaste kuvaa siis sitä, miten hyvin malli selittää selitettävän muuttujan vaihtelua.

REGRESSIOKERTOIMIEN OTOSJAKAUMA

Regressiokertoimien PNS-estimaattorit ovat regressiomallin standardioletuksien 1—5 pätiessä *harhattomia*:

$$\mathrm{E}(b_i) = \beta_i.$$

Voidaan osoittaa, että standardioletuksien 1—5 pätiessä regressiokertoimien estimaattorit ovat suurissa otoksissa approksimatiivisesti normaalisia $\mathrm{N}(\beta_i, \sigma_{b_i}^2)$, jossa $\sigma_{b_i}^2$ on kerroinestimaattorin b_i varianssi:

$$b_i \sim_a \mathrm{N}(\beta_i, \sigma_{\beta_i}^2).$$

Jos jäännöstermit ovat normaalisia eli, jos standardioletukset 1—6 pätevät, regressiokertoimen estimaattorin jakaumaa koskeva tulos on eksakti eli tarkka.

REGRESSIOKERTOIMIEN LUOTTAMUSVÄLI

Rergressiokertoimien estimaattoreiden varianssit voidaan estimoida havainnoista. Sivuutamme tässä varianssien laskukaavat. Merkitään kertoimen b_i estimoitua varianssia symbolilla $s^2_{b_i}$. Kertoimen β_i luottamusväli luottamustasolla $1-\alpha$ on standardioletusten 1—5 pätiessä muotoa

$$b_i \pm z s_{b_i},$$

jossa z on luottamustasoa $1-\alpha$ vastaava *luottamuskerroin*. Luottamuskerroin z saadaan normaalijakaumasta, jos regressiokertoimen otosjakauma on suurissa otoksissa approksimatiivisesti normaalinen. Jos regressiokertoimen otosjakauma on eksaktisti normaalinen, z saadaan t-jakaumasta, jossa vapausasteiden lukumäärä on $n-p-1$.

5.2.3 REGRESSIOMALLIIN LIITTYVÄT TESTIT

Käsittelemme seuraavassa regressiomallin kertoimille tehtäviä testejä. Regressiokertoimille on aina tapana tehdä seuraavat testit:

1. *Yleistestinä* kertoimille käytetään testiä, jolla tutkitaan auttavatko malliin valitut *selittävät muuttujat yhdessä* selittämään selitettävän muuttujan arvojen vaihtelua. Tällä testillä tutkitaan *onko mallilla ilmaistua regressiota olemassa.*
2. Jokaiselle kertoimelle on tapana tehdä *kerroinkohtainen testi*, jolla tutkitaan auttaako *se selittävä muuttuja*, jonka kerrointa testataan, selittämään selitettävän muuttujan arvojen vaihtelua.

ONKO REGRESSIOTA?

Yleistestissä regressiokertoimille testataan nollahypoteesia

$$H_0: \quad \beta_1 = \beta_2 = \ldots = \beta_p = 0.$$

Jos H_0 pätee, selitettävä muuttuja Y ei riipu lineaarisesti *yhdestäkään* selittäjästä $x_1, x_2, \ldots, x_p$. Tässä on esimerkki tilanteesta, jossa tavallisesti toivotaan nollahypoteesin kumoutumista.

Nollahypoteesia H_0 voidaan testata testisuureella

$$F = \frac{n-p-1}{p} \cdot \frac{SSM}{SSE}$$
$$= \frac{n-p-1}{p} \cdot \frac{R^2}{1-R^2},$$

jossa SSM on *mallineliösumma*, SSE on *jäännösneliösumma* ja R^2 on *selitysaste* (kts. edellinen kappale). Testisuure F on jakautunut kuten F-jakauma vapausastein p ja $n-p-1$, *jos nollahypoteesi* H_0 *pätee* ja, jos standardioletukset 1—6 pätevät. Vaikka normaalisuusoletus 6 ei olisikaan tosi, tulos pätee approksimatiivisesti suurissa otoksissa, jos regressiokertoimien jakauma on approksimatiivisesti normaalinen.

Nollahypoteesin pätiessä testisuureen odotusarvo on

$$E(F) = \frac{n-p-1}{n-p-3} \approx 1.$$

Jos testisuure saa odotettavissa olevaa arvoaan kyllin paljon suurempia arvoja, nollahypoteesi joudutaan hylkäämään.

Testiä voidaan pitää *varianssien vertailutestinä*, joka perustuu *jäännösvarianssin*

$$s^2 = \frac{1}{n-p-1} SSE$$

ja *mallivarianssin*

$$\frac{1}{p} SSM$$

vertailuun.

YKSITTÄISEN SELITTÄJÄN MERKITYS

Edellisessä kappaleessa esitetyn yleistestin *jälkeen* on tapana tutkia jokaisen selittäjän merkitystä mallissa. Testattavat nollahypoteesit ovat muotoa

$$H_{0i}: \quad \beta_i = 0, \qquad i = 1,2,\ldots,p.$$

Jos nollahypoteesi H_{0i} *pätee*, selitettävä muuttuja Y ei riipu lineaarisesti selittäjästä x_i. Tällöin selittäjä x_i on mallissa turha ja se voidaan poistaa mallista.

Muodostetaan testisuure

$$t_i = \frac{b_i}{s_{b_i}},$$

jossa b_i on kertoimen β_i PNS-estimaattori ja s_{b_i} on estimaattorin b_i hajonnan estimaattori. Testisuuretta t_i kutsutaan tavallisesti kertoimen β_i *t-arvoksi*.

Testisuure t_i on jakautunut suurissa otoksissa approksimatiivisesti kuten normaalijakauma, *jos nollahypoteesi pätee* ja, jos estimaattorin b_i otosjakauma on suurissa otoksissa approksimatiivisesti normaalinen. Testisuure t_i on jakautunut kuten t-jakauma $(n-p-1)$:llä vapausasteella, *jos nollahypoteesi pätee* ja, jos estimaattorin b_i otosjakauma on eksaktisti normaalinen. Nollahypoteesin pätiessä testisuureen odotusarvo on

$$E(t_i) = 0.$$

Jos kertoimen β_i t-arvo poikkeaa nollasta kyllin paljon, joudutaan nollahypoteesi hylkäämään.

Selittäjät voidaan jakaa niiden regressiokertoimien t-arvojen avulla *tilastollisesti merkityksellisiin ja merkityksettömiin*. Valitaan sopiva, pieni todennäköisyys α.[1] Jos regressiokertoimen β_i t-arvoa t_i vastaava P-arvo on *suurempi* kuin

[1] Yleensä $\alpha = 0.05$ tai 0.01.

α, sanotaan, että kerrointa β_i vastaavaa selittäjä x_i on tilastollisesti *merkityksetön*. Jos taas regressiokertoimen β_i t-arvoa t_i vastaava *P*-arvo on *pienempi* kuin α, sanotaan, että kerrointa β_i vastaava selittäjä x_i on tilastollisesti *merkityksellinen*.

Tavallisesti regrssiomalleja käytetään sellaisissa tilanteissa, joissa selittäjiksi on tarjolla useita *selittäjäkandidaatteja* ja tutkimus kohdistuu siihen, mistä kandidaateista selitettävä muuttuja Y todellakin riippuu.[1] Malliin pyritään valitsemaan ne selittäjät, jotka auttavat selittämään vastemuuttujan Y arvojen vaihtelua kerroinkohtaisten t-testien avulla. Selittäjien valinta voidaan tehdä *askeltaen*. Käytössä on useita erilaisia *askellusstrategioita*. Seuraavassa kuvataan kahta yleisimmin käytettyä strategiaa, joita kutsutaan suuntansa takia *alaspäiseksi* ja *ylöspäiseksi askellukseksi*. Alaspäisessä askelluksessa mallista poistetaan merkityksettömiä selittäjiä yksi kerrallaan, kunnes saadaan malli, jossa kaikki selittäjät ovat merkityksellisiä. Ylöspäisessä askelluksessa malliin lisätään merkityksellisiä selittäjiä yksi kerrallaan, kunnes saadaan malli, jonka ulkopuolelle jätettyjen selittäjäkandidaattien joukossa ei ole enää merkityksellisiä selittäjiä.

ALASPÄINEN ASKELLUS

Lähdetään liikkeelle mallista, jossa ovat mukana kaikki selittäjäkandidaatit ja estimoidaan mallin kertoimet. Poistetaan mallista turhat selittäjät vaiheittain siten, että jokaisessa vaiheessa poistetaan tilastollisesti *merkityksettömistä* selittäjistä se, jota vastaava t-arvo on itseisarvoltaan *pienin*. Jokaisen poiston jälkeen malli estimoidaan uudelleen. Tätä jatketaan kunnes on saatu malli, jossa kaikki selittäjät ovat tilastollisesti merkitseviä. Näin saadulle mallille tehdään lopuksi kertoimien merkin ja suuruuden järkevyystarkastelu.

YLÖSPÄINEN ASKELLUS

Lähdetään liikkeelle mallista, jossa on mukana pelkkä vakiotermi. Estimoidaan kaikki mahdolliset *yhden* selittäjän mallit ja valitaan malliin ensimmäisenä *merkityksellisistä* selittäjäkandidaateista se, jota vastaava t-arvo on itseisarvoltaan *suurin*. Estimoidaan tämän jälkeen kaikki mahdolliset *kahden* selittäjän mallit, joissa on mukana ensimmäisenä valittu selittäjä ja valitaan malliin toisena merkityksellisistä selittäjäkandidaateista se, jota vastaava t-arvo on itseisarvoltaan suurin. Tehdään sama kaikille mahdollisille kolmen selittäjän malleille, joissa on mukana kaksi ensimmäisenä valittua selittäjää. Tätä jatketaan kunnes on saatu malli, jonka ulkopuolelle jätetyistä selittäjistä yksikään ei ole merkityksellinen. Näin saadulle mallille tehdään lopuksi kertoimien merkin ja suuruuden järkevyystarkastelu.

Ylöspäisen askelluksen tuloksena saatavan mallin kaikki selittäjät eivät ole välttämättä tilastollisesti merkityksellisiä. Tämä huomio on johtanut sellaisten

---

[1] Huomaa, että regressiomallin avulla mallitettavan *tilastollisen riippuvuuden* olemassaolosta ei saa tehdä johtopäätöksiä *kausaalisen riippuvuuden* olemassaolosta.

valikointimenetelmien kehittämiseen, joissa sovelletaan jollakin strategialla askellusta molempiin suuntiin. Sivuutamme tällaisten menetelmien kuvaamisen tässä yhteydessä. Todettakoon kuitenkin, että monet tilasto-ohjelmistot sisältävät mahdollisuuden soveltaa *automaattista askellusta*. On kuitenkin syytä tietää, että *automaattisen askelluksen tulos ei ole millään muotoa objektiivinen*.

Todettakoon vielä lopuksi seuraavat seikat: Askelluksen tulos riippuu todennäköisyydestä α. Lisäksi erilaisten askellusstrategioiden käyttö saattaa johtaa erilaisiin malleihin. Askelluksen tuloksena syntyvälle mallille on aina tehtävä kertoimia koskevat järkevyystarkistukset ja standardioletuksien vaatimat diagnostiset tarkistukset. Jos malli havaitaan näissä tarkistuksissa joiltakin osiltaan puutteelliseksi, mallia saatetaan joutua korjaamaan monellakin tavalla. Tämä merkitsee sitä, että hyvän regressiomallin rakentaminen on vaativa tehtävä, johon täytyy tavallisesti uhrata paljon työtä ja aikaa.

ESIMERKKI 1.

Kulutustutkimus on eräs empiirisen taloustieteen osa-alueista. Yksityiset kulutusmenot voidaan jakaa sellaisiin ryhmiin kuten kulutusmenot ruokaan, vaatteisiin, kestokulutushyödykkeisiin, liikennevälineisiin jne. Kulutusmenoja voidaan kussakin ryhmässä selittää esimerkiksi hinnoilla ja käytettävissä olevilla tuloilla.

Seuraavassa tarkastellaan kulutusmenoja kotitalouden tekstiileihin. Kotitalouden kokonaiskulutusmenoista kuluu kotitalouden tekstiilien ostamiseen keskimäärin noin 1%.

Olkoon mallina

$$\log(q_j) = \beta_0 + \beta_1 \log(p_j) + \beta_2 \log(Q_j) + \varepsilon_j,$$

jossa

q_j = kulutusmenot kotitalouden tekstiileihin (kiinteisiin eli reaalisiin hintoihin) per capita (mk),

p_j = kotitalouden tekstiilien reaalihinta (mk),

Q_j = kokonaiskulutusmenot (kiinteisiin eli reaaliisiin hintoihin) per capita (mk).

Oletetaan lisäksi, että jäännöstermit ε_j ovat riippumattomia ja normaalisti jakautuneita kaikille j:

$$\varepsilon_j \sim N(0, \sigma^2).$$

Kokonaiskulutusmenot edustavat mallissa tuloja. Huomaa, että sekä hintamuuttuja että kokonaiskulutusmenot ovat satunnaismuuttujia. Oletetaan, että jäännöstermistä ε_j voidaan tehdä sellaiset *muunnetut* oletukset, että selitettävän muuttujan *ehdollinen* odotusarvo hintojen ja kokonaiskulutusmenojen suhteen on muotoa

$$E(\log(q_j)) = \beta_0 + \beta_1 \log(p_j) + \beta_2 \log(Q_j).$$

Miksi mallin muuttujat on logaritmoitu? Tämä johtuu siitä, että tällöin regressiokertoimet voidaan tulkita ns. *joustoiksi*:

β_1 on *hintajousto*:

Kerroin kuvaa *kuinka monta prosenttia* kulutusmenot kotitalouden tekstiileihin muuttuvat, kun niiden hinta kasvaa 1%:n (ja kokonaistulot pysyvät ennallaan).

β_2 on *kokonaiskulutusmenojen jousto* (tulojousto):

Kerroin kuvaa *kuinka monta prosenttia* kulutusmenot kotitalouden tekstiileihin muuttuvat, kun tulot kasvavat 1%:n (ja tekstiilien hinta pysyy ennallaan).

Estimointitulokset perustuvat Tilastokeskuksen tuottamiin aikasarjoihin vuosilta 1950 — 1978 (n =29). Kulutusmenot ja hinnat on määrätty vuoden 1975 rahassa.

Estimointitulokset voidaan esittää seuraavana taulukkona:

Kerroin	Estimaatti	Hajonta
β_0	−10.3	1.72
β_1	0.136	0.190
β_2	1.55	0.0967

Mallin selitysaste on

$$R^2 = 0.980.$$

Kerroinestimaateista voidaan tehdä seuraavat johtopäätökset:

Jos tekstiilien hinta nousee 1%:n, tekstiilien kulutus nousee 0.136%.

Jos kulutusmenot nousevat 1%:n, tekstiilien kulutus nousee 1.55%.

Hinnan vaikutusta koskeva johtopäätös on järjen vastainen, mutta onneksi hinta ei ole tilastollisesti merkitsevä kulutusta määräävä tekijä, kuten alla nähdään.

Tarkastellaan ensin nollahypoteesia

$$\mathrm{H}_0:\quad \beta_1 = \beta_2 = 0.$$

Jos nollahypoteesi pätee, vastemuuttuja $\log(q_j)$ ei riipu lineaarisesti kummastakaan selittäjästä.

Testisuure tälle nollahypoteesille on

$$\begin{aligned} F &= \frac{n-p-1}{p}\cdot\frac{R^2}{1-R^2} \\ &= \frac{29-2-1}{2}\cdot\frac{0.980}{1-0.980} \\ &\approx 637. \end{aligned}$$

Jos nollahypoteesi pätee, testisuure F on jakautunut kuten F-jakauma 2:lla ja 26:lla vapausasteella. Testisuureen arvoa vastaava P-arvo on 0 kaikilla tavanomaisilla lukutarkkuuksilla, ts.

$$\mathrm{P}(F > 637) \approx 0.$$

Nollahypoteesi siis hylätään: Vaste riippuu molemmista selittäjistä yhdessä.

Tutkitaan seuraavaksi vasteen riippuvuutta yksittäisistä selittäjistä.

Tarkastellaan ensin nollahypoteesia

$$\mathrm{H}_{01}:\quad \beta_1 = 0.$$

Testisuure tälle nollahypoteesille on

$$\begin{aligned} t_1 &= \frac{b_1}{s_{b_1}} \\ &= \frac{0.136}{0.190} \\ &\approx 0.716. \end{aligned}$$

Jos nollahypoteesi pätee, testisuure t_1 on jakautunut kuten t-jakauma 26:lla vapausasteella. Testisuureen arvoa vastaava P-arvo on 0.24. Siten nollahypoteesi voidaan jättää voimaan: Vaste ei riipu hintamuuttujasta.

Tarkastellaan toiseksi nollahypoteesia

$$H_{02}: \quad \beta_2 = 0.$$

Testisuure tälle nollahypoteesille on

$$\begin{aligned} t_2 &= \frac{b_2}{s_{b_2}} \\ &= \frac{1.55}{0.0967} \\ &\approx 16.0. \end{aligned}$$

Jos nollahypoteesi pätee, testisuure t_2 on jakautunut kuten t-jakauma 26:lla vapausasteella. Testisuureen arvoa vastaava P-arvo on 0 kaikilla tavallisilla lukutarkkuuksilla. Siten nollahypoteesi voidaan hylätä: Vaste riippuu tulomuuttujasta.

Yllä esitetyistä estimointituloksista voidaan siis päätellä, että kulutusmenot kotitalouden tekstiileihin riippuvat pelkästään käytettävissä olevista tuloista. Jos malli estimoidaan uudelleen muodossa

$$\log(q_j) = \beta_0 + \beta_2 \log(Q_j) + \varepsilon_j,$$

saadaan seuraavat estimointitulokset:

Kerroin	Estimaatti	Hajonta
β_0	–9.05	0.364
β_2	1.48	0.0407

Mallin selitysaste on

$$R^2 = 0.980.$$

Huomaa, että uudelleenestimoinnin tuloksena jäljelle jääneiden regressiokertoimien estimaatit muuttuivat. Jäljelle jääneiden regressiokertoimien estimaatit muuttuvatkin tavallisesti aina poistettiinpa mallista tai lisättiinpä malliin selittäjiä.

Tarkastellaan nyt nollahypoteesia

$$H_{02}: \quad \beta_2 = 0.$$

Testisuure tälle nollahypoteesille on

$$
\begin{aligned}
t_2 &= \frac{b_2}{s_{b_2}} \\
&= \frac{1.48}{0.0407} \\
&\approx 36.4.
\end{aligned}
$$

Jos nollahypoteesi pätee, testisuure t_2 on jakautunut kuten t-jakauma 27:llä vapausasteella. Testisuureen arvoa vastaava P-arvo on 0 kaikilla tavallisilla lukutarkkuuksilla. Siten nollahypoteesi voidaan hylätä: Vaste riippuu tulomuuttujasta.

Huomaa, että tässä tapauksessa ei ole tarvetta tehdä F-testiä yleishypoteesille, koska testin tulos on sama kuin ylläkuvatun t-testin. Tämä johtuu siitä, että yhden selittäjän tapauksessa

$$F = t^2.$$

Todettakoon vielä, että regressiomallia koskevat estimointitulokset esitetään usein seuraavassa muodossa:

$$
\begin{array}{lll}
\log(q_j) = \underset{(0.364)}{-9.05} + \underset{(0.0407)}{1.48}\cdot\log(Q_j) & \qquad & R^2 = 0.980
\end{array}
$$

Kerroinestimaattien alapuolella on ilmoitettu vastaavat hajonnat. Toinen paljon käytetty tapa on ilmoittaa kerroinestimaattien alla niitä vastaavat t-arvot.

95%:n luottamusväli kertoimelle β_2 on muotoa

$$1.48 \pm z\cdot 0.0407,$$

jossa z valitaan t-jakauman taulukosta niin, että

$$P(-z \le t \le +z) = 0.95,$$

jossa t-jakautuneen satunnaismuuttujan t vapausasteiden luku on 27. t-jakauman taulukoista saadaan

$$z = 1.703.$$

95%:n luottamusväli kertoimelle β_2 saa siis muodon

$$1.48 \pm 0.069$$

eli

$$[1.41,\ 1.55].$$

Olemme käyttäneet sekä testeissä, että luottamusvälin konstruktiossa t-jakaumaa. t-jakauman käyttö on luvallista, koska jäännöstermistä ε_t tehtiin normaalisuusoletus. Jos normaalisuusoletus ei päde, joudutaan vetoamaan suurten otosten approksimaatiotulokseen ja turvautumaan normaalijakauman taulukoihin. Näin saatava luottamusväli on vain approksimatiivisesti 95%:n luottamusväli. ●

On syytä huomata, että edellisessä esimerkissä kiinnitettiin huomiota ainoastaan *regressiokertoimia koskeviin testeihin*. Näiden testien käyttö on tarkasti ottaen luvallista vain, jos mallista tehdyt standardioletukset 1—5 (ja 6) pätevät. Siksi regressioanalyysiin pitää aina liittää *diagnostiset testit* oletusten tarkistamiseksi. Sivuutamme kuitenkin sellaiset testit tässä yhteydessä.

5.2.4 YHDEN SELITTÄJÄN MALLI

Käsittelemme seuraavassa *yhden selittäjän lineaarisen regressiomallin* estimointiin liittyviä kaavoja.

Olkoon

$$Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j, \qquad \varepsilon_j \sim N(0, \sigma^2), \qquad j = 1,2,\ldots,n.$$

Parametrien β_0 ja β_1 PNS-estimaattorit saadaan kaavoista

$$b_0 = \bar{Y} - b_1\bar{x},$$

$$b_1 = r\frac{s_Y}{s_x},$$

jossa

$$\bar{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j$$

on vastemuuttujan havaintoarvojen aritmeettinen keskiarvo,

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

on selittäjän havaintoarvojen aritmeettinen keskiarvo,

$$s_Y^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \bar{Y})^2$$

on vastemuuttujan havaintoarvojen otosvarianssi,

$$s_x^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2$$

on selittäjän havaintoarvojen otosvarianssi ja

$$r = \frac{s_{Yx}}{s_Y s_x}$$

on vastemuuttujan ja selittäjän havaintoarvojen otoskorrelaatiokerroin, jossa

$$s_{Yx} = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \bar{Y})(x_j - \bar{x})$$

on vastemuuttujan ja selittäjän havaintoarvojen otoskovarianssi.

Selitysaste R^2 yhtyy yhden selittäjän mallin tapauksessa vastemuuttujan ja selittäjän havaintoarvojen korrelaatiokertoimen neliöön:

$$R^2 = r^2.$$

Jäännöstermin varianssin σ^2 harhaton estimaattori saadaan esimerkiksi kaavasta

$$s^2 = \frac{n-1}{n-2}(1-r^2)s_Y^2.$$

Kertoimen b_1 varianssin harhaton estimaattori saadaan kaavasta

$$s_{b_1}^2 = \frac{s^2}{(n-1)s_x^2}.$$

Olkoon nollahypoteesina

$$H_0: \quad \beta_1 = 0.$$

Tavanomainen t-testisuure nollahypoteesille H_0 on muotoa

$$t = \frac{b_1}{s / \sqrt{(n-1)s_x^2}}.$$

Jos nollahypoteesi pätee ja jäännöstermi ε_j on normaalinen, testisuure t noudattaa t-jakaumaa vapausastein n–2. Jos jäännöstermistä ei voida tehdä normaalisuusoletusta, joudutaan soveltamaan suurten otosten tulosta, jonka mukaan testisuure t noudattaa nollahypoteesin pätiessä suurissa otoksissa approksimatiivisesti standardoitua normaalijakaumaa N(0,1).

Tässä tapauksessa F-testisuure nollahypoteesille H_0 on muotoa

$$F = \frac{b_1^2}{s^2 / ((n-1)s_x^2)}.$$

Jos jäännöstermi on normaalisti jakautunut, testisuure F on jakautunut kuten F-jakauma vapausastein 1 ja n–2. F-testisuureen käyttö johtaa tässä tapauksessa aivan samaan tulokseen kuin t-testin käyttö normaalijakautuneen jäännöstermin tapauksessa. Tämä nähdään siitä, että tässä tapauksessa

$$F = t^2$$

ja seuraavasta tuloksesta: Jos satunnaismuuttuja noudattaa t-jakaumaa vapausastein f, niin sen neliö noudattaa F-jakaumaa vapausastein 1 ja f.

Yllä esitetty F-testisuure voidaan kirjoittaa myös tutumpaan muotoon

$$F = (n-2)\frac{r^2}{1-r^2},$$

kun käytetään apuna yhtälöä $R^2 = r^2$. On syytä huomata, että tämän testisuureen neliöjuuri on t-testisuure nollahypoteesille

$$H_0: \quad \rho = 0,$$

jossa ρ on selitettävän muuttujan Y ja selittäjän X välinen korrelaatiokerroin perusjoukossa.

Edellä esitetystä nähdään, että korreloimattomuushypoteesi $\rho = 0$ ja hypoteesi $\beta_1 = 0$ ovat yhden selittäjän regressiomallin tapauksessa yhtäpitäviä. Edellä esitetystä nähdään myös se, että testisuureet näille yhtäpitäville hypoteeseille voidaan formuloida monilla yhtäpitävillä tavoilla.

ESIMERKKI 1.

Menestyminen opinnoissa saattaa vaikuttaa vastavalmistuneen alkupalkkaan. Asiaa tutkittiin USA:ssa poimimalla eräästä korkeakoulusta valmistuneiden joukosta yksinkertainen satunnaisotos, jonka koko oli 15.

Otokseen poimituilta kysyttiin heidän arvosanapisteidensä keskiarvoa x ja alkupalkkaansa Y (vuosipalkka tuhansina dollareina). Oletetaan, että regressiomalli alkupalkan riippuvuudelle arvosanapisteiden keskiarvosta on muotoa

$$Y_j = \beta_0 + \beta_1 x_j + \varepsilon_j, \qquad j = 1,2,\ldots,15,$$

jossa jäännöstermistä ε_j voidaan tehdä tavanomaiset oletukset.

Otosta kuvaavat perustunnusluvut olivat

$$\bar{x} = 3.04, \qquad \bar{Y} = 18.05,$$

$$s_x^2 = 0.063, \qquad s_Y^2 = 5.81,$$

$$r = 0.848.$$

Näistä tunnusluvuista saadaan seuraavat estimaatit regressiokertoimille:

$$b_1 = r\frac{s_Y}{s_x} = 8.12,$$

$$b_0 = \bar{Y} - b_1\bar{x} = -6.63.$$

Selitysasteeksi saadaan

$$R^2 = r^2 = 0.718.$$

Siten malli selittää 71.8% selitettävän muuttujan arvojen vaihtelusta.

Testisuure korreloimattomuushypoteesille

$$H_0: \quad \rho = 0$$

saa arvon

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = 5.77.$$

Vastaava P-arvo on 0.000033, joten nollahypoteesi voidaan hylätä: Menestyminen opinnoissa parantaa alkupalkkaa.

Jäännösvarianssin estimaatiksi saadaan

$$s^2 = \frac{n-1}{n-2}(1-r^2)s_Y^2 = 1.76.$$

Estimaattorin b_1 varianssin estimaattori on

$$s_{b_1}^2 = \frac{s^2}{(n-1)s_x^2} = 2.00.$$

Siten testisuureen arvoksi hypoteesille

$$H_0: \quad \beta_1 = 0$$

saadaan

$$t = \frac{b_1}{s_{b_1}} = 5.76.$$

Tämä testisuureen arvo on käytännössä sama kuin ylläesitetyn korreloimattomuushypoteesin testisuureen arvo kuten pitääkin (ero johtuu käytetystä laskutarkkuudesta).

Neliösummista

$$SST = (n-1)s_Y^2 = 81.3,$$

$$SSE = (n-2)s^2 = 22.9,$$

$$SSM = SST - SSE = 58.4,$$

voidaan edellä esitetyn nollahypoteesin $\beta_1 = 0$ t-testisuureelle muodostaa sen kanssa yhtäpitävä F-testisuure

$$F = (n-2)\frac{SSM}{SSE} = 33.2.$$

Huomaa, että (laskutarkkuudesta johtuvaa eroa lukuunottamatta)

$$F = t^2.$$

Regressioanalyysin tulokset voidaan esittää esimerkiksi seuraavassa muodossa:

$$\underset{}{Y_j = -6.63} + \underset{}{8.12 \cdot x_j} \qquad R^2 = 0.718$$

$$(4.30) \quad (1.41)$$

Sulkuihin on merkitty kerroinestimaattoreiden hajontojen estimaattorit.

Estimointituloksista voidaan tehdä seuraava johtopäätös: Jos kahdella otokseen poimitun henkilön arvosanapisteiden keskiarvot eroavat yhdellä pisteellä, heidän alkupalkkansa eroavat 8,120 dollaria. ●

Tässäkin tapauksessa analyysiä olisi täydennettävä mallin oletuksia koskevilla *diagnostisilla testeillä*.

6. VARIANSSIANALYYSI

6.1 JOHDANTO

Testausta käsitelleen luvun johdannossa käsiteltiin seuraavaa testausasetelmaa: Oletetaan, että perusjoukko voidaan jakaa kahteen ryhmään ja, että kiinnostus kohdistuu jonkin satunnaismuuttujan odotusarvoon näissä ryhmissä. Päämääränä on muodostaa testi odotusarvojen yhtäsuuruudelle. Jos odotusarvot olisivat samoja, voitaisiin ryhmiä tarkastella tutkimuksen kohteena olevan satunnaismuuttujan kannalta yhtenä ryhmänä. Nollahypoteesille esitettiin luvussa 4 testi, joka perustuu ryhmäkohtaisten otoskeskiarvojen vertailuun ja on muodoltaan t-testi.

Varianssianalyysi on tämän testausasetelman yleistys tilanteeseen, jossa perusjoukko voidaan jakaa kahteen tai *useampaan* ryhmään. Nollahypoteesina on se, että tarkastelun kohteena olevan satunnaismuuttujan odotusarvot ovat eri ryhmissä yhtä suuria. Jos nollahypoteesi jää voimaan, perusjoukkoa voidaan tarkastella ko. satunnaismuuttujan kannalta yhtenä ryhmänä.

Odotusarvojen vertailu perustuu ryhmäkohtaisten otoskeskiarvojen vertailuun. *Varianssianalyysilla on harhaanjohtava nimitys*, joka johtuu siitä, että vertailussa käytettävä F-testisuure voidaan tulkita kahdella eri tavalla määrätyn varianssin vertailutestiksi.

Perusjoukko voidaan jakaa ryhmiin usealla eri periaatteella. Jos jako perustuu *yhteen* luokittelevaan tekijään, puhutaan *yksisuuntaisesta varianssianalyysista*. Jos luokittelevia tekijöitä on m, puhutaan *m-suuntaisesta varianssianalyysista*. Seuraavassa rajoitutaan pelkästään yksi- ja kaksisuuntaisten varianssianalyysien tarkasteluun.

6.2 YKSISUUNTAINEN VARIANSSIANALYYSI

6.2.1 YKSISUUNTAISEN VARIANSSIANALYYSIN PERUSASETELMA

Oletetaan, että tutkimuksen kohteena oleva perusjoukko voidaan jakaa k ryhmään. Oletetaan lisäksi, että kustakin ryhmästä $i = 1,2,\ldots n$ on poimittu toisistaan riippumattomat yksinkertaiset satunnaisotokset, joiden koot ovat $n_1, n_2, \ldots, n_k$. Olkoon ryhmään i kuuluva j. havainto Y_{ij}. Oletetaan, että havainnon Y_{ij} odotusarvo on

$$E(Y_{ij}) = \mu_i, \qquad j = 1,2,\ldots,n_i, \qquad i = 1,2,\ldots,k$$

ja varianssi on

$$D^2(Y_{ij}) = \sigma^2, \qquad j = 1,2,\ldots,n_i, \qquad i = 1,2,\ldots,k.$$

Nämä oletukset merkitsevät sitä, että *kaikilla samaan ryhmään i kuuluvilla havainnoilla on sama odotusarvo* μ_i. Sen sijaan odotusarvot eri ryhmiin kuuluvilla havainnoilla voivat erota toisistaan. Kaikilla havainnoilla oletetaan kuitenkin olevan ryhmästä riippumatta sama varianssi σ^2. Edellä esitettyjen oletuksien lisäksi tehdään oletus, että havainnot ovat jakautuneet kuten normaalijakauma ym. parametrein.

Kiinnostuksen kohteena oleva nollahypoteesi on muotoa

$$H_0: \quad \mu_1 = \mu_2 = \ldots = \mu_k = \mu,$$

jossa μ on havaintojen yhteinen odotusarvo, *jos nollahypoteesi pätee*. Jos siis nollahypoteesi pätee, noudattavat havainnot eri ryhmissä samaa jakaumaa ja ryhmät voidaan muuttujaa Y koskevissa tarkasteluissa yhdistää yhdeksi ryhmäksi. Sen sijaan, jos nollahypoteesi ei päde, muuttujan Y odotettavissa olevat arvot eri ryhmissä eroavat toisistaan. Tällöin tieto siitä, mihin ryhmään havainto kuuluu, antaa tietoa muuttujan Y keskimääräisestä arvosta. Seuraavassa tarkastellaan testiä ym. nollahypoteesille.

6.2.2 TESTI ODOTUSARVOJEN YHTÄSUURUUDELLE

Havainnot Y_{ij} voidaan ryhmitellä seuraavalla tavalla:

Ryhmä 1: $\quad Y_{11}, Y_{12}, \ldots, Y_{1n_1},$

Ryhmä 2: $\quad Y_{21}, Y_{22}, \ldots, Y_{2n_2},$

...

Ryhmä k: $\quad Y_{k1}, Y_{k2}, \ldots, Y_{kn_k}.$

Ryhmäkohtaiset otoskeskiarvot $\bar{Y}_i$, $i = 1,2,\ldots,k$ voidaan määrätä kaavoilla

Ryhmä 1: $\quad \bar{Y}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1} Y_{1j},$

Ryhmä 2: $\quad \bar{Y}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} Y_{2j},$

...

Ryhmä k: $\quad \bar{Y}_k = \frac{1}{n_k}\sum_{j=1}^{n_k} Y_{kj}.$

Ryhmät yhdistämällä saatava *kokonaiskeskiarvo* voidaan määrätä kaavalla

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} Y_{ij},$$

jossa

$$n = n_1 + n_2 + \cdots + n_k$$

on havaintojen kokonaislukumäärä.

Kirjoitetaan identiteetti

$$Y_{ij} - \bar{Y} = (\bar{Y}_i - \bar{Y}) + (Y_{ij} - \bar{Y}_i),$$

jossa

$Y_{ij} - \overline{Y}$ = havaintoarvon poikkeama kokonaiskeskiarvosta,

$\overline{Y}_i - \overline{Y}$ = ryhmäkeskiarvon poikkeama kokonaiskeskiarvosta,

$Y_{ij} - \overline{Y}_i$ = havaintoarvon poikkeama ryhmäkeskiarvosta.

Testi nollahypoteesille H_0 ryhmäkohtaisten odotusarvojen samuudesta voidaan perustaa näille poikkeamille. Jos nollahypoteesi pätee, ryhmäkeskiarvojen poikkeamat kokonaiskeskiarvosta eli erotukset

$$\overline{Y}_i - \overline{Y}$$

eivät voi olla kovin suuria.

Muodostetaan yllä esitettyä hajoitelmaa vastaava *neliösummahajoitelma*:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\overline{Y}_i - \overline{Y})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2 .$$

Neliösummahajoitelmassa havaintoarvojen havaintojen vaihtelua kokonaiskeskiarvon ympärillä kuvaava *kokonaisneliösumma*

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y})^2$$

on jaettu kahteen osaan (SST = total sum of squares of):

Ryhmäneliösumma

$$SSG = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\overline{Y}_i - \overline{Y})^2$$

$$= \sum_{i=1}^{k} n_i(\overline{Y}_i - \overline{Y})^2$$

kuvaa ryhmäkeskiarvojen vaihtelua kokonaiskeskiarvon ympärillä (SSG = sum of squares due to groups).

Jäännösneliösumma

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2$$

kuvaa havaintojen vaihtelua ryhmäkeskiarvojen ympärillä (SSE = sum of squares due to errors). Yllä esitetty neliösummahajoitelma voidaan esittää em. määritelmiä käyttäen muodossa

$$SST = SSG + SSE.$$

On syytä huomata, että

$$s^2 = \frac{SST}{n-1}$$

on havaintojen kokonaisvaihtelua kuvaava tavanomainen otosvarianssi, joka saadaan, jos ryhmäkohtaiset otokset yhdistetään yhdeksi otokseksi. Vastaavasti havaintojen vaihtelua ryhmässä i kuvaava ryhmäkohtainen otosvarianssi s_i^2 voidaan määritellä jäännösneliösumman SSE lausekkeessa esiintyvien osasummien

$$\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2$$

avulla. Siten

$$s_i^2=\frac{1}{n_i-1}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2.$$

Käyttämällä ryhmäkohtaisten otosvarianssien s_i^2 lausekkeita jäännösneliösumma SSE voidaan kirjoittaa muotoon

$$SSE=\sum_{i=1}^{k}(n_i-1)s_i^2.$$

Testisuure nollahypoteesille H_0 ryhmäkohtaisten odotusarvojen yhtäsuuruudesta voidaan perustaa em. neliösummahajoitelmalle. Siinä kokonaisneliösumma *SST* on jaettu ryhmäneliösumman *SSG* ja jäännösneliösumman *SSE* summaksi. Varianssianalyysin tulkinnan kannalta on keskeistä se, että ryhmäneliösumma *SSG* kuvastaa *ryhmien välistä vaihtelua*, kun taas jäännösneliösumma *SSE* kuvastaa *ryhmien sisäistä vaihtelua*. Koska em. neliösummahajoitelman mukaan

$$SST = SSG + SSE,$$

voidaan tehdä seuraava johtopäätös:

Jos ryhmien välinen vaihtelu on suurta verrattuna ryhmien sisäiseen vaihteluun, nollahypoteesi on syytä asettaa epäilyksen alaiseksi.

Muodostetaan testisuure

$$F = \frac{n-k}{k-1} \cdot \frac{SSG}{SSE}.$$

Jos nollahypoteesi H_0 ryhmäkohtaisten odotusarvojen yhtäsuuruudesta *pätee*, testisuure F noudattaa F-jakaumaa vapausastein $k-1$ ja $n-k$. Nollahypoteesin pätiessä testisuureen F odotusarvo on

$$\mathrm{E}(F) = \frac{n-k}{n-k-2} \approx 1.$$

Jos testisuureen F arvo on kyllin paljon odotusarvoaan suurempi, hylätään nollahypoteesi. Jos nollahypoteesi hylätään, hyväksytään vaihtoehtoinen hypoteesi, jonka mukaan ryhmäkohtaisissa odotusarvoissa on eroja.

Edellä esitetyn testin tulokset esitetään tavallisesti ns. *varianssianalyysitaulukkona*:

Taulukko 1. Varianssianalyysitaulukko 1-suuntaisessa varianssianalyysissa.

Vaihtelun lähde	Neliö-summa	Vapaus-asteet	Varianssi-estimaattori	F-testisuure
Ryhmien välinen vaihtelu	*SSG*	$k-1$	$\frac{SSG}{k-1}$	$F = \frac{n-k}{k-1} \cdot \frac{SSG}{SSE}$
Ryhmien sisäinen vaihtelu	*SSE*	$n-k$	$\frac{SSE}{n-k}$	
Kokonais-vaihtelu	*SST*	$n-1$	$\frac{SST}{n-1}$	

Varianssianalyysitaulukko perustuu siihen, että taulukon neliösummat toteuttavat em. neliösummahajoitelman

$$SST = SSG + SSE,$$

ja taulukon vapausasteet toteuttavat hajoitelman

$$n-1 = (k-1) + (n-k).$$

Kokonaisvarianssi

$$s^2 = \frac{SST}{n-1}$$

on aina perusjoukon varianssin σ^2 harhaton estimaattori. Sen sijaan

$$\mathrm{E}\left(\frac{SSG}{k-1}\right) = \sigma^2$$

vain, jos nollahypoteesi pätee.

VARIANSSIANALYYSI JA YLEINEN LINEAARINEN MALLI

Tilastotieteen kannalta varianssianalyysi on erikoistapaus *yleisestä lineaarisesta mallista*. Yleinen lineaarinen malli on lineaarinen regressiomalli sopivalla tulkinnalla varustettuna. Yksisuuntaista varianssianalyysiasetelmaa kuvaava yleinen lineaarinen malli voidaan esittää muodossa

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \qquad j = 1,2,\ldots,n_i, \qquad i = 1,2,\ldots,k,$$

jossa

Y_{ij} = Y-muuttujan j. havaintoarvo ryhmässä i,

μ_i = Y-muuttujan j. havaintoarvon odotusarvo ryhmässä i,

ε_{ij} = jäännöstermi.

Jäännöstermeistä ε_{ij} tehdään seuraavat kaksi oletusta:

1. Kaikki ε_{ij}:t noudattavat normaalijakaumaa $N(0,\sigma^2)$.
2. Kaikki ε_{ij}:t ovat keskenään korreloimattomia.

Yllä esitetty malli vastaa sellaista lineaarista regressiomallia, jossa selitettävänä muuttujana on Y ja selittäjinä ovat *indikaattorimuuttujat* I_i, jotka kuvaavat sitä, mihin ryhmään havainto Y_{ij} kuuluu. Indikaattorimuuttujat I_i, $i = 1,2,\ldots,k$ voidaan määritellä seuraavalla kaavalla:

$$I_{ij} = \begin{cases} 1, \text{ jos } Y_{ij} \text{ kuuluu ryhmään } i \\ 0, \text{ jos } Y_{ij} \text{ ei kuulu ryhmään } i \end{cases}$$

Mallin regressiokertoimina ovat ryhmäkohtaiset odotusarvot μ_i. 1-suuntaista varianssianalyysia vastaava yleinen lineaarinen malli on regressioanalyysimuodossaan seuraava:

$$Y_{ij} = \mu_1 I_{1j} + \mu_2 I_{2j} + \ldots + \mu_k I_{kj} + \varepsilon_{ij}.$$

Voidaan osoittaa, että regressiokertoimien $\mu_1, \mu_2, \ldots, \mu_k$ PNS-estimointi johtaa estimaattoreihin

$$\begin{aligned}\hat{\mu}_i &= \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij} \\ &= \overline{Y}_{i\cdot}\end{aligned}$$

Edelleen voidaan osoittaa, että nollahypoteesia H_0 regressiokertoimien μ_i samuudesta eli hypoteesia

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_k = \mu$$

voidaan testata edellä esitetyllä F-testillä.

Edellä sanottu merkitsee sitä, että 1-suuntainen varianssianalyysi on tilastollisena mallina erikoistapaus yleisestä lineaarisesta mallista. Tämä pätee yleisemminkin: Myös m-suuntainen varianssianalyysi on lineaarisen mallin erikoistapaus. Lineaarisia malleja käsittelevässä kirjallisuudessa varianssianalyysi esitetäänkin yleensä tässä yleisemmässä kehikossa.

6.2.3 SOVELLUS

Odottavien äitien paino lisääntyy selvästi odotusajan viimeisellä kolmanneksella. Painonlisäyksen suuruus vaikuttaa vauvan syntymäpainoon: Jos äidin paino ei lisäänny tarpeeksi, vauva syntyy pienikokoisena ja on usein sairaalloinen. Tällä seikalla on erityistä merkitystä kehitysmaissa.

Seuraavassa on esitetty äitien keskimääräiset painonlisäykset ja lisäysten hajonnat raskausajan viimeisellä kolmanneksella Egyptissä, Keniassa ja Meksikossa. Tiedot perustuivat toisistaan riippuumattomiin yksinkertaisiin satunnaisotoksiin, joiden koot käyvät ilmi taulukosta.

Maa	Otoskoko n_i	Otoskeskiarvo $\overline{Y}_i$	Otoshajonta s_i
Egypti	46	3.7	2.5
Kenia	111	3.1	1.8
Meksiko	52	2.9	1.8

Tässä tapauksessa siis

$$k = 3$$

ja

$$n = n_1 + n_2 + n_3 = 46 + 111 + 52 = 209.$$

Olkoon

$$\mu_i = \text{äidin painonlisäys maassa } i.$$

Varianssianalyysissa testattava nollahypoteesi on muotoa

$$H_0: \quad \mu_1 = \mu_2 = \mu_3 = \mu.$$

Testin tekemiseksi muodostetaan varianssianalyysitaulukko. Varianssianalyysitaulukkoa varten tarvittavat neliösummat voidaan määrätä yo. taulukossa esitetyistä tiedoista seuraavalla tavalla:

Määrätään ensin *kokonaiskeskiarvo*:

$$\begin{aligned}\overline{Y} &= \frac{1}{n}\sum_{i=1}^{k} n_i\overline{Y}_i \\ &= \frac{1}{209}(46\cdot 3.7 + 111\cdot 3.1 + 52\cdot 2.9) \\ &= 3.2.\end{aligned}$$

Ryhmien välistä vaihtelua kuvaava neliösumma saadaan ryhmäkeskiarvoista ja kokonaiskeskiarvosta seuraavalla laskutoimituksella:

$$
\begin{aligned}
SSG &= \sum_{i=1}^{n} n_i(\overline{Y}_i - \overline{Y})^2 \\
&= 46 \cdot (3.7 - 3.2)^2 + 111 \cdot (3.1 - 3.2)^2 + 52 \cdot (2.9 - 3.2)^2 \\
&= 17.3.
\end{aligned}
$$

Ryhmien sisäistä vaihtelua kuvaava neliösumma saadaan ryhmävariansseista seuraavalla laskutoimituksella:

$$
\begin{aligned}
SSE &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2 \\
&= \sum_{i=1}^{k} (n_i - 1)s_i^2 \\
&= 45 \cdot 2.5^2 + 110 \cdot 1.8^2 + 51 \cdot 1.8^2 \\
&= 802.9.
\end{aligned}
$$

Näiden tulosten perusteella *kokonaisvaihteluksi* saadaan

$$SST = SSG + SSE = 17.3 + 802.9 = 820.2.$$

Nollahypoteesille H_0 tarkoitetun F-testisuureen arvoksi saadaan yllä johdettujen tietojen perusteella

$$
\begin{aligned}
F &= \frac{n-k}{k-1} \cdot \frac{SSG}{SSE} \\
&= \frac{206}{2} \cdot \frac{17.3}{802.9} \\
&= 2.22.
\end{aligned}
$$

Jos nollahypoteesi odottavien äitien painonlisäysten keskimääräisestä yhtäsuuruudesta pätee, testisuure F on jakautunut kuten F-jakauma vapausastein 2 ja 206. Koska 5%:n merkitsevyystasoa vastaava kriittinen arvo on 3.00, voidaan nollahypoteesi jättää voimaan 5%:n merkitsevyystasolla: Äitien odotettavissa olevat painonlisäykset eivät eroa toisistaan tutkimuksen kohteena olleissa kolmessa maassa.

Varianssianalyysin tulokset voidaan esittää seuraavana varianssianalyysitaulukkona:

Vaihtelun lähde	Neliösumma	Vapaus-asteet	Varianssi-estimaatti	F-testisuure
Ryhmien välinen vaihtelu	17.3	2	8.65	2.22
Ryhmien sisäinen vaihtelu	802.9	206	3.90	
Kokonais-vaihtelu	820.2	208	3.94	

Muutetaan esimerkissä esitettyjä lukuja niin, että ryhmäkohtaisia otoshajontoja pienennetään hieman, mutta ryhmäkohtaiset otoskeskiarvot pidetään ennallaan.

Seuraavassa on esitetty äitien keskimääräiset painonlisäykset ja lisäysten hajonnat raskausajan viimeisellä kolmanneksella *muunnetun* aineiston tapauksessa:

Maa	Otoskoko n_i	Otoskeskiarvo $\overline{Y}_i$	Otoshajonta s_i
Egypti	46	3.7	1.8
Kenia	111	3.1	1.1
Meksiko	52	2.9	1.1

Miten varianssianalyysin tulokset muuttuvat?

Kokonaiskeskiarvo säilyy ennallaan:

$$\overline{Y} = 3.2.$$

Myös *ryhmien välistä vaihtelua kuvaava neliösumma* säilyy ennallaan:

$$SSG = 17.3.$$

Sen sijaan *ryhmien sisäistä vaihtelua kuvaava neliösumma* pienenee, koska ryhmäkohtaiset otoshajonnat ovat pienentyneet:

$$\begin{aligned} SSE &= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \overline{Y}_i)^2 \\ &= \sum_{i=1}^{k}(n_i - 1)s_i^2 \\ &= 45 \cdot 1.8^2 + 110 \cdot 1.1^2 + 51 \cdot 1.1^2 \\ &= 340.6. \end{aligned}$$

Näiden tulosten perusteella *kokonaisvaihteluksi* saadaan nyt

$$SST = SSG + SSE = 17.3 + 340.6 = 357.9.$$

Nollahypoteesille takoitettu F-testisuureen arvo kasvaa, koska ryhmien sisäistä vaihtelua kuvaava neliösumma SSE on pienentynyt:

$$\begin{aligned} F &= \frac{n-k}{k-1} \cdot \frac{SSG}{SSE} \\ &= \frac{206}{2} \cdot \frac{17.3}{340.6} \\ &= 5.24. \end{aligned}$$

Jos nollahypoteesi H_0 odottavien äitien painonlisäysten keskimääräisestä yhtäsuuruudesta *pätee*, testisuure F on jakautunut kuten F-jakauma vapausastein 2 ja 206. Koska 1%:n merkitsevyystasoa vastaava kriittinen arvo on 4.61, nollahypoteesi voidaan nyt hylätä 1%:n merkitsevyystasolla: Äitien odotettavissa olevat painonlisäykset eroavat toisistaan tutkimuksen kohteena olleissa kolmessa maassa.

Tulokset voidaan esittää seuraavana varianssianalyysitaulukkona:

Vaihtelun lähde	Neliösumma	Vapaus-asteet	Varianssi-estimaatti	F-testisuure
Ryhmien välinen vaihtelu	17.3	2	8.65	5.24
Ryhmien sisäinen vaihtelu	340.6	206	1.65	
Kokonais-vaihtelu	357.9	208	1.72	

Verrataanpa saatua tulosta aikaisempaan tulokseen. Mistä johtuu, että nolla-hypoteesi voidaan nyt hylätä, kun aikaisemmin näin ei voitu tehdä? Selityksenä on se, että edellisessä tapauksessa ryhmäkohtainen vaihtelu oli niin suurta, että otoskeskiarvojen väliset erot *peittyivät* tämän vaihtelun alle. Ryhmäkohtaisen vaihtelun pienentäminen nosti esiin otoskeskiarvojen väliset erot.

Varianssianalyysin nimitys selittyykin sillä, että keskiarvojen välisten erojen tilastollisen merkitsevyyden kannalta on ratkaisevaa ryhmäkohtaisten varianssien suuruus ryhmäkohtaisten keskiarvojen välisiin eroihin verrattuna. Tilanne on aivan samanlainen kuin aikaisemmin käsitellyssä kahden perusjoukon odotusarvoja koskevassa t-testissä: Testisuure mittasi havaitun maailmantilan ja nollahypoteesin mukaisen maailmantilan yhteensopivuutta tai etäisyyttä, kun etäisyyden mittaaminen perustettiin havaintojen hajaantuneisuuden sopivaan mittariin.

6.3 KAKSISUUNTAINEN VARIANSSIANALYYSI

6.3.1 JOHDANTO

Yksisuuntaisessa varianssianalyysissa perusjoukko jaetaan kahteen tai useampaan ryhmään *yhden* tekijän suhteen. Tarkoituksena on vertailla johonkin perusjoukon ominaisuuteen liittyvän satunnaismuuttujan odotusarvoa näissä ryhmissä. Testauksen kohteena on nollahypoteesi, jonka mukaan odotusarvo on ryhmissä sama. Nollahypoteesin pätiessä ryhmät voidaan yhdistää tutkimuksen kohteena olevaa perusjoukon ominaisuutta tarkasteltaessa yhdeksi ryhmäksi.

Kaksisuuntaisessa varianssianalyysissa perusjoukko jaetaan ryhmiin *kahden* tekijän suhteen. Nytkin tarkoituksena on vertailla tutkimuksen kohteena olevaan perusjoukon ominaisuuteen liittyvän satunnaismuuttujan odotusarvoa jaon tuloksena syntyvissä ryhmissä. Koska perusjoukko on jaettu ryhmiin *kahden* tekijän suhteen, kohdataan analyysissa sellaisia ongelmia, joita ei esiinny yksisuuntaisessa varianssi-analyysissa. Nämä ongelmat liittyvät tekijöiden mahdolliseen *interaktioon* eli *yhdysvaikutukseen*.

Jos perusjoukko jaetaan ryhmiin *m* tekijän suhteen, puhutaan *m-suuntaisesta varianssianalyysista*. Mitä useamman tekijän suhteen perusjoukko jaetaan ryhmiin, sitä vaativammaksi muuttuu yhdysvaikutusten analysointi. Tässä esityksessä rajoitumme pelkästään kaksisuuntaiseen varianssianalyysiin.

Ennen yhdysvaikutuskäsitteen havainnollistamista tarkastelemme esimerkkejä, joiden tehtävänä on havainnollistaa tutkimusasetelmia, joissa perusjoukko jaetaan ryhmiin kahden selittävän tekijän suhteen. Ensimmäinen esimerkki käsittelee *koetta*, jossa selitettävä tekijää pyritään selittämään kahden selittävän tekijän avulla.

ESIMERKKI 1.

Oletetaan, että tehtävänä on suunnitella koe, jonka päämääränä on selvittää miten eräiden kivennäisaineiden määrät ravinnossa vaikuttavat verenpaineeseen. Kiinnostuksen kohteena olevat kivennäisaineet ovat kalsium ja magnesium. Koe tehdään rotille. Kaikkien rottien ruokavalio on muuten samanlainen, mutta rotat saavat erilaisia määriä mainittuja kivennäisaineita. Tällaiselle kokeelle tyypilliseen tapaan kummankin kivennäisaineen määrälle valitaan kolme tasoa: matala, normaali, korkea.

Kokeen *kohteina* ovat siis rotat ja kokeessa on *vastemuuttujana* rotan verenpaine. Kokeen *selittävinä tekijöinä* ovat kalsiumin ja magnesiumin määrät rotan ravinnossa. Kummallakin selittävällä tekijällä on kolme *tasoa*: matala, normaali, korkea. Tämä merkitsee sitä, että rotat voidaan alistaa $3 \times 3 = 9$ erilaiseen *käsittelyyn*. Rotat jaetaan *satunnaistusta* käyttäen eri käsittelyn kohteiksi joutuviin ryhmiin. Rotan verenpaine *mitataan*, kun rotta on ollut sille arvotulla ruokavaliolla 2 viikkoa.

Seuraava kaavio kuvaa käytettyä koeasetelmaa ja niitä 9 käsittelyryhmää, joihin kokeen kohteeksi valitut rotat arvotaan:

	Kalsiumin määrä		
Magnesiumin määrä	Matala	Normaali	Korkea
Matala	1	2	3
Normaali	4	5	6
Korkea	7	8	9

Esimerkiksi sellaisen rotan, joka joutuu arvonnassa ryhmään 7, ruokavaliossa kalsiumin määrä on normaalia matalampi ja magnesiumin määrä normaalia korkeampi. Jos kuhunkin 9 ryhmään arvotaan 9 rottaa, kokeen kohteiksi joutuu 81 rottaa.

Kokeessa vertaillaan rottien verenpaineita em. kaaviolla kuvatuissa 9 käsittelyryhmässä. ●

Esimerkissä 1 kuvatussa kokeessa syntyvää aineistoa voidaan analysoida 2-suuntaisella varianssianalyysillä. Siinä rottien *keskimääräisiä* verenpaineita verrataan kuvatuissa 9 käsittelyryhmässä toisiinsa. Kalsiumin ja magnesiumin määrien vaikutusta rottien verenpaineeseen voidaan vertailla *erillisinä* tekijöinä käyttämällä 1-suuntaista varianssianalyysiä, kuten edellisessä kappaleessa on selitetty. Selittävän tekijän erillisiä vaikutuksia on tapana kutsua kaksisuuntaisen varianssianalyysin yhetydessä *päävaikutuksiksi*. Jos analyysi kohdistuu pelkästään päävaikutuksiin, jää kuitenkin huomiotta eräs tärkeä tekijä, jolla saattaa olla ratkaiseva vaikutus vastemuuttujan käyttäytymiseen. Tämä tekijä on ko. selittävien tekijöiden *interaktio* eli *yhdysvaikutus*.

Mitä kahden selittävän tekijän yhdysvaikutuksella tarkoitetaan? Tarkastellaan vielä esimerkkiä 1. Oletetaan, että sekä kalsiumin että magnesiumin normaalia suuremmat määrät ruokavaliossa alentavat rottien verenpainetta. Jos molempien kivennäisaineiden normaalia suuremmat määrät alentavat verenpainetta selvästi *enemmän* kuin vain toisen normaalia suurempi määrä, ko. kivennäisaineiden normaalia suuremmilla määrillä on *positiivinen* interaktio. Jos taas molempien kivennäisaineiden normaalia suuremmat määrät alentavat verenpainetta selvästi *vähemmän* kuin vain toisen normaalia suurempi määrä, ko. kivennäisaineiden normaalia suuremmilla määrillä on *negatiivinen* interaktio. Toisena esimerkkinä interaktiosta mainittakoon eräiden lääkkeiden ja alkoholin yhdysvaikutus: Esimerkiksi monilla kipulääkkeillä ja alkoholilla on sellainen interaktio, että henkilö, joka ottaa tällaista lääkettä alkoholin kanssa saattaa menehtyä, vaikka hänen nauttimansa alkoholimäärä aiheuttaa normaalioloissa vain lievän humalatilan. Interaktioiden tutkiminen on tilastollisen analyysin keskeisiä tehtäviä.

Seuraavassa esimerkissä tarkastellaan *suoriin havaintoihin* perustuvaa *otantatutkimusta*, jossa kiinnostuksen kohteena olevia tekijöitä pyritään selittämään kahdella selittävällä tekijällä.

ESIMERKKI 2.

Erään kunnan koululautakunta haluaa ottaa selville kunnan lukioiden oppilaiden alkoholin käytön. Käyttö päätetään kartoittaa kyselytutkimuksella. Lautakunta ajattelee, että riittää, että kerätään yksinkertainen satunnaisotos oppilaiden joukosta ja kohdistetaan kysely otokseen poimituille oppilaille. Tehtävä päätetään antaa kunnan tilastotoimiston tehtäväksi. Kunnan tilastoimisto toteaa kuitenkin heti, että tällainen tutkimussuunnitelma on puutteellinen ja, että oppilaiden alkoholin käytöstä saadaan enemmän informaatiota, jos aineisto kerätään käyttämällä *ositettua otantaa*:

Ensinnäkin se millä luokalla oppilaat ovat (tai oppilaiden ikä) saattaa vaikuttaa heidän alkoholin käyttöönsä. Samoin tiedetään, että tytöt ja pojat suhtautuvat erilailla alkoholin käyttöön. Niinpä tilastotoimisto ehdottaa, että tutkimusaineisto kerätään soveltamalla *ositettua otantaa*, jossa tutkimuksen kohteena oleva *perusjoukko* eli kunnan lukioiden oppilaat jaetaan $3 \times 2 = 6$ *ositteeseen* luokan ja sukupuolen suhteen. Jokaisesta ositteesta poimitaan yksinkertainen satunnaisotos. Tässä kyselytutkimuksessa ovat *vasteina* vastaukset alkoholin käyttöä kartoittaviin kysymyksiin. *Selittäviä tekijöitä* on kaksi: luokka ja sukupuoli. Tekijällä luokka on kolme *tasoa* ja tekijällä sukupuoli on kaksi *tasoa*. Siten otos jakaantuu 6 ryhmään (ositteeseen) ja tehtävänä on vertailla oppilaiden alkoholin käyttöä näissä ryhmissä.

Tutkimusasetelmaa ja ositteita voidaan kuvata seuraavalla kaaviolla:

	Luokka		
Sukupuoli	1	2	3
Poika	1	2	3
Tyttö	4	5	6

Nytkin on kiinnostavaa tietää sukupuolen ja luokan erillisten vaikutusten lisäksi, millainen mahdollinen yhdysvaikutus näillä tekijöillä on vasteeseen. ●

Seuraavilla esimerkeillä kuvataan tarkemmin yhdysvaikutuksen käsitettä.

ESIMERKKI 3.

Tarkastellaan mies- ja naispuolisten matemaatikoiden ja fyysikoiden vuosipalkkoja vuonna 1985 USA:ssa. Vuotuiset keskipalkat (1000$) on ilmaistu seuraavassa taulukossa:

Ala	Naiset	Miehet	Keskiarvo
Fysiikka	41.2	48.6	44.9
Matematiikka	34.7	42.3	38.5
Keskiarvo	37.95	45.45	41.7

Mitä johtopäätöksiä tästä taulukosta voidaan tehdä? Ensinnäkin fyysikot ansaitsivat enemmän kuin matemaatikot ja miehet ansaitsivat enemmän kuin naiset. Tämän lisäksi taulukosta nähdään, että miesten ja naisten keskipalkkojen ero oli n. 7,500$ riippumatta siitä olivatko he fyysikoita tai matemaatikoita ja se, että fyysikoiden ja matemaatikoiden keskipalkkojen ero oli n. 6,500$ olivatpa he naisia tai miehiä.

Siten taulukosta voidaan tehdä seuraavat johtopäätökset:

(1) Ala *ei vaikuttanut* sukupuolikohtaisiin palkkaeroihin.

(2) Sukupuoli *ei vaikuttanut* alakohtaisiin palkkaeroihin.

Tällaisessa tilanteessa *päävaikutukset* selittävät palkkaerot täydellisesti. Taulukossa esitetyt *marginaali-* eli *reunakeskiarvojen* erot kuvaavat hyvin palkkaeroja. Koska päävaikutukset riittävät selittämään palkkaerot, selittävien tekijöiden välillä ei ole interaktiota. *Tilastollista testiä* interaktiaktion olemassaololle tarkastellaan myöhemmin tässä kappaleessa. ●

Edellisessä esimerkissä selittävillä tekijöillä (sukupuolella ja koulutuksella) ei ollut interaktiota. Sen sijaan seuraavassa esimerkissä selittävillä tekijöillä on interaktiota.

ESIMERKKI 4.

Tarkastellaan mies- ja naispuolisten biologien ja lääkäreiden vuosipalkkoja vuonna 1985 USA:ssa. Vuotuiset keskipalkat (1000$) on ilmaistu seuraavassa taulukossa:

Ala	Naiset	Miehet	Keskiarvo
Biologia	34.5	42.0	38.25
Lääketiede	36.2	50.4	43.30
Keskiarvo	35.35	46.2	40.775

Mitä johtopäätöksiä tästä taulukosta voidaan tehdä? Ensinnäkin lääkärit ansaitsivat enemmän kuin biologit ja miehet ansaitsivat enemmän kuin naiset. Tämän lisäksi taulukosta nähdään, että miesten ja naisten keskipalkkojen ero oli 7,500$, jos he olivat biologeja ja 14,200$, jos he olivat lääkäreitä. Edelleen lääkäreiden ja biologien keskipalkkojen ero oli 1,700$ jos he olivat naisia ja 8,400$, jos he olivat miehiä.

Siten taulukosta voidaan tehdä seuraavat johtopäätökset:

(1) Alalla *oli vaikutusta* sukupuolikohtaisiin palkkaeroihin.

(2) Sukupuolella *oli vaikutusta* alakohtaisiin palkkaeroihin.

(3) Alalla ja sukupuolella *oli yhdysvaikutusta* palkkaeroihin.

Tällaisessa tilanteessa *päävaikutukset* eivät riitä selittämään palkkaeroja. Koska päävaikutukset eivät riitä selittämään palkkaeroja, selittävien tekijöiden välillä on interaktiota. *Tilastollista testiä* interaktiaktion olemassaololle tarkastellaan myöhemmin tässä kappaleessa. ●

Edellisessä esimerkissä sillä, että henkilö oli mies ja sillä, että henkilö oli lääkäri oli *positiivinen* interaktio siinä mielessä, että näiden tekijöiden esiintyminen yhdessä lisäsi palkkaeroa. Interaktiot voivat tulla esiin usealla hyvin monilla erilaisilla tavoilla. Seuraavassa esimerkkissä selittävien tekijöiden interaktion luonne on aivan erilainen kuin edellisessä esimerkissä.

ESIMERKKI 5.

Eräässä kokeessa tutkittiin miten melu vaikuttaa hyperaktiiviisten ja normaaliaktiivisten lasten suorituskykyyn. Lapset poimittiin kokeeseen satunnaisesti erään peruskoulun toisilta luokilta. Hyperaktiiviset lapset muodostivat koeryhmän ja normaaliaktiviset lapset muodostivat kontrolliryhmän. Eräs lapsille tehdyistä kokeista mittasi kykyä ratkaista aritmeettisia tehtäviä meluisassa ja hiljaisessa luokassa. Seuraava taulukko kuvaa kokeen tuloksista määrättyjä keskimääräisiä koepisteitä (korkeampi pistemäärä ~ parempi tulos):

Ryhmä	Meluisa luokka	Hiljainen luokka	Keskiarvo
Normaaliaktiiviset lapset	214	170	192
Hyperaktiiviset lapset	120	140	130
Keskiarvo	167	155	161

Mitä johtopäätöksiä tästä taulukosta voidaan tehdä? Ensinnäkin normaaliaktiiviset lapset saavat testissä keskimäärin korkeampia pistemääriä kuin hyperaktiiviset lapset. Toiseksi meluisassa luokassa saadaan keskimäärin korkeampia pistemääriä kuin hiljaisessa luokassa. Mutta, kun normaaliaktiivisten lasten suoritus on meluisassa luokassa parempi kuin hiljaisessa luokassa, hyperaktiivisilla lapsilla suoritus on parempi hiljaisessa luokassa.

Siten taulukosta voidaan tehdä seuraavat johtopäätökset:

(1) Luokan meluisuudella on vaikutusta lasten suorituskykyyn.

(2) Lasten aktiivisuuden asteella on vaikutusta lasten suorituskykyyn.

(3) Lasten aktiivisuuden asteella ja luokan meluisuudella on yhdysvaikutusta lasten suorituskykyyn.

Kumpikaan päävaikutuksista ei riitä yksinään selittämään suorituskyvyn vaihtelua, vaan selittävillä tekijöillä on tässä tilanteessa yhdysvaikutusta. Lisäksi luokan hiljaisuus vaikuttaa normaaleihin ja hyperaktiivisiin lasten suorituskykyyn eri suuntiin. *Tilastollista testiä* interaktiaktion olemassaololle tarkastellaan myöhemmin tässä kappaleessa. ●

Edellisessä esimerkissä luokan hiljaisuudella ja lapsen aktiivisuudella oli *negatiivinen* interaktio siinä mielessä, että näiden tekijöiden yhdysvaikutus vähensi lasten suorituskyvyssä havaittua eroa.

6.3.2 2-SUUNTAISEN VARIANSSIANALYYSIN PERUSASETELMA

Oletetaan, että tutkimuksen kohteena oleva perusjoukko jaetaan ryhmiin kahden *tekijän* A ja B suhteen. Oletetaan, että tekijällä A on I *tasoa* ja, että tekijällä B on J *tasoa*. Siten tutkimuksen kohteena oleva perusjoukko jaetaan IJ ryhmään tekijöiden A ja B suhteen. Tutkimuksen kohdistuu satunnaismuuttujan Y *keskimäärin* saamiin arvoihin tekijöiden A ja B tasojen määräämissä ryhmissä.

Olkoot Y_{ijk} muuttujan Y tekijän A tason i ja tekijän B tason j määräämään ryhmään kuuluva k. havainto ja olkoot

$$E(Y_{ijk}) = \mu_{ij}, \quad i = 1,2,\ldots,I, \quad j = 1,2,\ldots,J, \quad k = 1,2,\ldots,K$$

havainnon Y_{ijk} odotusarvo ja

$$D^2(Y_{ijk}) = \sigma^2, \quad i = 1,2,\ldots,I, \quad j = 1,2,\ldots,J, \quad k = 1,2,\ldots,K$$

havainnon Y_{ijk} varianssi.

Huomaa, että sallimme sen, että ryhmäkohtaiset odostusarvot μ_{ij} vaihtelevat ryhmästä toiseen, mutta oletamme, että kaikilla havainnoilla on sama varianssi σ^2. Huomaa edelleen, että ryhmäkohtainen havaintojen lukumäärä K on sama kaikissa ryhmissä. Tavallisesti oletamme lisäksi, että havainnot Y_{ijk} ovat jakautuneet kuten normaalijakauma ym. parametrein.

2-suuntaisessa varianssianalyysissa kiinnostuksen kohteena on *kolme* erilaista nollahypoteesia. Nollahypoteeseista kaksi koskee tekijöiden A ja B määräämiä *päävaikutuksia*. Nämä nollahypoteesit ovat muotoa

H_A: Ei A-vaikutusta,

H_B: Ei B-vaikutusta.

Kolmas nollahypoteeseista koskee tekijöiden A ja B yhdysvaikutusta:

H_{AB}: Ei yhdysvaikutusta.

Nämä hypoteesit ovat siinä mielessä erilaisessa asemassa, että hypoteeseista ensisijainen on H_{AB} ja sitä on syytä testata aina ensimmäisenä. Tämä johtuu siitä, että mahdollisen *yhdysvaikutuksen olemassaolon huomiotta jättäminen tekee melko* *mahdottomaksi ymmärtää oikein vastemuuttujan käyttäytymistä*. Vasta jos hypoteesi H_{AB} jää voimaan, voidaan päävaikutuksien olemassaoloa tutkia *erillisinä* ongelmina testaamalla hypoteeseja H_A ja H_B. Jos kaikki mainitut hypoteesit jäävät voimaan, voidaan ryhmät yhdistää muuttujaa Y koskevissa tarkasteluissa.

Seuraavassa kappaleessa todetaan, että mainittuja hypoteeseja voidaan testata tavanomaisilla varianssien vertailuun tarkoitetuilla F-testeillä.

6.3.3 2-SUUNTAISEN VARIANSSIANALYYSIIN LIITTYVÄT TESTIT

Havainnoista Y_{ijk} voidaan määrätä tekijän A tasoon i ja tekijän B tasoon j liittyvä *ryhmäkeskiarvo*

$$\overline{Y}_{ij.} = \frac{1}{K}\sum_{k=1}^{K}\overline{Y}_{ijk}$$

ja *ryhmävarianssi*

$$s_{ij}^2 = \frac{1}{K-1}\sum_{k=1}^{K}(Y_{ijk} - \overline{Y}_{ij.})^2 .$$

2-suuntaisen varianssianalyysin tarkoituksena on ryhmäkeskiarvojen $\overline{Y}_{ij.}$ vertailu sekä selvittää voidaanko ryhmäkeskiarvojen eroja selittää selittävillä tekijöillä A ja B.

Havainnoista voidaan määrätä myös *yleiskeskiarvo*

$$\overline{Y}_{-} = \frac{1}{IJK}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{ijk} ,$$

joka on kaikkien havaintojen Y_{ijk} keskiarvo, kun ryhmät on yhdistetty yhdeksi perusjoukoksi sekä *marginaali-* eli *reunakeskiarvot*

$$\overline{Y}_{i\_} = \frac{1}{JK}\sum_{j=1}^{J}\sum_{k=1}^{K}Y_{ijk} ,$$

$$\overline{Y}_{.j.} = \frac{1}{IK}\sum_{i=1}^{I}\sum_{k=1}^{K}Y_{ijk} .$$

Reunakeskiarvo $\overline{Y}_{i\_}$ on havaintojen Y_{ijk} keskiarvo tekijän A tason i määräämässä ryhmässä, kun perusjoukko on jaettu ryhmiin pelkästään A-tekijän suhteen. Reunakeskiarvo $\overline{Y}_{.j.}$ on havaintojen Y_{ijk} keskiarvo tekijän B tason j määräämässä ryhmässä, kun perusjoukko on jaettu ryhmiin pelkästään B-tekijän suhteen.

Kannattaa huomata, että yleiskeskiarvo $\overline{Y}_{-}$ on kaikkien ryhmäkeskiarvojen keskiarvo, ts.

$$\overline{Y}_{-} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\overline{Y}_{ij.} .$$

Myös reunakeskiarvot voidaan määrätä ryhmäkeskiarvoista:

$$\overline{Y}_{i\_} = \frac{1}{J}\sum_{j=1}^{J}\overline{Y}_{ij.} ,$$

$$\overline{Y}_{.j.} = \frac{1}{I}\sum_{i=1}^{K}\overline{Y}_{ij.} .$$

Varianssianalyysin lähtökohtana ovat edellä määritellyt ryhmäkeskiarvot ja ryhmävarianssit (ryhmähajonnat). Jos ryhmäkeskiarvot eivät eroa toisistaan kovin paljon ryhmien sisäistä vaihtelua kuvaaviin ryhmähajontoihin nähden, voidaan ryhmät yhdistää muuttujaa Y koskevissa tarkasteluissa yhdeksi perusjoukoksi. Analyysia hankaloittaa selittävien tekijöiden mahdollinen interaktio.

Varianssianalyysin lähtökohtana olevat ryhmäkeskiarvot, ryhmähajonnat, yleis- ja reunakeskiarvot voidaan järjestää taulukoksi seuraavaan tapaan:

	Tekijä B				
Tekijä A	1	2	...	J	Reuna-keskiarvo
1	$\bar{Y}_{11\bullet}$ s_{11}	$\bar{Y}_{12\bullet}$ s_{12}	...	$\bar{Y}_{1J\bullet}$ s_{1J}	$\bar{Y}_{1\bullet\bullet}$
2	$\bar{Y}_{21\bullet}$ s_{21}	$\bar{Y}_{22\bullet}$ s_{22}	...	$\bar{Y}_{2J\bullet}$ s_{2J}	$\bar{Y}_{2\bullet\bullet}$
...	...	...	...	...	...
I	$\bar{Y}_{I1\bullet}$ s_{I1}	$\bar{Y}_{I2\bullet}$ s_{I2}	...	$\bar{Y}_{IJ\bullet}$ s_{IJ}	$\bar{Y}_{I\bullet\bullet}$
Reuna-keskiarvo	$\bar{Y}_{\bullet 1\bullet}$	$\bar{Y}_{\bullet 2\bullet}$	...	$\bar{Y}_{\bullet J\bullet}$	$\bar{Y}_{\bullet\bullet\bullet}$

Tästä on nähty esimerkkejä edellisessä kappaleessa.

Kirjoitetaan identiteetti

$$Y_{ijk} - \overline{Y}_{\_} = (\overline{Y}_{i\_} - \overline{Y}_{\_}) + (\overline{Y}_{.j.} - \overline{Y}_{\_}) + (\overline{Y}_{ij.} - \overline{Y}_{i\_} - \overline{Y}_{.j.} + \overline{Y}_{\_}) + (Y_{ijk} - \overline{Y}_{ij.}).$$

Tätä identiteettiä vastaa *neliösummahajoitelma*

$$\Sigma\Sigma\Sigma(Y_{ijk} - \overline{Y}_{\_})^2 = \Sigma\Sigma\Sigma(\overline{Y}_{i\_} - \overline{Y}_{\_})^2 + \Sigma\Sigma\Sigma(\overline{Y}_{.j.} - \overline{Y}_{\_})^2$$
$$+ \Sigma\Sigma\Sigma(\overline{Y}_{ij.} - \overline{Y}_{i\_} - \overline{Y}_{.j.} + \overline{Y}_{\_})^2 + \Sigma\Sigma\Sigma(Y_{ijk} - \overline{Y}_{ij.})^2.$$

Tämä neliösummahajoitelma kirjoitetaan tavallisesti symbolisessa muodossa

$$SST = SSA + SSB + SSAB + SSE,$$

jossa *kokonaisneliösumma* SST on siis jaettu neljään osaan:

SSA on tekijän A *päävaikutukseen* liittyvä neliösumma,

SSB on tekijän B *päävaikutukseen* liittyvä neliösumma,

$SSAB$ on tekijöiden A ja B *yhdysvaikutukseen* liittyvä neliösumma,

SSE on *jäännösneliösumma*.

Merkitään jatkossa havaintojen kokonaislukumäärää seuraavalla tavalla:

$$N = IJK.$$

Edellä mainittuja neliösummia vastaavat *vapausasteet* toteuttavat yhtälön

$$IJK - 1 = (I - 1) + (J - 1) + (I - 1)(J - 1) + IJ(K - 1),$$

jossa yhtälön vasemmalla puolella oleva termi voidaan kirjoittaa muotoon

$$IJK - 1 = N - 1$$

ja yhtälön oikealla puolella viimeisenä mainittu termi voidaan kirjoittaa muotoon

$$IJ(K - 1) = N - IJ.$$

Kannattaa huomata, että kaikki edellä mainitut neliösummat voidaan lausua havaintojen, ryhmäkeskiarvojen, ryhmähajontojen, reunakeskiarvojen ja yleiskeskiarvon avulla:

$$SST = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk} - \overline{Y}_{\bullet\bullet\bullet})^2,$$

$$SSA = JK\sum_{i=1}^{I}(\overline{Y}_{i\bullet\bullet} - \overline{Y}_{\bullet\bullet\bullet})^2,$$

$$SSB = IK\sum_{j=1}^{J}(\overline{Y}_{\bullet j\bullet} - \overline{Y}_{\bullet\bullet\bullet})^2,$$

$$SSAB = K\sum_{i=1}^{I}\sum_{j=1}^{J}(\overline{Y}_{ij\bullet} - \overline{Y}_{i\bullet\bullet} - \overline{Y}_{\bullet j\bullet} + \overline{Y}_{\bullet\bullet\bullet})^2,$$

$$SSE = (K-1)\sum_{i=1}^{I}\sum_{j=1}^{J}s_{ij}^2.$$

Huomaa myös, että havaintojen Y_{ijk} kokonaisvarianssi s^2 saadaan kokonaisneliösummasta kaavalla

$$s^2 = \frac{SST}{N-1}.$$

Käyttämällä edelläesitettyjä määritelmiä ja identiteettejä voidaan rakentaa 2-suuntaisen varianssianalyysin *varianssianalyysitaulukko*:

Taulukko 1. Varianssianalyysitaulukko 2-suuntaisessa varianssianalyysissa.

Vaihtelun lähde	Neliösumma	Vapausasteet	Varianssiestimaattori	F-testisuure
A	SSA	$I-1$	$\frac{SSA}{I-1}$	$F_A = \frac{N-IJ}{I-1} \cdot \frac{SSA}{SSE}$
B	SSB	$J-1$	$\frac{SSB}{J-1}$	$F_B = \frac{N-IJ}{J-1} \cdot \frac{SSB}{SSE}$
AB	$SSAB$	$(I-1)(J-1)$	$\frac{SSAB}{(I-1)(J-1)}$	$F_{AB} = \frac{N-IJ}{(I-1)(J-1)} \cdot \frac{SSAB}{SSE}$
Virhe	SSE	$N-IJ$	$\frac{SSE}{N-IJ}$	
Kokonaisvaihtelu	SST	$N-1$	$\frac{SST}{N-1}$	

Taulukosta 1 saadaan testisuure hypoteesille H_{AB} sekä testisuureet hypoteeseille H_A ja H_B:

1. Testataan ensin *yhdysvaikutuksen* olemassaoloa:

   Jos hypoteesi H_{AB} pätee, testisuure F_{AB} noudattaa F-jakaumaa vapausastein $(I-1)(J-1)$ ja $N-IJ$.

2. *Jos hypoteesi* H_{AB} *jää voimaan*, voidaan testata *päävaikutusten* olemassaoloa:

   Jos hypoteesi H_A pätee, testisuure F_A noudattaa F-jakaumaa vapausastein $I-1$ ja $N-IJ$.

   Jos hypoteesi H_B pätee, testisuure F_B noudattaa F-jakaumaa vapausastein $J-1$ ja $N-IJ$.

Kaikissa kolmessa tapauksessa testisuureen odotettavissa oleva arvo on vastaavan hypoteesin pätiessä

$$\mathrm{E}(F) = \frac{N-IJ}{N-IJ-2}$$

ja testisuureen *suuret* arvot viittaavat siihen, että ko. hypoteesi ei päde.

6.3.4 SOVELLUS

Eräässä kokeessa tutkittiin liikunnan harrastamisen ja sukupuolen vaikutuksia sydämen lyöntitiheyteen eli sykkeeseen fyysisessä rasituksessa. Syke mitattiin määräämällä sydämen lyöntien lukumäärä minuutissa. Kokeessa verrattiin säännöllisesti hölkkäävien, joilla tarkoitettiin yli 25 km viikossa hölkkääviä, miesten ja naisten sykkeitä liikuntaa harrastamattomien miesten ja naisten sykkeisiin, kun koehenkilöiden oli annettu juosta 6 minuuttia juoksumatolla.

Tässä tutkimusasetelmassa hölkkää harrastaneet muodostivat koeryhmän ja liikuntaa harrastamattomat kontrolliryhmän. Tutkimusasetelmassa oli vasteena koehenkilön syke. Selittävinä tekijöinä oli kaksi tekijää: liikuntaharrastus (A) ja sukupuoli (B). Kummallakin selittävällä tekijällä oli kaksi tasoa (hölkkäävä/liikuntaa harrastamaton, mies/nainen). Koeryhmiä oli siten $2 \times 2 = 4$. Koehenkilöiden sykkeitä verrattiin toisiinsa näissä ryhmissä. Kuhunkin ryhmään poimittiin 200 koehenkilöä.

On syytä huomata se, että itse asiassa tällainen koe sisältää harhan lähteen, koska sekä hölkkääjät että liikuntaa harrastamattomat ovat itse päättäneet miten paljon he harrastavat liikuntaa. Emme kuitenkaan analysoi tätä harhan mahdollisuutta jatkossa.

Koetulokset on annettu seuraavassa taulukossa:

	Sukupuoli	
Liikunnan harrastus	Mies	Nainen
Säännöllisesti hölkkäävä	103.1 (12.67)	117.9 (15.22)
Liikuntaa harrastamaton	129.4 (16.92)	150.2 (15.91)

Taulukossa on kerrottu sykkeiden keskiarvot kussakin ryhmässä. Vastaavat keskihajonnat on kerrottu suluissa keskiarvojen alapuolella. Taulukosta voidaan tehdä seuraavat johtopäätökset:

(1) Miesten syke on keskimäärin alempi kuin naisten syke.

(2) Hölkkääjien syke on keskimäärin alempi kuin naisten syke.

(3) Sukupuolen ja hölkän harrastamisen välillä saattaa olla yhdysvaikutusta.

Yhdysvaikutuksen olemassaoloon viittaa se, että hölkkäävien ja liikuntaa harrastamattomien miesten keskisykkeiden ero oli 26.3 lyöntiä minuutissa, kun taas hölkkäävien ja liikuntaa harrastamattomien naisten keskisykkeiden ero oli 32.3 lyöntiä minuutissa. Onko tämä ero niin suuri, että yhdysvaikutus on tilastollisesti merkitsevää, nähdään alla.

Taulukossa annettujen tietojen avulla voidaan muodostaa seuraava varianssianalyysitaulukko:

Vaihtelun lähde	Neliö-summa	Vapaus-asteet	Varianssi-estimaattori	F-testisuure
A	171698	1	171698	737.2
B	63368	1	63368	272.1
AB	1800	1	1800	7.73
Jäännös	185387	796	232.9	
Kokonais-vaihtelu	422253	799	528.5	

Kaikki taulukon F-testisuureet ovat merkitseviä. Erityisesti yhdysvaikutuksen olemassaoloa testaava testisuureella F_{AB} on arvo 7.73, jota vastaava *P*-arvo on 0.006. Siten saamme vahvistuksen esitetylle epäilylle, että liikunnanharrastuksella ja sukupuolella on yhdysvaikutusta sydämen lyöntitiheyteen. Koska nollahypoteesi H_{AB} tulee siis hylätyksi, ei päävaikutusten merkitsevyyttä tarvitse tutkia: Päävaikutusten olemassaolo on selvää, jos niillä on yhdysvaikutusta.

HARJOITUSTEHTÄVIÄ

LUKU 1.

1. Eräässä maassa suunnitellaan auton rekisterikilpien uudistamista. Uudistuksessa halutaan varautua tulevaisuutta varten tarpeeksi suurella määrällä erilaisia kilpiä. On ehdotettu kahta erilaista kilpeä:

    (a) Xxnnn,
    (b) XXXnn,

joissa X = kirjain a,...,ö ja n = numero 0,1,2,...,9. Kumpia kilpiä on enemmän?

2. Kuinka monta erilaista sanaa on mahdollista muodostaa sanan

    (a) apua,
    (b) saippuakauppias

kirjaimista, kun sanoihin käytetään aina kaikki kirjaimet?

3. Kuinka monella tavalla 7 henkilöä voidaan istuttaa

    (a) penkille,
    (b) pyöreään pöytään?

4. Suomessa on normaalikäytössä 6 erilaista kolikkoa (10p, 20p, 50p, 1mk, 5mk, 10mk). Oletetaan, että sinulla on käytössäsi täsmälleen 1 kolikko kutakin lajia. Kuinka monta erilaista rahasummaa voidaan näistä kolikoista saada aikaan?

5. Ihmisten verityyppi jakaantuu ryhmiin O, A, B, AB. Satunnaisesti valitun afroamerikkalaisen verityypin todennäköisyydet on taulukoitu alla.

    (a) Mikä on todennäköisyys, että satunnaisesti valitun mustan verityyppi on AB?
    (b) Mikä on todennäköisyys, että satunnaisesti valitun mustan verityyppi on A *tai* B?

6. Eräässä suureen otokseen perustuvassa tutkimuksessa kysyttiin 1. vuoden college-opiskelijoilta minkälainen heidän opintomenestyksensä oli ollut high schoolin viimeisellä luokalla. Opintomenestystä mitattiin 5-portaisella asteikolla. Todennäköisyydet on taulukoitu alla. Olkoon A="Opiskelija sijoittui parhaaseen 40%:iin" ja B="Opiskelija sijoittui huonoimpaan 40%:iin".

    (a) Määrää P(A) ja P(B).
    (b) Kuvaa A^C sanoin. Määrää $P(A^C)$ sekä suoraan taulukosta että käyttämällä komplementtitodennäköisyyden kaavaa.
    (c) A ja B ovat toisensa poissulkevia. Määrää P(A *tai* B).
    (d) Määrää tapahtuman (A *tai* B)^C todennäköisyys suoraan taulukosta ja käyttämällä apuna kohtaa (c) ja komplementtitodennäköisyyden kaavaa.

7. Eräissä arpajaisissa on myytävänä 200 arpaa. Jotta arpajaiset olisivat osallistujia kohtaan reilut, voittojen arvonta pitää järjestää niin, että jokaisen arvan voiton todennäköisyys on yhtä suuri. Mikä on siis satunnaisesti valitun arvan voiton todennäköisyys? Kuvaa arvontatapahtumaan liittyvä todennäköisyysmalli?

8. Rulettipyörässä on 37 lokeroa, jotka on numeroitu numeroilla 0 ja 1–36. Numero 0 on vihreä, numeroiden 1–36 joukossa on 18 mustaa ja 18 punaista numeroa.

Rulettipyörää pyöräytetään ja pieni pallo pannaan pyörimään ruletin kehälle vastakkaiseen suuntaan. Pyörän pysähdyttyä pallo asettuu johonkin lokeroista. Pyörä on tasapainotettu niin, että pallolla on yhtä suuri todennäköisyys päätyä mihin tahansa lokeroista.

(a) Eräs tapa pelata on asettaa panos "numerolle". Mikä on todennäköisyys, että tulee valittu numero?

(b) Eräs tapa pelata on asettaa panos "värille". Mikä on todennäköisyys, että tulee punainen numero?

(c) Numerot on asetettu sarakkeiksi niin, että esimerkiksi 3:n monikerrat 3,6,9,...,36 muodastavat yhden sarakkeista. Eräs tapa pelata on asettaa panos tällaiselle sarakkeelle. Mikä on todennäköisyys, että tulee numero mainitusta sarakkeesta?

Tehtävä 5:

O	A	B	AB
0.49	0.27	0.20	?

Tehtävä 6:

Paras 20%	Toinen 20%	Kolmas 20%	Neljäs 20%	Huonoin 20%
0.41	0.23	0.29	0.06	0.01

9. Heitetään kahta noppaa. Olkoon x_1 = 1. nopan silmäluku ja x_2 = 2. nopan silmäluku. Määritellään tapahtumat

$$A = \{(x_1,x_2) | \, x_1 \geq 5\}$$
$$B = \{(x_1,x_2) | \, 2 \leq x_1 \leq 3,\ 3 \leq x_2 \leq 5\}$$
$$C = \{(x_1,x_2) | \, x_1 + x_2 > 8\}$$
$$D = \{(x_1,x_2) | \, x_1 - x_2 \leq 2\}$$

Piirrä ruudulliselle paperille 10 kappaletta 6×6-ruudukoita, jotka kuvaavat kahden nopanheiton tuloksiin liittyvää otosavaruutta. Esitä näissä ruudukoissa tapahtumat

A, B, C, D
$A \cup B,\ C \cup D$
$A \cap B,\ A \cap C,\ A \cap D$
D^C

10. Määrää kaikkien tehtävässä 1 määriteltyjen joukkojen todennäköisyydet.

11. Näytä, että yleinen yhteenlaskusääntö pätee tehtävän 1 joukolle $A \cup C$.

12. Heitetään kahta noppaa. Olkoon x_1 = 1. nopan silmäluku ja x_2 = 2. nopan silmäluku. Määritellään tapahtumat

$$A = \{(x_1,x_2) | \, x_1 + x_2 \geq 10\}$$
$$B = \{(x_1,x_2) | \, x_1 < 3\}$$
$$C = \{(x_1,x_2) | \, 4 \leq x_2 \leq 6\}$$

(a) Ovatko A ja B toisensa poissulkevia?
(b) Ovatko A ja C toisensa poissulkevia?
(c) Ovatko A ja C riippumattomia?
(d) Ovatko B ja C riippumattomia?

13. Eräässä 200 henkilön ryhmässä on 80 miestä ja 120 naista. Ryhmässä on 60 ydinvoiman kannattajaa ja 140 vastustajaa. Olkoon x satunnaisesti valittu ryhmän jäsen ja oletetaan, että seuraavat tapahtumat ovat riippumattomia:

$M = \{x \text{ on mies}\}$
$N = \{x \text{ on nainen}\}$
$K = \{x \text{ kannattaa ydinvoimaa}\}$
$V = \{x \text{ vastustaa ydinvoimaa}\}$

(a) Määrää P(M), P(N), P(K), P(V).
(b) Määrää P($M \cap K$), P($M \cap V$), P($N \cap K$), P($M \cap V$).
(c) Kuinka moni mies kannattaa ja kuinka moni vastustaa ydinvoimaa?
(d) Kuinka moni nainen kannattaa ja kuinka moni vastustaa ydinvoimaa?

14. Toisessa maailmansodassa amerikkalaisella hävittäjälentäjällä oli noin 2%:n todennäköisyys tulla ammutuksi alas taistelulennon aikana. Lentäjät uskoivat, että todennäköisyys tulla ammutuksi alas kasvoi lentojen lukumäärän kanssa ja olisi varmaa tulla ammutuksi alas, kun taistelulentojen lukumäärä saavutti 50:n rajan. Pitikö lentäjien pelko paikkaansa? Perustele vastaus.

15. Kokonaislukuihin 1,2,...,n liitetään todennäköisyydet, jotka ovat verrannollisia lukujen suuruuteen.

(a) Määrää lukuihin liittyvät todennäköisyydet.
(b) Määrää 1:n ehdollinen todennäköisyys ehdolla, että joko 1 tai *n* on saatu tulokseksi.

16. Alla on taulukko, joka kuvaa todennäköisyyksiä, että satunnaisesti valittu amerikkalainen college-opiskelija kuuluu taulukossa ilmoitettuihin ikä- ja sukupuoliryhmiin.

(a) Mikä on todennäköisyys, että opiskelija on nainen?
(b) Mikä on ehdollinen todennäköisyys, että opiskelija on nainen ehdolla, että hän on yli 35-vuotias?
(c) Mikä on todennäköisyys, että opiskelija on nainen tai vähintään 35-vuotias?

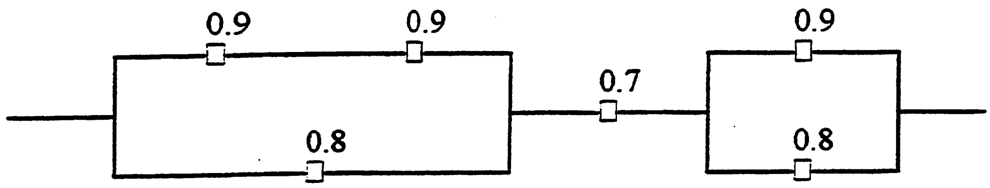
Tehtävä 16:

Ikä	14-17	18-24	25-34	≥35
Mies	.01	.30	.12	.04
Nainen	.01	.30	.13	.09

17. Eräässä tehtaassa on 3 valmistuslinjaa, joilla tehdään samanlaisia CD-soittimia. Linja A valmistaa soittimista 30%, linja B 25% ja linja C 45%. A:n valmistamista soittimista keskimäärin 2% osoittautuu tarkistuksissa viallisiksi, B:n valmistamista soittimista 3% ja C:n valmistamista soittimista 4%. Valitaan satunnaisesti yksi soitin tarkistusta varten.

(a) Mikä on todennäköisyys, että soitin on viallinen?
(b) Mikä on todennäköisyys, että soitin on tehty linjalla A, jos se on viallinen?

18. Seuraava kaavio kuvaa toimintaverkkoa, johon on merkitty kunkin komponentin toimintatodennäköisyys. Mikä on todennäköisyys, että verkon kuvaama systeemi toimii?



19. USA:ssa high schoolissa kilpaurheilua harrastaneiden opiskelijoiden mahdollisuuksista päästä ammattilaisiksi ja heidän koulutuksestaan on olemassa seuraavat tiedot:

High school-kilpaurheilijoista 5% pääsee collegeen.
Collegeen päässeistä high school-kilpaurheilijoista 1.7% pääsee ammattilaisiksi.
Niistä high school-kilpaurheilijoista, jotka eivät pääse collegeen, 0.01% pääsee ammattilaisiksi.

(a) Kuvaa high school-kilpaurheilijoiden "uravaihtoehdot" puudiagrammina.
(b) Määrää todennäköisyys, että high school-kilpaurheilija pääsee ammattilaiseksi.
(c) Määrää todennäköisyys, että ammattilaiseksi päässyt high school-kilpaurheilija on käynyt collegen.

20. Uurnassa on 3 valkoista ja 2 mustaa kuulaa. Uurnasta otetaan satunnaisesti 2 kuulaa ilman takaisinpanoa. Mikä on todennäköisyys, että toinen kuulista on musta?

(a) Ratkaise tehtävä käyttämällä puudiagrammia.
(b) Ratkaise tehtävä kokonaistodennäköisyyden kaavan avulla.

21. Eräs koe koostuu monivalintatehtävistä, joissa jokaisessa on 5 vastausvaihtoehtoa. Pekka arvioi, että hänellä on 75%:n todennäköisyys tietää oikea vastaus. Jos hän ei tiedä oikeata vastausta, hän joutuu arvaamaan, jolloin hänellä on 20%:n todennäköisyys arvata oikein (miksi?). Määrää todennäköisyys, että Pekka antaa oikean vastauksen satunnaisesti valittuun tehtävään. Käytä ratkaisussa puudiagrammia.

22. Jussilla on sepelvaltimotauti. Hänen on päätettävä hoidetaanko häntä lääkkeillä vai tehdäänkö hänelle ohitusleikkaus. Jussi pitää tärkeänä jäljellä olevan elämän

pituuden rinnalla myös jäljellä olevan elämän laatua. Hän haluaa maksimoida todennäköisyyden tapahtumalle

A = "Jussi on hengissä 5 vuoden kuluttua ja voi viettää kohtuullisen aktiivista elämää".

Jussin lääkäri esittää seuraavat todennäköisyysarviot Jussin ikäiselle ja kuntoiselle potilaalle:

- Lääkityksessä P(A) = 0.7.
- Todennäköisyys, että potilas kuolee leikkauksen aiheuttamiin komplikaatioihin on 0.05. Todennäköisyys, että leikkauksen seurauksena on komplikaatioita, mutta potilas jää henkiin on 0.10. Todennäköisyys, että potilaalla ei ole leikkauksen jälkeen komplikaatioita on 0.85.
Jos potilaalla on leikkauksen seurauksena komplikaatioita, tapahtuman A todennäköisyys on 0.73. Jos potilaalla ei ole leikkauksen jälkeen komplikaatioita, tapahtuman A todennäköisyys on 0.76.

(a) Piirrä Jussin vaihtoehdoista puudiagrammi.
(b) Mikä on P(A), jos Jussi valitsee leikkauksen?
(c) Tarjoaako leikkaus vai lääkehoito paremmat mahdollisuudet saavuttaa A?

23. Oletetaan, että 10%:lla niistä potilaista, jotka lähetetään sairaalaan tutkimusta varten epäiltynä erään vaarallisen taudin kantajiksi todellakin on ko. tauti. Ensidiagnoosiin käytetään yksinkertaista ja halpaa koetta, joka paljastaa 90% niistä, jotka todella kanatava tautia. Toisaalta 10% niistä, jotka eivät ole taudin kantajia tulevat diagnostisoiduiksi virheellisesti taudin kantajiksi. Määrää todennäköisyys, että henkilö, joka tulee ensidiagnoosissa diagnostisoiduiksi taudin kantajiksi onkin terve.

(a) Käytä Bayesin kaavaa.
(b) Ratkaise tehtävä puudiagrammin avulla.

24. Tehtävä 23 voidaan ratkaista myös käyttämättä eksplisiittisesti todennäköisyyslaskennan sääntöjä. Ajattele, että sairaalaan tulee 1000 potilasta, joista siis 10% on taudin kantajia. Käytä puudiagrammia esittämään miten ko. 1000 potilaalle käy eri diagnoosivaihtoehdoissa. Tehtävä on helppo ratkaista puudiagrammin avulla. Vrt. ratkaisumenetelmää tehtävän 23 (b)-kohdassa käytettyyn menetelmään.

Huomaa, että tehtävässä 24 esitettyä menetelmää voidaan usein käyttää apuna yksinkertaisten todennäköisyyslaskennan päättelytehtävien ratkaisemisessa.

25. Tarkastellaan 5-lapsisia perheitä ja oletetaan, että lapset ovat syntyneet yksi kerrallaan ja jokaisen syntyneen lapsen sukupuoli on riippumaton siitä, mitä sukupuolta aikaisemmin syntyneet lapset ovat olleet. Oletetaan lisäksi, että P(tyttö) = P(poika) = 0.5. Olkoon satunnaismuuttuja X = tyttöjen lukumäärä perheessä.

(a) Määrää X:n pistetodennäköisyysfunktio.
(b) Määrää X:n kertymäfunktio.
(c) Hahmottele kummankin funktion kuvaajat paperille.
(d) Määrää tapahtuman $X > 2$ todennäköisyys sekä tiheysfunktion että kertymäfunktion avulla.

26. Tarkastellaan luvun valitsemista satunnaisesti väliltä $[-1,0]$ ja olkoon satunnaismuuttuja X = valittu luku. Tämän satunnaismuuttujan tiheysfunktio on muotoa $f(x) = 1$, kun $x \in [-1,0]$ ja $f(x) = 0$ muulloin. Määrää seuraavat todennäköisyydet:

(a) $P(-0.5 \le X \le -0.1)$,
(b) $P(-1 \le X \le 0)$,
(c) $P(0.5 \le X \le 0.6)$,
(d) $P(X = -0.1)$.

27. Olkoon X normaalinen $N(-1,2)$. Määrää normaalijakauman taulukoista (tai jonkin tilasto-ohjelman avulla) todennäköisyydet

(a) $P(-3 \le X \le +1)$,
(b) $P(-3 \le X \le 0)$,
(c) $P(X \le -2$ *tai* $X \ge 0)$,
(d) $P(X = -1)$.

28. Satunnaismuuttuja X:n kertymäfunktio on muotoa

$$
\begin{aligned}
F(x) &= ax^2, && \text{kun } 0 \le x \le 2,\\
F(x) &= 0, && \text{kun } x \le 0,\\
F(x) &= 1, && \text{kun } 2 \le x.
\end{aligned}
$$

Mikä on vakion a arvo?.

29. Satunnaismuuttuja X:n tiheysfunktio on muotoa

$$
\begin{aligned}
f(x) &= 1 + kx, \text{ kun } 0 \le x \le 2,\\
f(x) &= 0, \text{ muulloin.}
\end{aligned}
$$

Mikä on vakion k arvo?

30. Olkoon $E(X_i) = 18$ ja $D^2(X_i) = 9$, kun $i = 1,\ldots,10$. Määritellään

$$
\begin{aligned}
Z_i &= (X_i - 18)/3,\\
Y_i &= (X_i + 9)/9,\\
M &= (X_1 + \ldots + X_{10})/10.
\end{aligned}
$$

Määrää

(a) $E(Z_i)$ ja $D^2(Z_i)$,
(b) $E(Y_i)$ ja $D^2(Y_i)$,
(c) $E(M)$ ja $D^2(M)$.

31. Viinilasien valmistuksessa käytetään kaasuliekkiä, jonka lämpötila vaihtelee hieman. Liekin lämpötila X on siten satunnaismuuttuja, jonka jakauma on karkealla asteikolla mitattuna diskreetti. Jakauma on esitetty tehtävien lopussa.

(a) Määrää odotusarvo $E(X)$ ja teoreettinen hajonta $D(X)$.

(b) Tavoitelämpötila on 550°C. Poikkeama tavoitelämpötilasta on $550 - X$. Mikä on poikkeaman odotusarvo ja varianssi?

(c) USA:ssa tieto liekin lämpötilasta halutaan Fahrenheit-asteissa. Muunnoskaava on Celsius-asteista Fahrenheit-asteisiin on

°F = 9°C/5 + 32.

Olkoon Y = liekin lämpötila Fahrenheit-asteissa. Määrää $E(Y)$ ja $D(Y)$.

32. Eräässä kemian laboratoriossa on käytössä 2 vaakaa. Kummankin vaa'an punnitustulos vaihtelee hieman, kun punnituksia toistetaan. Kummallakin vaa'alla punnitaan 2 g:n punnus. Toinen vaaka tuottaa lukeman X, jonka odotusarvo on 2.000 g ja standardipoikkeama on 0.002 g. Toinen vaaka tuottaa lukeman Y, jonka odotusarvo on 2.001 g ja standardipoikkeama on 0.001 g. Oletetaan, että eri vaaoilla tehdyt punnitukset ovat toisistaan riippumattomia.

(a) Mikä on vaakojen lukemien erotuksen $X - Y$ odotusarvo ja hajonta?

(b) Teet yhden punnituksen kummallakin vaa'alla ja lasket lukemien keskiarvon. Tulos on $Z = (X + Y)/2$. Mikä on $E(Z)$ ja $D(Z)$? Onko keskiarvon Z vaihtelu suurempaa vai pienempää kuin vähemmän vaihtelevan vaa'an lukeman Y vaihtelu?

Tehtävä 31:

Lämpötila (°C)	540	545	550	555	560
Todennäköisyys	0.1	0.25	0.3	0.25	0.1

33. Sijoituksen riskiä arvioidaan usein sijoituksen tuoton standardipoikkeamalla: Mitä suurempi on standardipoikkeama, sitä suurempi on riski. Määrää vakuutusyhtiön riski 21 vuotiaille miehille suunnatussa henkivakuutuksessa, jossa vakuutetun omaisille maksetaan 100,000 mk, jos vakuutettu kuolee seuraavan 5 vuoden aikana. Vakuutusmaksu on 250 mk vuodessa. Yhtiön tuotto koostuu 250 mk vakuutusmaksuista vähennettynä 100,000 mk:lla, jos vakuutettu kuolee. Kuoleman todennäköisyydet on ilmoitettu alla esitetyssä taulukossa. Huomaa, että joudut ensin määräämään 21 vuotiaiden miesten todennäköisyyden kuolla 26 vuotiaana tai sitä vanhempana.

Kuolin-ikä	21	22	23	24	25	26 taiyli
Vakuutuksen tuotto	–99,750	–99,500	–99,250	–99,000	–98,750	1250
Todennäköisyys kuolla	0.00183	0.00186	0.00189	0.00191	0.00193	?

34. Jatkoa tehtävälle 33. Vakuutusyhtiön riski pienenee, jos useampi henkilö ottaa vakuutuksen. Oletetamme seuraavassa, että 21-vuotiaiden kuoliniät eivät riippu toisistaan.

(a) Jos X ja Y ovat yhtiön tuotot kahdesta vakuutuksesta, kokonaistuotto on $X + Y$. Mikä on kokonaistuoton odotusarvo ja standardipoikkeama?

(b) Keskimääräinen tuotto on $(X + Y)/2$. Mikä on keskimääräisen tuoton odotusarvo ja standardipoikkeama? Vertaa tulosta yksittäisen vakuutuksen tuoton odostusarvoon ja standardipoikkeamaan.

(c) Oletetaan, että neljä 21-vuotiasta miestä ottaa vakuutuksen. Mikä on keskimääräisen tuuoton odotusarvo ja standardipoikkeama? Vertaa tulosta kohdassa (b) saatuun tulokseen.

35. Heitetään kahta noppaa. Mikä on noppien silmälukujen summan ja erotuksen odotusarvo ja standardipoikkeama?

36. Jatkoa tehtävälle 35. Pelaaja saa voittona silmäslukujen summan markkoina 10-kertaisena, mutta peliin osallistuminen maksaa 80 mk. Mikä on pelaajan voiton odotusarvo ja standardipoikkeama, kun panos otetaan huomioon? Kannattaako peliin osallistua?

37. Olkoon $E(X) = \mu$ ja $D^2(X) = \sigma^2$. Mikä on standardoidun muuttujan

$$Z = (X - \mu)/\sigma \ .$$

odotusarvo ja standardipoikkeama?

38. Monivalintakokeessa on 12 kysymystä. Jokaisessa tehtävässä on 3 vastausvaihotehtoa. Eräs huonosti kokeeseen valmistautunut opiskelija joutuu vastaamaan kysymyksiin arvaamalla. Olkoon satunnaismuuttuja X oikeiden vastausten lukumäärä. Määrää $P(X = 12)$, $P(X = 6)$, $P(X < 4)$, $E(X)$, $D(X)$.

39. Olkoon $X \sim \mathrm{Bin}(5, p)$, jossa $p = 1/5,\ 1/2,\ 4/5$. Määrää kullekin p:n arvolle $\mathrm{E}(X)$, $\mathrm{D}(X)$, moodi ja mediaani. Mihin suuntaan jakaumat ovat vinoja?

Huomaa:

Diskreetin todennäköisyysjakauman *moodi* on se jakaumaan liittyvän satunnaismuuttujan arvo, jota vastaava todennäköisyys on suurin.

Diskreetin todennäköisyysjakauman *mediaani* on se jakaumaan liittyvän satunnaismuuttujan arvo, jonka alapuolelle jää 50% todennäköisyysmassasta.

40. Syöt luonaalla kahta einesteollisuuden tuotetta A (pääruoka) ja B (jälkiruoka). Kummassakin tuotteessa käytetään erästä lisäainetta parantamaan tuotteen säilyvyyttä. Ko. lisäaineen määrä on normaaliannoksessa tuotetta A satunnaismuuttuja X, joka on $\mathrm{N}(10\mathrm{mg}, (5\mathrm{mg})^2)$ ja normaaliannoksessa tuotetta B satunnaismuuttuja Y, joka on $\mathrm{N}(15\mathrm{mg}, (5\mathrm{mg})^2)$. Määrää:

(a) Luonaalla saamasi lisäaineen kokonaismäärän Z jakauma.
(b) $\mathrm{P}(20\mathrm{mg} \le Z \le 30\mathrm{mg})$.
(c) Luku z siten, että $\mathrm{P}(-z \le Z \le +z) = 0.98$.

41. Erään tölkitetyn tuoremehun sisältämän C-vitamiinin määrä on litrassa mehua satunnaismuuttuja X, joka on $\mathrm{N}(350\mathrm{mg}, (10\mathrm{mg})^2)$. Juot mehua 2dl/päivä. Mikä on todennäköisyys, että saat viikossa C-vitamiiniannoksen, joka ylittää 500mg?

42. Olkoon satunnaismuuttuja X poikien lukumäärä 1000 syntyneen lapsen joukossa. Määrää:

(a) $\mathrm{P}(X \le 480)$.
(b) $\mathrm{P}(485 \le X \le 510)$.

Oletetaan, että P(Syntyy poika) = P(Syntyy tyttö) = 0.5.

43. Ostat 12 arpaa arpajaisissa, joissa luvataan, että joka toinen arpa voittaa. Olkoon satunnaismuuttuja X voittojen lukumäärä. Määrää $\mathrm{P}(5 \le X \le 7)$:

(a) Binomijakaumasta.
(b) Normaalijakaumasta käyttäen jatkuvuuskorjausta ja ilman sitä.

LUKU 2.

1. Eräässä 15 henkilön ryhmästä 6 henkilöä kannattaa 5. ydinvoimalan rakentamista. Ryhmästä valitaan palauttaen otos, jonka koko on 5. Olkoon satunnaismuuttuja X rakentamisen kannattajien lukumäärä otoksessa. Määrää $P(X = 1)$, $P(X < 5)$, $E(X)$, $D(X)$.

2. Kuten tehtävä 1, mutta otos poimitaan palauttamatta.

3. Perusjoukon muodostaa 4 opiskelijapoikaa, joiden painot ovat 70kg, 75kg, 80kg ja 90kg. Olkoon satunnaismuuttuja X perusjoukosta satunnaisesti valitun pojan paino. Määrää X:n jakauma, odotusarvo ja varianssi.

4. Jatkoa tehtävälle 3. Perusjoukosta poimitaan palauttaen yksinkertainen satunnaisotos, jonka koko on 2. Määrää otokseen poimittujen poikien painojen keskiarvon jakauma ja esitä se myös graafisesti. Määrää myös keskiarvon odotusarvo ja varianssi.

5. Erään hedelmäsäilykepurkin sisältämä sokerimäärä on satunnaismuuttuja X, joka on $N(150\text{g}, (10\text{g})^2)$. Poimitaan purkkien joukosta palauttaen yksinkertainen satunnaisotos, jonka koko on (a) 25, (b) 100. Määrää $P(146\text{g} \le \overline{X} \le 154\text{g})$. Vertaa tuloksia toisiinsa.

6. Murojen pakkaaja väittää, että joka kolmannessa pakkauksessa on eräs lasten suuresti himoitsema lelu. Kuluttujansuojavirasto tutkii 600 pakkausta ja lelu löytyy 160 pakkauksesta. Kuinka yleinen on tämä tapahtuma, jos murojen pakkaajan väite pätee?

7. Kunnan väestöstä 25% on saamelaisia. Kuntalaisten joukosta poimitaan satunnaisotos, jonka koko on 400. Olkoon satunnaismuuttuja P otokseen poimittujen saamelaisten suhteellinen osuus. Määrää $P(0.22 \le P \le 0.28)$.

8. Jatkoa tehtävälle 7. Kunnan töissä olevista asukkaista 50% saa toimeentulonsa kunnan tai valtion virkamiehinä. Olkoon satunnaismuuttuja Q otokseen poimittujen virkamiesten suhteellinen osuus. Määrää $P(0.47 \le Q \le 0.53)$. Vertaa tulosta tehtävän 7 tulokseen. Huomaa, että välit (0.22, 0.28) ja (0.47, 0.53) ovat saman mittaisia.

9. Erään sähkölampputyypin paloaika X on $N(1000\text{h}, (200\text{h})^2)$. Lamppujen joukosta tehdään yksinkertainen satunnaisotos, jonka koko on 100. Olkoon $\overline{X}$ otoksesta määrätty otoskeskiarvo. Määrää vakio b siten, että

   (a) $P(1000 - b < \overline{X} < 1000 + b) = 0.95$,
   (b) $P(1000 - b < \overline{X} < 1000 + b) = 0.99$.

10. Vaaleissa ehdokas A sai 55% äänistä. Mikä on todennäköisyys, että ennen vaaleja tehdyssä kyselyssä A:n kannattajiksi ilmoittautui alle 50% otoksesta, kun otoskoko oli

   (a) 100, (b) 1000?

11. Tiilitehtaalla valmistetaan tiiliä, joiden paino X on $N(5\text{kg}, (100\text{g})^2)$. Tiilet pakataan 200 kappaleen eriksi. Mikä on todennäköisyys, että erän kokonaispaino on

   (a) < 995kg, (b) > 1002kg?

12. Tehdas väittää, että korkeintaan 3% tehtaan tuotteista on viallisia. Mikä on todennäköisyys, että 500 tuotteen erässä on vähemmän kuin 13 viallista?

LUKU 3.

1. Jatkoa tehtävälle 2.9.

   (a) Mikä on otoskeskiarvon $\overline{X}$ odotusarvo ja varianssi, jos otoskoko on 100?
   (b) Mitä arvoa kohti otoskeskiarvo $\overline{X}$ tarkentuu, jos otoskoon annetaan kasvaa rajatta?
   (c) Mikä on otoskeskiarvon $\overline{X}$ jakauma, jos otoskoko on 1000?
   (d) Mikä on otosmediaanin Md odotusarvo ja varianssi, jos otoskoko on 200?

2. Jatkoa tehtävälle 2.10.

   (a) Mikä on otoksesta määrätyn suhteellisen frekvenssin $\hat{p}$ odotusarvo ja varianssi, jos otoskoko on 200?
   (b) Mitä arvoa kohti otoksesta määrätty suhteellinen frekvenssi $\hat{p}$ tarkentuu, jos otoskoon annetaan kasvaa rajatta?
   (c) Mikä on otoksesta määrätyn frekvenssin X jakauma, jos otoskoko on 100?
   (d) Mikä on otoksesta määrätyn suhteellisen frekvenssin $\hat{p}$ approksimatiivinen jakauma, jos otoskoko on 1000?

3. Olkoot $X_1,\ldots,X_n$ riippumattomia havaintoja, jotka noudattavat normaalijakaumaa $N(\mu,\sigma^2)$. Parametrin μ SU-estimaattori on $\overline{X}$. Johda parametrin σ^2 SU-estimaattori. Ohje: Etsi uskottavuusfunktion maksimi σ^2:n suhteen derivoimalla, kun sijoitat μ:n paikalle sen suurimman uskottavuuden estimaattorin. Totea myös, että olet saanut maksimin.

4. Johda binomijakauman $\mathrm{Bin}(n,p)$ parametrin p SU-estimaattori. Ohje: Etsi uskottavuusfunktion (= binomijakauman pistetodennäköisyysfunktio) maksimi derivoimalla. Totea myös, että olet saanut maksimin.

5. 25:stä satunnaisesti valitusta savukkeesta määrättiin CO-pitoisuudet (CO = hiilimonoksidi). Tulokseksi saatiin $\overline{X} = 9.7\text{mg}$ ja $s = 1.4\text{mg}$. Määrää

(a) 95%:n, (b) 99%:n

luottamusväli satunnaisesti valitun savukkeen CO-pitoisuuden odotusarvolle, jos oletetaan, että CO-pitoisuuden jakaumaa perusjoukossa voidaan pitää normaalisena.

6. Oletetaan, että tuotteen paino on normaalijakautunut. Tuotteista poimitusta otoksesta saatiin otoskeskiarvoksi $\overline{X} = 2550\text{g}$ ja otoshajonnaksi $s = 100\text{g}$. Määrää 95%:n luottamusväli painon odotusarvolle, kun otoskoko oli

(a) 10, (b) 100, (c) 1000.

7. Kohdistetaan tutkimus erään varuskunnan varusmiehiin ja tutkittavana varusmiesten ominaisuutena on varusmiesten kokonaispaino. Olkoon perusjoukon koko 1000. Varusmiesten joukosta poimitaan 50 alokkaan otos, johon tulleet varusmiehet punnitaan. Tulokseksi saadaan $\overline{X} = 75\text{kg}$ ja $s = 10\text{kg}$. Mikä on varusmiesten *kokonaispainon* 95%:n luottamusväli, jos oletetaan, että paino on jakautunut normaalisesti perusjoukossa?

8. 100:n suuruisessa otoksessa 55% äänioikeutetuista kannatti ehdokasta A. Määrää

(a) 90%:n, (b) 95%:n, (c) 99%:n

luottamusväli A:n todelliselle kannatusosuudelle.

9. Jatkoa tehtävälle 8. Kuinka suuri otos tarvitaan, että saadaan 99%:n varmuus siitä, että A tulee todella valituksi, jos A:n kannattajien osuudeksi otoksessa saadaan 55%? Tällä tarkoitetaan seuraavaa: Kuinka suuri otos tarvitaan, että 99%:n luottamusväli A:n todelliselle kannatukselle perusjoukossa on muotoa 55%±5%.

10. 40:ssä rahanheitossa saatiin 24 klaavaa. Määrää 95%:n luottamusväli klaavojen suhteelliselle osuudelle äärettömän pitkässä heittosarjassa.

11. Laboratorio selvittää erään savukemerkin nikotiinipitoisuutta. Arvioksi pitää olla niin tarkka, että 95%:n varmuudella voidaan päätellä, että otoksesta laskettu keskimääräinen nikotiinipitoisuus ei poikkea todellisesta nikotiinipitoisuudesta yli 1 mg. Aikaisempien tutkimusten perusteella tiedetään, että nikotiinipitoisuusmittauksissa otoshajonta on tavallisesti n. 5 mg. Mikä on tarvittava otoskoko?

12. Erään tuotteen valmistaja väittää, että tuotteista korkeintaan 5% on viallisia. Asiakas poimii otoksen, jonka koko on 200 ja löytää 19 viallista tuotetta. Onko valmistajan väite oikeutettu?

LUKU 4.

1. 600 erääseen vakavaan tautiin sairastunutta potilasta jaettiin satunnaisesti kahteen ryhmään A ja B, joissa kummassakin oli 300 potilasta. Ryhmälle A annettiin tautiin kehitettyä uutta lääkettä ja ryhmälle B vanhaa paljon käytettyä lääkettä. Ryhmässä A taudista parani 195 potilasta ja ryhmässä B 225 potilasta. Suosittelisitko uuden lääkkeen ottamista käyttöön tulosten perusteeella?

2. Jatkoa tehtävälle 1. Mitä tekisit, jos ryhmässä A taudista olisikin parantunut 225 potilasta ja ryhmässä B 195 potilasta?

3. Eräällä tilastotieteen kurssilla 1. välikokeen tulokset ilmenevät tehtävien lopussa esitetystä taulukosta. Testaa nollahypoteesia, että naisten ja miesten suorituksissa ei ole eroa.

4. Alkoholin vaikutusta erään suorituksen kestoaikaan tutkittiin vertailemalla selvien ja alkoholia nauttineiden suoritusaikoja. Tulokset ilmenevät tehtävien lopussa olevasta taulukosta. Testaa nollahypoteesia, että alkoholia nauttineiden ja selvien suoritusajoissa ei ole eroa käyttäen 1%:n merkitsevyystasoa sekä kaksisuuntaista vaihtoehtoista hypoteesia että yksisuuntaista vaihtoehtoista hypoteesia vastaan. Yksisuuntainen vaihtoehtoinen hypoteesi: Alkoholia nauttineiden suoritusajat ovat pitempiä kuin selvien.

5. Esitä kriittisten rajojen avulla testi eli päätössääntö tilanteessa, jossa testataan onko raha harhaton 100:n rahanheiton perusteella. Raha on harhaton, jos P(kruuna) = P(klaava) = 0.5. Käytä merkitsevyystasoa

(a) 0.05, (b) 0.01.

6. Tutkittaessa E.S.P.-kyvyn olemassaoloa (extra-sensory perception) erästä koehenkilöä pyydetään arvaamaan toisessa huoneessa nostettavien korttien väri (sininen tai punainen). Kortit nostetaan peräjälkeen hyvin sekoitetusta 50 kortin pakasta. Henkilö ei tiedä kuinka monta kummankin väristä korttia pakassa on. Oletetaan, että henkilö arvaa 30 korttia oikein. Testaa poikkeaako tulos puhtaasta arvaamisesta.

7. Eräässä USA:n collegessa todettiin, että 50:n aktiivisisesti urheilevan miesopiskelijan pituuden keskiarvo oli 88.2 tuumaa ja keskihajonta oli 2.5 tuumaa. Vastaavat luvut ei-urheileville 50:lle miesopiskelijalle olivat 87.5 tuumaa ja 2.8 tuumaa. Testaa nollahypoteesia, että urheilevien ja ei-urheilevien pituudet eivät poikkea toisistaan.

8. Jatkoa tehtävälle 7. Kuinka suuri otos on poimittava ryhmiin (yhtä monta kumpaankin), jotta havaittu 0.7 tuuman ero olisi merkitsevä

(a) 5%:n, (b) 1%:n

merkitsevyystasolla?

9. Alueella A 300:sta satunnaisotokseen poimituista äänioikeutetuista 56% kannatti ehdokasta X. Alueella B ehdokkaan X kannatus 200:n satunnaisotokseen poimitun äänioikeutetun joukossa oli 48%. Määrää testin P-arvo. Testaa 5%:n merkitsevyystasolla, että

(a) X:n kannatus eroaa alueilla A ja B,
(b) X:n kannatus on alueella A suurempaa kuin alueella B.

10. Tehtävissä 1,2,3,4,7,8 ja 9 on tahallisesti jätetty sanomatta eräs olennainen oletus, jonka on oltava voimassa, jotta testit voidaan suorittaa luennolla esitetyllä tavalla. Mikä on tämä oletus?

Tehtävä 3:

	lkm	keskiarvo	hajonta
naiset	61	24.4	3.9
miehet	71	23.2	3.8

Tehtävä 4:

	lkm	keskiarvo	hajonta
juoneet	50	2.9	1.1
selvät	40	2.4	0.6

11. Erään rokotukokeen tulokset on esitetty alla. Testaa hypoteesia, että rokotettujen ja rokottamattomien sairastuvuudessa ei ole eroa käyttäen suhteellisten osuuksien vertailuun tarkoitettua testiä.

12. Käytä tehtävän 11 rokotuskokeen tuloksiin χ^2-testiä, kun nollahypoteesi olettaa, että sairastuvuus ei riipu rokotuksesta.

13. Riipuvatko silmien ja hiusten väri toisistaan? Onko tummahiuksisilla tummat silmät ja vaaleahiuksisilla siniset silmät? Asian tutkimiseksi poimittiin 95 henkilön otos. Tutkimustulokset on esitetty alla olevassa taulukossa. Testaa hypoteesia, että hiusten väri ja silmien väri ovat toisistaan riippumattomia.

14. Henkilöille A ja B annetaan kummallekin noppa ja kumpaakin pyydetään heittämään sitä 120 kertaa. A ja B kertovat saaneensa heittojen tuloksena alla esitetyt silmäluvut.

(a) Ovatko sekä A että B rehellisiä kertoessaan heittojen tuloksista?
(b) Voidaanko B:n heittämää noppaa pitää symmetrisenä?

15. Henkilöille A ja B annetaan kummallekin noppa ja kumpaakin pyydetään heittämään sitä 120 kertaa. A ja B kertovat saaneensa heittojen tuloksena alla esitetyt silmäluvut. Onko mahdollista, että A ja B ovat käyttäneet samaa noppaa?

16. Alla oleva taulukko koskee USA:n äänioikeutettujen joukosta poimittua otosta. Otoksesta on määrätty äänioikeutettujen puoluekanta ja suhtautuminen käsiaseiden rajoituksiin. Ovatko puoluekanta ja suhtautuminen aserajoituksiin riippumattomia?

17. Eräässä sairaalassa synnytykset jakaantuivat vuonna 1995 eri kuukausille alla esitetyn taulukon mukaisesti. Jakaantuvatko synnytykset tasaisesti eri kuukausille?

18. Kauppatieteilijät lähettävät usein kyselyitä liikeyrityksille. Vaikuttaako yritysten koko vastausalttiuteen? Vastausalttiutta tutkittiin otoksen avulla, jonka koko oli 600 yritystä. Aineistoon liittyvä frekvenssitaulukko on annettu alla.

Tehtävä 11:

Sairastuminen → Rokotus ↓	Sairastui	Ei sairastunut
Rokotettiin	9	42
Ei rokotettu	17	28

Tehtävä 13:

Silmät → Hiukset ↓	Siniset	Ruskeat	Muun väriset
Vaaleat	32	14	6
Tummat	12	22	9

Tehtävä 14:

Silmäluku	1	2	3	4	5	6
A:n tulokset	19	21	19	21	19	21
B:n tulokset	12	16	20	17	22	33

Tehtävä 15:

Silmäluku	1	2	3	4	5	6
A:n tulokset	16	18	19	22	19	26
B:n tulokset	12	16	20	17	22	33

Tehtävä 16:

Suhtautuminen aserajoituksiin → Puoluekanta ↓	Puoltaa	Vastustaa	Ei kantaa
Demokraatti	110	64	26
Republikaani	90	116	14
Riippumaton	55	35	10

Tehtävä 17:

Kuukausi	1	2	3	4	5	6	7	8	9	10	11	12
Synnytysten lkm	95	105	95	105	90	95	105	110	105	100	95	100

Tehtävä 18:

Yrityskoko → Vastaaminen ↓	Pieni	Keskisuuri	Suuri
Vastasi	125	81	40
Ei vastannut	75	119	160

19. Eräässä tutkimuksessa pyrittiin osoittamaan, että älykkyyden ja rasististen asenteiden välillä vallitsee negatiivinen riippuvuus. Eräiltä henkilöiltä mitattiin sekä älykkyysosamäärä (ÄO) että rasististen asenteiden voimakkuutta kuvaava pistemäärä (RP). Pistemäärät on esitetty alla olevassa taulukossa. Laske Spearmanin järjestyskorrelaatiokertoimen arvo ja testaa nollahypoteesia, että älykkyys ja rasististen asenteiden voimakkuus ovat toisistaan riippumattomia.

20. Kuten tehtävä 19, mutta käytä Kendallin järjestyskorrelaatiokerrointa.

21. Vuonna 1986 Valtiotieteelliseen tiedekuntaan pyrkineiden tyttöjen joukosta poimittiin otos, josta tutkittiin pyrkineiden aloitus- ja koepisteet. Pisteet on esitetty alla olevassa taulukossa. Määrää sekä Spearmanin että Kendallin kertoimet ja testaa nollahypoteesia, että alku- ja koepisteet ovat riippumattomia.

22. Säilykepurkkien sisäpinta on käsitelty lakalla tarkoituksena estää metallin liukeneminen säilykkeeseen. Tehdas on kokeillut kahta lakkaa A ja B ja elintarvikelaboratorio on poiminut asiaa tutkiakseen kaksi riippumatonta satunnaisotosta mainituilla lakoilla käsitellyistä purkeista. Liuenneen metallin määrät (mg) on esitetty alla olevassa taulukossa. Testaa Mannin ja Whitneyn testillä ovatko liuenneet metallimäärät keskimäärin yhtäsuuria.

23. Kahden savukemerkin A ja B nikotiinipitoisuuksien mittaamiseksi poimittiin kaksi riippumatonta satunnaisotosta. Pitoisuudet (mg) on esitetty alla olevassa taulukossa. Testaa hypoteesia, että keskimääräiset pitoisuudet ovat samansuuruiset. Käytä:

(a) Testiä, jossa perujoukon variansseista ei tehdä yhtäsuuruusoletusta ja perusjoukon jakaumasta ei tehdä normaalisuusoletusta.

(b) Testiä, jossa varianssеista tehdään yhtäsuuruusoletus ja perusjakaumasta tehdään normaalisuusoletus.

Vertaa (a)- ja (b)-kohdan tuloksia keskenään.

24. Kuten tehtävä 23, mutta oletetaan, että otoskoko oli kummassakin otoksessa 500.

25. Eräässä kokeessa verrattiin kahta sademäärän mittaukseen käytettyä laitetta. Kummallakin laitteella mitattiin sademäärät 10 sadepäivän aikana ja sademäärät (mm) on esitetty alla olevassa taulukossa. Testaa hypoteesia, että mittarit toimivat samalla tavalla.

26. Verenpainelääkkeen testauksessa samojen potilaiden verenpaine mitattiin ennen ja jälkeen lääkkeen nauttimisen. Verenpaineet (mm/Hg) on esitetty alla olevassa taulukossa. Testaa hypoteesia, että lääke ei vaikuta verenpaineeseen.

Tehtävä 19:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ÄO	102	116	108	139	94	99	109	123	124	113	103	100	125	117
RP	101	84	76	26	98	102	66	32	59	83	54	112	47	43

Tehtävä 21:

Alkupist.	19.2	16.4	15.9	11.5	13.3	19.9	10.0	9.1	18.5	14.2	14.9
Koepist.	45	43	39	40	26	42	19	24	3	40	32

Tehtävä 22:

A	0.42	0.58	0.76	0.44	0.62	0.74	0.38	5.5

B	0.64	0.89	1.02	0.70	0.75	0.99	0.86

Tehtävät 23:

	Otos-keskiarvo	Otos-keskihajonta	Otos-koko
A	23.4	5.3	5
B	27.0	8.5	5

Tehtävä 25:

	1	2	3	4	5	6	7	8	9	10
A	1.38	9.69	0.39	1.42	0.54	5.94	0.59	2.63	2.44	0.56
B	1.42	10.37	0.39	1.46	0.55	6.15	0.61	2.69	2.68	0.53

Tehtävä 26:

	1	2	3	4	5	6	7	8
Ennen	128	176	110	149	183	136	118	158
Jälkeen	134	174	118	152	187	136	125	168

LUKU 5.

1. Muuttujien x ja Y havaitut arvot ovat

x	1	3	4	6	8	9	11	14
Y	1	2	4	4	5	7	8	9

Määrää regressiomallin $Y_i = \alpha + \beta x_i + \varepsilon_i$ regressiokertoimien PNS-estimaattorit.

2. Määrää tehtävässä 1 estimoidun regressiomallin sovitteet ja residuaalit.

3. Määrää tehtävässä 1 estimoidun regressiomallin jäännösvarianssin estimaatti ja selitysaste.

4. Piirrä tehtävän 1 havaintoja (X_i,Y_i) esittävään pistedigrammiin tehtävässä 1 estimoitu regressiosuora. Merkitse kuvioon sovitteita vastaavat pisteet $(X_i,\hat{Y}_i)$. Piirrä samaan kuvioon myös residuaaleja kuvaavat janat.

5. Testaa tehtävän 1 mallin kerrointa β koskevaa nollahypoteesia H_0: $\beta = 0$ käyttäen tavanomaista t-testiä.

6. Eräässä 50 kunnan otoksessa suhteellisen rikollisuuden (rikoksia per 1000 asukasta) ja asukastiheyden (asukasta per km^2) välinen Pearsonin tulomomenttikorrelaatiokertoimen arvo oli $r = 0.157$. Testaa hypoteesia, että ko. muuttujat ovat korreloimattomia.

7. Tehtävässä 4.3. verrattiin miesten ja naisten suorituksia erään tilastotieteen kurssin 1. välikokeessa. Tulokset olivat

	Lkm	Keskiarvo	Keskihajonta
Naiset	61	24.4	3.9
Miehet	71	23.2	3.8

Testaa hypoteesia, että varianssit ovat yhtä suuria. Jos tämä hypoteesi jää voimaan, muodosta t-testi koepisteiden odotusarvojen samuudelle, kun nollahypoteesi otetaan huomioon.

8. Tehtävässä 4.4. vertailtiin erään tehtävän suoritusaikoja alkoholia nauttineiden ja selvien välillä. Tulokset olivat

	Lkm	Keskiarvo	Keskihajonta
Juoneet	50	2.9	1.1
Selvät	40	2.4	0.6

Testaa hypoteesia, että varianssit ovat yhtä suuria. Jos tämä hypoteesi jää voimaan, muodosta t-testi koepisteiden odotusarvojen samuudelle, kun nollahypoteesi otetaan huomioon.

9. Teräsvaijereita valmistava tehdas ilmoittaa, että valmistettujen vaijereiden murtumalujuus on 1800kg. Testattaessa 50 vaijerin otosta keskimääräiseksi murtumalujuudeksi saadaan 1750 kg keskihajonnan ollessa 100kg. Testaa nollahypoteesia, että valmistajan väite murtumalujuudesta on oikea.

10. Eräässä tutkimuksessa tutkittiin rekistöröityjen muskeliveneiden lukumäärän ja veneiden tappamien merilehmien lukumäärän yhteyttä Floridassa 1977-90. Estimoitu malli oli muotoa

$$\underset{}{\text{Kuolleiden lkm}} = \underset{(7.41)}{-41.4} + \underset{(0.0129)}{0.125}\ \text{Veneiden lkm}$$

Testaa nollahypoteesia, että veneiden lukumäärä ei vaikuta merilehmien kuolleisuuteen. Ohje: Käytä t-testiä.

11. Myös regressioanalyysin tuloksista voidaan muodostaa varianssianalyysitaulukko. Edelliseen tehtävään liittyvä varianssianalyysitaulukko on

Vaihtelun lähde	Neliösumma	Vapausasteet	F-testisuure
Malli	1712	1	
Jäännös		12	
Kok.vaiht.	1931		

(a) Täytä taulukon puuttuvat kohdat.
(b) Muodosta F-testi hypoteesille: Ei regressiota.
(c) Vertaa sitä edellisen tehtävän t-testisuureeseen.
(d) Määrää mallin selitysaste.
(e) Määrää veneiden lukumäärän ja kuolleiden merilehmien lukumäärän välinen korrelaatiokerroin.

LUKU 6.

1. Seuraavassa taulukossa on esitetty 18-30 kuukauden ikäisten lasten keskimäärin saamat päivittäiset energiamäärät (kcal). Määrät on laskettu taulukossa mainituista kolmesta maasta poimituista riippumattomista yksinkertaisista satunnaisotoksista. Testaa hypoteesia, että ravinnon energiamäärät ovat ko. maissa yhtä suuria.

	Otoskoko	Keskiarvo	Keskihajonta
Egypti	88	1217	327
Kenia	91	844	184
Meksiko	54	1119	285

2. Eräässä tutkimuksessa verrattiin miespuolisten opiskelijoiden kuntoa 4:ssä ryhmässä eräällä kuntoindeksillä. Ryhmät olivat:

1:	Erään kunto-ohjelman läpikäyneet	(10 opiskelijaa)
2:	Em. ohjelman aloittaneet, mutta keskeyttäneet	(5 opiskelijaa)
3:	Säännöllisesti hölkkäävät	(11 opiskelijaa)
4:	Täysin urheilemattomat	(10 opiskelijaa)

Opiskelijoista saaduista tiedoista voitiin muodostaa seuraava varianssianalyysitaulukko:

Vaihtelun lähde	Neliösumma	Vapauasteet	F-testisuure
Ryhmät	104855	3	
Jäännös	70501	32	
Kok.vaiht.			

(a) Täytä taulukon puuttuvat kohdat.
(b) Mikä on nollahypoteesi?
(c) Mikä on F-testisuureen jakauma, jos nollahypoteesi pätee?
(d) Testaa nollahypoteesia käyttäen 5%:n merk.tasoa.

3. Eräässä tutkimuksessa haluttiin verrata kolmea menetelmää, joita oli ehdotettu käytettäväksi peruskoulussa lukutaidon opettamisessa. Tutkimukseen valittiin 66 lasta, jotka jaettiin satunnaisesti kolmeen saman suuruiseen ryhmään. Ennen opetuskokeilua haluttiin varmistautua siitä, että lasten kielellinen lahjakkuus ei ollut ryhmissä erilaista. Tätä varten ryhmät alistettiin erilliseen testiin, jossa mitattiin sanojen ymmärätämistä. Testin tulokset on ilmaistu seuraavassa varianssianalyysitaulukossa, jossa sarakkeessa Keskiarvo on kerrottu lahjakkuustestin keskimääräset pisteet ko. ryhmissä.

	Keskiarvo	Keskihajonta
Ryhmä 1	10.50	2.972
Ryhmä 2	9.727	2.694
Ryhmä 3	9.136	3.342

Testaa annettujen tietojen perusteella hypoteesia, että lasten kielellinen lahjakkuus on ryhmissä keskimäärin samanlaista.

4. Jatkoa tehtävälle 3. Lukutaidon opettamisen jälkeen lapset alistettiin testiin, jossa mitattiin luetun ymmärtämistä. Testin tulokset on kerrottu seuraavassa taulukossa:

	Keskiarvo	Keskihajonta
Ryhmä 1	41.05	5.636
Ryhmä 2	46.73	7.388
Ryhmä 3	44.27	5.767

Testaa annettujen tietojen perusteella hypoteesia, että lasten luetun ymmärtämisen taju on ryhmissä keskimäärin samanlaista.

5. Mitkä ovat seuraavissa esimerkeissä vastemuuttujat, selittävät tekijät, selittävien tekijöiden tasojen lukumäärät ja havaintojen lukumäärät?

(a) Tomaatin kasvatusta koskevassa kokeessa koepuutarha vertailee kahden eri lannoitesekoitteen vaikutusta neljän tomattilajikkeen satoisuuteen. Kustakin lajikkeesta kasvatetaan viisi tainta kummakin lannoitteella. Kokeessa rekistöröidään taimien tuottama tomaattisato kiloissa.

(b) Markkinatutkimuksessa verrataan pyykinpesuaineen pakkauksia. Vertailtavana on kuusi erilaista pakkausta. Pakkausten vetovoimaa ostajiin kartoitetaan kyselyllä neljässä osassa maata. Jokainen pakkaus näytetään 30 kuluttajalle jokaisessa osassa maata. Vertailussa kuluttajia pyydetään antamaan pakkauksille arvosanat asteikolla 1-10.

(c) Kokeessa vertaillaan kolmea laihdutusohjelmaa. Viisi ylipainoista naista ja viisi ylipainoista miestä valitaan satunnaisesti kuhunkin ohjelmaan. Ohjelman jälkeen rekistöröidään kunkin laihduttajan painonmuutos kiloissa.

6. Mitä mitta-asteikkoja edellisen tehtävän muuttujat noudattavat?

7. Eräässä tutkimuksessa verrattiin toisiinsa mies- ja naispuolisten matemaatikoiden ja fyysikoiden palkkoja. Tutkimus perustui 800 henkilön ositettuun otokseen, jossa kussakin ositteessa oli 200 henkilöä. Palkkojen keskiarvot (ylempi luku) ja keskihajonnat (alempi luku) on annettu seuraavassa taulukossa:

Ka Kh	Sukupuoli	
Ala	Naiset	Miehet
Matematiikka	41.2 12.6	48.6 15.2
Fysiikka	34.7 16.9	42.3 15.9

Tee taulukon tietojen perusteella 2-suuntainen varianssianalyysi, ts. formuloi ja testaa hypoteesi H_{AB} sekä (jos tarvitaan) hypoteesit H_A ja H_B.

8. Eräässä tutkimuksessa verrattiin toisiinsa mies- ja naispuolisten biologien ja lääkäreiden palkkoja. Tutkimus perustui 800 henkilön ositettuun otokseen, jossa kussakin ositteessa oli 200 henkilöä. Palkkojen keskiarvot (ylempi luku) ja keskihajonnat (alempi luku) on annettu seuraavassa taulukossa. Tee taulukon tietojen perusteella kaksisuuntainen varianssianalyysi, ts. formuloi ja testaa hypoteesi H_{AB} sekä (jos tarvitaan) hypoteesit H_A ja H_B.

Ka Kh	Sukupuoli	
Ala	Naiset	Miehet
Biologia	34.5 12.6	42.0 15.2
Lääketiede	36.2 16.9	50.4 15.9

9. Eräässä tutkimuksessa verrattiin toisiinsa normaali- ja hyperaktiivisten lasten suorituskykyä meluisassa ja hiljaisessa luokassa. Suorituskykyä mitattiin menestymisellä matematiikan kokeessa. Tutkimus perustui 800 lapsen ositettuun otokseen, jossa kussakin ositteessa oli 200 lasta. Kokeesta saatujen pisteiden keskiarvot (ylempi luku) ja keskihajonnat (alempi luku) on annettu seuraavassa taulukossa:

Ka Kh	Luokka	
Ryhmä	Meluisa	Hiljainen
Normaali aktiivinen	214 126	170 152
Hyper-aktiivinen	120 169	140 159

Tee taulukon tietojen perusteella 2-suuntainen varianssianalyysi, ts. formuloi ja testaa hypoteesi H_{AB} sekä (jos tarvitaan) hypoteesit H_A ja H_B.

10. Kromin määrä ruokavaliossa vaikuttaa insuliinin tuotantoon. Eräässä kokeessa urosrotat alistettiin neljään erilaiseen ruokavalioon. Ruokavaliot erosivat toisistaan kahden eri tekijän suhteen: Kromin määrällä oli kaksi eri tasoa: matala (M) ja normaali (N). Samoin rotille tarjotun ravinnon määrällä oli kaksi tasoa: ei ylärajaa (E) ja rajoitettu määrä (R). Kokeessa rekistöröitiin erään entsymin (GITH) määrä veressä. Kokeen kohteina oli 40 rottaa. Keskimääräiset entsymin määrät on esitetty seuraavassa taulukossa:

Ka	Ravinnonmäärä	
Kromin määrä	E	R
M	4.545	5.175
N	4.425	5.317

(a) Kuvaa sanallisesti koetuloksia.

(b) Määrää ryhmäkeskiarvot ja yleiskeskiarvo.

11. Edelliseen tehtävään liittyvän varianssianalyysitaulukon neliösummista osa on ilmoitettu seuraavassa taulukossa:

Vaihtelun lähde	Neliösumma
A (kromin määrä)	0.00121
B (ravinnon määrä)	5.79121
AB	0.17161
Virhe	1.08084

(a) Määrää taulukon puuttuvat osat (kokonaisneliösumma, vapausasteet, varianssiestimaattorit ja F-testisuureet).

(b) Mikä havaintojen hajonta yhdistetyssä otoksessa?

(c) Tee täydennetyn taulukon tietojen perusteella 2-suuntainen varianssianalyysi, ts. formuloi ja testaa hypoteesi H_{AB} sekä (jos tarvitaan) hypoteesit H_A ja H_B.

LIITTEET

LIITE 1. KREIKAN KIELEN AAKKOSET.

Iso kirjain	Pieni kirjain	Nimi	Vastine
Α	α	alpha	a
Β	β	beta	b
Γ	γ	gamma	g
Δ	δ	delta	d
Ε	ε	epsilon	e
Ζ	ζ	dzeta	z
Η	η	eta	e
Θ	θ, ϑ	theta	th
Ι	ι	iota	i
Κ	κ	kappa	k
Λ	λ	lambda	l
Μ	μ	my	m
Ν	ν	ny	n
Ξ	ξ	ksi	ks
Ο	ο	omikron	o
Π	π	pi	p
Ρ	ρ	rho	r
Σ	σ	sigma	s
Τ	τ	tau	t
Υ	υ	ypsilon	y
Φ	φ, ϕ	phi	f
Χ	χ	khi	kh
Ψ	ψ	psi	ps
Ω	ω	omega	o

LIITE 2. SATUNNAISLUKUJEN TAULUKKO.

Ensimmäinen sarake (lihavoidut numerot) muodostuu rivinumeroista.
Satunnaisluvut on ryhmitelty 5:n numeron ryhmiin lukemisen helpottamiseksi.

1	33395	29691	07067	16478	90450	95214	68094	97921	65852	04465
2	90058	91876	42153	72260	29977	39646	75278	59845	75633	80904
3	68668	59073	16382	23456	18659	17936	94624	99933	95957	65811
4	58401	38021	17034	82756	96436	33036	73526	87135	30307	75090
5	42737	32135	67674	98210	43269	15147	57414	50907	79850	85209
6	03534	10891	76167	97748	17619	95704	83934	37042	94071	10254
7	06337	76318	65560	05348	65491	01949	93097	01020	96961	90710
8	97402	98282	93426	06134	51988	40832	98368	48989	76179	36277
9	38660	27385	56831	87850	17681	19145	31301	31718	98399	46367
10	07458	78817	33024	38108	90662	96114	92733	97840	94233	45095
11	63913	97894	54475	90352	31047	75409	78312	07675	28335	98064
12	57900	40263	84583	50395	90543	26345	61073	71694	10981	32069
13	24391	47015	74120	96760	98433	01274	87371	92459	53986	18023
14	26955	37524	19646	81158	24612	11712	84077	71173	65271	88008
15	15924	00409	60634	33579	18839	75072	24833	12647	95206	95186
16	75914	18465	31818	17933	30878	39869	88936	06037	44389	34665
17	22545	44486	88590	57749	22108	02189	16925	03723	18980	17274
18	12366	17769	93126	50697	30788	52467	51701	03245	58519	12286
19	71934	44339	59972	84450	71774	26425	52437	71661	71564	65610
20	58794	90306	96337	77979	02348	44075	44301	98233	15277	31982
21	56961	95118	71972	70698	81462	83314	59639	95250	16927	90613
22	83196	27983	31808	13245	14026	75392	60650	48857	72325	79525
23	73052	16601	25404	44423	96782	46704	86693	70823	86457	31494
24	44523	73469	49077	11204	37698	02440	35121	94407	96207	08221
25	28626	90631	34072	74751	55759	58337	73931	17489	62719	35255

LIITE 3.1. STANDARDOITU NORMAALIJAKAUMA N(0,1).

Kertymäfunktion $\Phi(z)$ arvoja argumentin z eri arvoilla.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
–0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
–0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
–0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
–0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
–0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
–0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
–0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
–0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
–0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
–0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
–1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
–1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
–1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
–1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
–1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
–1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
–1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
–1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
–1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002

Esimerkki: Jos $z = -2.23$, niin $\Phi(z) = 0.0129$.

LIITE 3.2. STANDARDOITU NORMAALIJAKAUMA N(0,1).

Kertymäfunktion Φ(*z*) arvoja argumentin *z* eri arvoilla.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

Esimerkki: Jos $z = +0.49$, niin $\Phi(z) = 0.6879$.

LIITE 4. STUDENTIN t-JAKAUMA.

Kriittisiä arvoja eri merkitsevyystasojen ja vapausasteiden f arvoilla.

	Merkitsevyystaso yksisuuntaisessa testissä					
f	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
1	6.314	12.706	31.821	63.657	318.309	636.619
2	2.920	4.303	6.965	9.925	22.327	31.599
3	2.353	3.182	4.541	5.841	10.215	12.924
4	2.132	2.776	3.747	4.604	7.173	8.610
5	2.015	2.571	3.365	4.032	5.893	6.869
6	1.943	2.447	3.143	3.707	5.208	5.959
7	1.895	2.365	2.998	3.499	4.785	5.408
8	1.860	2.306	2.896	3.355	4.501	5.041
9	1.833	2.262	2.821	3.250	4.297	4.781
10	1.812	2.228	2.764	3.169	4.144	4.587
11	1.796	2.201	2.718	3.106	4.025	4.437
12	1.782	2.179	2.681	3.055	3.930	4.318
13	1.771	2.160	2.650	3.012	3.852	4.221
14	1.761	2.145	2.624	2.977	3.787	4.140
15	1.753	2.131	2.602	2.947	3.733	4.073
16	1.746	2.120	2.583	2.921	3.686	4.015
17	1.740	2.110	2.567	2.898	3.646	3.965
18	1.734	2.101	2.552	2.878	3.610	3.922
19	1.729	2.093	2.539	2.861	3.579	3.883
20	1.725	2.086	2.528	2.845	3.552	3.850
21	1.721	2.080	2.518	2.831	3.527	3.819
22	1.717	2.074	2.508	2.819	3.505	3.792
23	1.714	2.069	2.500	2.807	3.485	3.768
24	1.711	2.064	2.492	2.797	3.467	3.745
25	1.708	2.060	2.485	2.787	3.450	3.725
26	1.706	2.056	2.479	2.779	3.435	3.707
27	1.703	2.052	2.473	2.771	3.421	3.690
28	1.701	2.048	2.467	2.763	3.408	3.674
29	1.699	2.045	2.462	2.756	3.396	3.659
30	1.697	2.042	2.457	2.750	3.385	3.646
40	1.684	2.021	2.423	2.704	3.307	3.551
50	1.676	2.009	2.403	2.678	3.261	3.496
60	1.671	2.000	2.390	2.660	3.232	3.460
80	1.664	1.990	2.374	2.639	3.195	3.416
100	1.660	1.984	2.364	2.626	3.174	3.390
200	1.653	1.972	2.345	2.601	3.131	3.340
500	1.648	1.965	2.334	2.586	3.107	3.310
∞	1.645	1.960	2.326	2.576	3.090	3.291
f	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
	Merkitsevyystaso kaksisuuntaisessa testissä					

LIITE 5. χ^2-JAKAUMA.

Yksisuuntaiseen testiin liittyviä kriittisiä arvoja eri merkitsevyystasojen ja vapausasteiden f arvoilla.

	Merkitsevyystaso yksisuuntaisessa testissä					
f	**0.99**	**0.95**	**0.10**	**0.05**	**0.01**	**0.001**
1	0.000	0.004	2.706	3.841	6.635	10.828
2	0.020	0.103	4.605	5.991	9.210	13.816
3	0.115	0.352	6.251	7.815	11.345	16.266
4	0.297	0.711	7.779	9.488	13.277	18.467
5	0.554	1.145	9.236	11.070	15.086	20.515
6	0.872	1.635	10.645	12.592	16.812	22.458
7	1.239	2.167	12.017	14.067	18.475	24.322
8	1.646	2.733	13.362	15.507	20.090	26.124
9	2.088	3.325	14.684	16.919	21.666	27.877
10	2.558	3.940	15.987	18.307	23.209	29.588
11	3.053	4.575	17.275	19.675	24.725	31.264
12	3.571	5.226	18.549	21.026	26.217	32.909
13	4.107	5.892	19.812	22.362	27.688	34.528
14	4.660	6.571	21.064	23.685	29.141	36.123
15	5.229	7.261	22.307	24.996	30.578	37.697
16	5.812	7.962	23.542	26.296	32.000	39.252
17	6.408	8.672	24.769	27.587	33.409	40.790
18	7.015	9.390	25.989	28.869	34.805	42.312
19	7.633	10.117	27.204	30.144	36.191	43.820
20	8.260	10.851	28.412	31.410	37.566	45.315
21	8.897	11.591	29.615	32.671	38.932	46.797
22	9.542	12.338	30.813	33.924	40.289	48.268
23	10.196	13.091	32.007	35.172	41.638	49.728
24	10.856	13.848	33.196	36.415	42.980	51.179
25	11.524	14.611	34.382	37.652	44.314	52.620
26	12.198	15.379	35.563	38.885	45.642	54.052
27	12.879	16.151	36.741	40.113	46.963	55.476
28	13.565	16.928	37.916	41.337	48.278	56.892
29	14.256	17.708	39.087	42.557	49.588	58.301
30	14.953	18.493	40.256	43.773	50.892	59.703
40	22.164	26.509	51.805	55.758	63.691	73.402
50	29.707	34.764	63.167	67.505	76.154	86.661
60	37.485	43.188	74.397	79.082	88.379	99.607
70	45.442	51.739	85.527	90.531	100.425	112.317
80	53.540	60.391	96.578	101.879	112.329	124.839
90	61.754	69.126	107.565	113.145	124.116	137.208
100	70.065	77.929	118.498	124.342	135.807	149.449
200	156.432	168.279	226.021	233.994	249.445	267.541
500	429.388	449.147	540.930	553.127	576.493	603.446

LIITE 6.1.1. F-JAKAUMA.

Yksisuuntaiseen testiin liittyviä kriittisiä arvoja 5%:n merkitsevyystasolla ja eri vapausasteiden f_1 ja f_2 arvoilla.

	f_1									
f_2	1	2	3	4	5	6	7	8	9	10
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543	241.882
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320	2.275
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265	2.220
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250	2.204
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236	2.190
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223	2.177
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993
70	3.978	3.128	2.736	2.503	2.346	2.231	2.143	2.074	2.017	1.969
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951
90	3.947	3.098	2.706	2.473	2.316	2.201	2.113	2.043	1.986	1.938
100	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927
200	3.888	3.041	2.650	2.417	2.259	2.144	2.056	1.985	1.927	1.878
500	3.860	3.014	2.623	2.390	2.232	2.117	2.028	1.957	1.899	1.850
∞	3.841	2.996	2.605	2.372	2.214	2.099	2.010	1.938	1.880	1.831

LIITE 6.1.2. F-JAKAUMA.

Yksisuuntaiseen testiin liittyviä kriittisiä arvoja 5%:n merkitsevyystasolla ja eri vapausasteiden f_1 ja f_2 arvoilla (jatkoa).

	f_1									
f_2	12	15	20	25	30	40	60	80	120	∞
1	243.906	245.950	248.013	249.260	250.095	251.143	252.196	252.724	253.253	254.314
2	19.413	19.429	19.446	19.456	19.462	19.471	19.479	19.483	19.487	19.496
3	8.745	8.703	8.660	8.634	8.617	8.594	8.572	8.561	8.549	8.526
4	5.912	5.858	5.803	5.769	5.746	5.717	5.688	5.673	5.658	5.628
5	4.678	4.619	4.558	4.521	4.496	4.464	4.431	4.415	4.398	4.365
6	4.000	3.938	3.874	3.835	3.808	3.774	3.740	3.722	3.705	3.669
7	3.575	3.511	3.445	3.404	3.376	3.340	3.304	3.286	3.267	3.230
8	3.284	3.218	3.150	3.108	3.079	3.043	3.005	2.986	2.967	2.928
9	3.073	3.006	2.936	2.893	2.864	2.826	2.787	2.768	2.748	2.707
10	2.913	2.845	2.774	2.730	2.700	2.661	2.621	2.601	2.580	2.538
11	2.788	2.719	2.646	2.601	2.570	2.531	2.490	2.469	2.448	2.404
12	2.687	2.617	2.544	2.498	2.466	2.426	2.384	2.363	2.341	2.296
13	2.604	2.533	2.459	2.412	2.380	2.339	2.297	2.275	2.252	2.206
14	2.534	2.463	2.388	2.341	2.308	2.266	2.223	2.201	2.178	2.131
15	2.475	2.403	2.328	2.280	2.247	2.204	2.160	2.137	2.114	2.066
16	2.425	2.352	2.276	2.227	2.194	2.151	2.106	2.083	2.059	2.010
17	2.381	2.308	2.230	2.181	2.148	2.104	2.058	2.035	2.011	1.960
18	2.342	2.269	2.191	2.141	2.107	2.063	2.017	1.993	1.968	1.917
19	2.308	2.234	2.155	2.106	2.071	2.026	1.980	1.955	1.930	1.878
20	2.278	2.203	2.124	2.074	2.039	1.994	1.946	1.922	1.896	1.843
21	2.250	2.176	2.096	2.045	2.010	1.965	1.916	1.891	1.866	1.812
22	2.226	2.151	2.071	2.020	1.984	1.938	1.889	1.864	1.838	1.783
23	2.204	2.128	2.048	1.996	1.961	1.914	1.865	1.839	1.813	1.757
24	2.183	2.108	2.027	1.975	1.939	1.892	1.842	1.816	1.790	1.733
25	2.165	2.089	2.007	1.955	1.919	1.872	1.822	1.796	1.768	1.711
26	2.148	2.072	1.990	1.938	1.901	1.853	1.803	1.776	1.749	1.691
27	2.132	2.056	1.974	1.921	1.884	1.836	1.785	1.758	1.731	1.672
28	2.118	2.041	1.959	1.906	1.869	1.820	1.769	1.742	1.714	1.654
29	2.104	2.027	1.945	1.891	1.854	1.806	1.754	1.726	1.698	1.638
30	2.092	2.015	1.932	1.878	1.841	1.792	1.740	1.712	1.683	1.622
40	2.003	1.924	1.839	1.783	1.744	1.693	1.637	1.608	1.577	1.509
50	1.952	1.871	1.784	1.727	1.687	1.634	1.576	1.544	1.511	1.438
60	1.917	1.836	1.748	1.690	1.649	1.594	1.534	1.502	1.467	1.389
70	1.893	1.812	1.722	1.664	1.622	1.566	1.505	1.471	1.435	1.353
80	1.875	1.793	1.703	1.644	1.602	1.545	1.482	1.448	1.411	1.325
90	1.861	1.779	1.688	1.629	1.586	1.528	1.465	1.429	1.391	1.302
100	1.850	1.768	1.676	1.616	1.573	1.515	1.450	1.415	1.376	1.283
200	1.801	1.717	1.623	1.561	1.516	1.455	1.386	1.346	1.302	1.189
500	1.772	1.686	1.592	1.528	1.482	1.419	1.345	1.303	1.255	1.113
∞	1.752	1.666	1.571	1.506	1.459	1.394	1.318	1.274	1.221	1.003

LIITE 6.1.3. F-JAKAUMA.

Yksisuuntaiseen testiin liittyviä kriittisiä arvoja 1%:n merkitsevyystasolla ja eri vapausasteiden f_1 ja f_2 arvoilla.

	f_1									
f_2	1	2	3	4	5	6	7	8	9	10
1	4052.2	4999.5	5403.4	5624.6	5763.7	5859.0	5928.4	5981.1	6022.5	6055.9
2	98.503	99.000	99.166	99.249	99.299	99.333	99.356	99.374	99.388	99.399
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345	27.229
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158	10.051
6	13.745	10.925	9.780	9.148	8.746	8.466	8.260	8.102	7.976	7.874
7	12.246	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719	6.620
8	11.259	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911	5.814
9	10.561	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351	5.257
10	10.044	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942	4.849
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632	4.539
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388	4.296
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191	4.100
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030	3.939
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895	3.805
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780	3.691
17	8.400	6.112	5.185	4.669	4.336	4.102	3.927	3.791	3.682	3.593
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597	3.508
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523	3.434
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457	3.368
21	8.017	5.780	4.874	4.369	4.042	3.812	3.640	3.506	3.398	3.310
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346	3.258
23	7.881	5.664	4.765	4.264	3.939	3.710	3.539	3.406	3.299	3.211
24	7.823	5.614	4.718	4.218	3.895	3.667	3.496	3.363	3.256	3.168
25	7.770	5.568	4.675	4.177	3.855	3.627	3.457	3.324	3.217	3.129
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182	3.094
27	7.677	5.488	4.601	4.106	3.785	3.558	3.388	3.256	3.149	3.062
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120	3.032
29	7.598	5.420	4.538	4.045	3.725	3.499	3.330	3.198	3.092	3.005
30	7.562	5.390	4.510	4.018	3.699	3.473	3.304	3.173	3.067	2.979
40	7.314	5.179	4.313	3.828	3.514	3.291	3.124	2.993	2.888	2.801
50	7.171	5.057	4.199	3.720	3.408	3.186	3.020	2.890	2.785	2.698
60	7.077	4.977	4.126	3.649	3.339	3.119	2.953	2.823	2.718	2.632
70	7.011	4.922	4.074	3.600	3.291	3.071	2.906	2.777	2.672	2.585
80	6.963	4.881	4.036	3.563	3.255	3.036	2.871	2.742	2.637	2.551
90	6.925	4.849	4.007	3.535	3.228	3.009	2.845	2.715	2.611	2.524
100	6.895	4.824	3.984	3.513	3.206	2.988	2.823	2.694	2.590	2.503
200	6.763	4.713	3.881	3.414	3.110	2.893	2.730	2.601	2.497	2.411
500	6.686	4.648	3.821	3.357	3.054	2.838	2.675	2.547	2.443	2.356
∞	6.635	4.605	3.782	3.319	3.017	2.802	2.639	2.511	2.407	2.321

LIITE 6.1.4. F-JAKAUMA.

Yksisuuntaiseen testiin liittyviä kriittisiä arvoja 1%:n merkitsevyystasolla ja eri vapausasteiden f_1 ja f_2 arvoilla.

	f_1									
f_2	12	15	20	25	30	40	60	80	120	∞
1	6106.3	6157.3	6208.7	6239.8	6260.6	6286.8	6313.0	6326.2	6339.4	6365.9
2	99.416	99.433	99.449	99.459	99.466	99.474	99.482	99.487	99.491	99.499
3	27.052	26.872	26.690	26.579	26.505	26.411	26.316	26.269	26.221	26.125
4	14.374	14.198	14.020	13.911	13.838	13.745	13.652	13.605	13.558	13.463
5	9.888	9.722	9.553	9.449	9.379	9.291	9.202	9.157	9.112	9.020
6	7.718	7.559	7.396	7.296	7.229	7.143	7.057	7.013	6.969	6.880
7	6.469	6.314	6.155	6.058	5.992	5.908	5.824	5.781	5.737	5.650
8	5.667	5.515	5.359	5.263	5.198	5.116	5.032	4.989	4.946	4.859
9	5.111	4.962	4.808	4.713	4.649	4.567	4.483	4.441	4.398	4.311
10	4.706	4.558	4.405	4.311	4.247	4.165	4.082	4.039	3.996	3.909
11	4.397	4.251	4.099	4.005	3.941	3.860	3.776	3.734	3.690	3.602
12	4.155	4.010	3.858	3.765	3.701	3.619	3.535	3.493	3.449	3.361
13	3.960	3.815	3.665	3.571	3.507	3.425	3.341	3.298	3.255	3.165
14	3.800	3.656	3.505	3.412	3.348	3.266	3.181	3.138	3.094	3.004
15	3.666	3.522	3.372	3.278	3.214	3.132	3.047	3.004	2.959	2.868
16	3.553	3.409	3.259	3.165	3.101	3.018	2.933	2.889	2.845	2.753
17	3.455	3.312	3.162	3.068	3.003	2.920	2.835	2.791	2.746	2.653
18	3.371	3.227	3.077	2.983	2.919	2.835	2.749	2.705	2.660	2.566
19	3.297	3.153	3.003	2.909	2.844	2.761	2.674	2.630	2.584	2.489
20	3.231	3.088	2.938	2.843	2.778	2.695	2.608	2.563	2.517	2.421
21	3.173	3.030	2.880	2.785	2.720	2.636	2.548	2.503	2.457	2.360
22	3.121	2.978	2.827	2.733	2.667	2.583	2.495	2.450	2.403	2.305
23	3.074	2.931	2.781	2.686	2.620	2.535	2.447	2.401	2.354	2.256
24	3.032	2.889	2.738	2.643	2.577	2.492	2.403	2.357	2.310	2.211
25	2.993	2.850	2.699	2.604	2.538	2.453	2.364	2.317	2.270	2.169
26	2.958	2.815	2.664	2.569	2.503	2.417	2.327	2.281	2.233	2.131
27	2.926	2.783	2.632	2.536	2.470	2.384	2.294	2.247	2.198	2.097
28	2.896	2.753	2.602	2.506	2.440	2.354	2.263	2.216	2.167	2.064
29	2.868	2.726	2.574	2.478	2.412	2.325	2.234	2.187	2.138	2.034
30	2.843	2.700	2.549	2.453	2.386	2.299	2.208	2.160	2.111	2.006
40	2.665	2.522	2.369	2.271	2.203	2.114	2.019	1.969	1.917	1.805
50	2.562	2.419	2.265	2.167	2.098	2.007	1.909	1.857	1.803	1.683
60	2.496	2.352	2.198	2.098	2.028	1.936	1.836	1.783	1.726	1.601
70	2.450	2.306	2.150	2.050	1.980	1.886	1.785	1.730	1.672	1.540
80	2.415	2.271	2.115	2.015	1.944	1.849	1.746	1.690	1.630	1.494
90	2.389	2.244	2.088	1.987	1.916	1.820	1.716	1.659	1.598	1.457
100	2.368	2.223	2.067	1.965	1.893	1.797	1.692	1.634	1.572	1.427
200	2.275	2.129	1.971	1.868	1.794	1.694	1.583	1.521	1.453	1.279
500	2.220	2.075	1.915	1.810	1.735	1.633	1.517	1.452	1.377	1.164
∞	2.185	2.039	1.878	1.773	1.696	1.592	1.473	1.404	1.325	1.005