

---

Aki Taanila

# TILASTOLLISEN PÄÄTTELYN ALKEET

---

9.5.2011

# SISÄLLYS

0 JOHDANTO .....	1
1 TILASTOLLINEN PÄÄTTELY .....	2
1.1 Tiekartta .....	4
2 YHTÄ MUUTTUJAA KOSKEVA PÄÄTTELY .....	5
2.1 Keskiarvon luottamusväli .....	5
2.2 Keskiarvon testaus (t-testi).....	6
2.3 Prosenttiluvun luottamusväli.....	7
2.4 Prosenttiluvun testaus (z-testi).....	7
3 KAHDEN RYHMÄN VERTAILU .....	9
3.1 Riippumattomien otosten t-testi .....	9
3.2 Riippuvien otosten t-testi .....	10
3.3 Khiin neliö -riippumattomuustesti .....	10
4 KAHDEN MUUTTUJAN VÄLINEN RIIPPUVUUS .....	11
4.1 Korrelaatiokertoimen testaus .....	11
4.2 Khiin neliö -riippumattomuustesti .....	11
5 TÄRKEITÄ HUOMIOITA.....	13
5.1 p-arvon tulkinta .....	13
5.2 Tilastollinen merkitsevyys ja käytännön merkitsevyys .....	13
5.3 Normaalijakautuneisuus ja otoskoko 30 .....	14
6 EXCEL JA TILASTOLLINEN PÄÄTTELY .....	15

# 0 JOHDANTO

Tässä monisteessa käsitellään tilastollista päättelyä. Tilastollinen päättely tarkoittaa otoksesta laskettujen tulosten yleistämistä laajempaan perusjoukkoon.

Asiat käsitellään käytännön soveltajan näkökulmasta. Tilastollisen päättelyn menetelmien taustalla olevaa matematiikkaa (todennäköisyysjakaumia jne.) ei käsitellä.

Lähtötietoina lukijalta edellytetään aineistojen esittämiseen ja kuvailuun käytettävien menetelmien hallinta.

**Opiskele huolellisesti monisteen ensimmäinen luku, koska se antaa perustan myöhempien lukujen ymmärtämiselle.**

Monisteeseen liittyviä Excel-esimerkkejä ja -laskentapohjia löydät työkirjasta

<http://myy.haaga-helia.fi/~taaak/p/p.xlsx>

Monisteessa käsiteltäviin menetelmiin liittyvät laskutoimitukset voidaan suorittaa Excelillä. Laskutoimitukset olisi helpointa suorittaa tilasto-ohjelmalla (esimerkiksi PASW/SPSS), mutta tässä monisteessä ei anneta tilasto-ohjelman käyttöön liittyviä ohjeita.

Laajemman valikoiman tilastollisia testimenetelmiä ja niihin liittyviä tilasto-ohjelman (PASW/SPSS) käyttöohjeita löydät verkkosivustolta <http://myy.haaga-helia.fi/~taaak/p>

Tutustu myös muihin oppimateriaaleihini <http://myy.haaga-helia.fi/~taaak>

# 1 TILASTOLLINEN PÄÄTTELY

Tilastollinen päättely tarkoittaa perusjoukkoa koskevien päätelmien tekemistä perusjoukosta poimitun otoksen perusteella. Otoksesta laskettuja tuloksia ei voida suoraan yleistää laajempaa perusjoukkoa koskeviksi, vaan päättelyssä täytyy huomioida otantavirheestä aiheutuva epävarmuus. Tilastollinen päättely voi sisältää

- virhemarginaalien/luottamusvälien laskemista
- hypoteesien testausta.

**Tilastollisen päättelyn käyttöedellytyksenä on, että otos on valittu satunnaisesti asianmukaista otantamenetelmää käyttäen.**

## Otantavirhe

Otoksesta lasketut taulukot ja tunnusluvut kuvailevat otosta. Otoksen perusteella voidaan tehdä päätelmiä perusjoukosta jos otos on satunnaisesti valittu. Jos otosta ei ole valittu satunnaisesti, niin on syytä kutsua näytteeksi.

Otoksen perusteella tehtyihin päätelmiin liittyy otantavirheen aiheuttamaa epävarmuutta. Otantavirhe seuraa siitä, että otoksen kokoonpano riippuu sattumasta ja näin ollen otoksesta lasketut tulokset vaihtelevat satunnaisesti otoksesta toiseen. Otantavirhe on luonnollisesti sitä pienempi mitä suurempaa otosta käytetään.

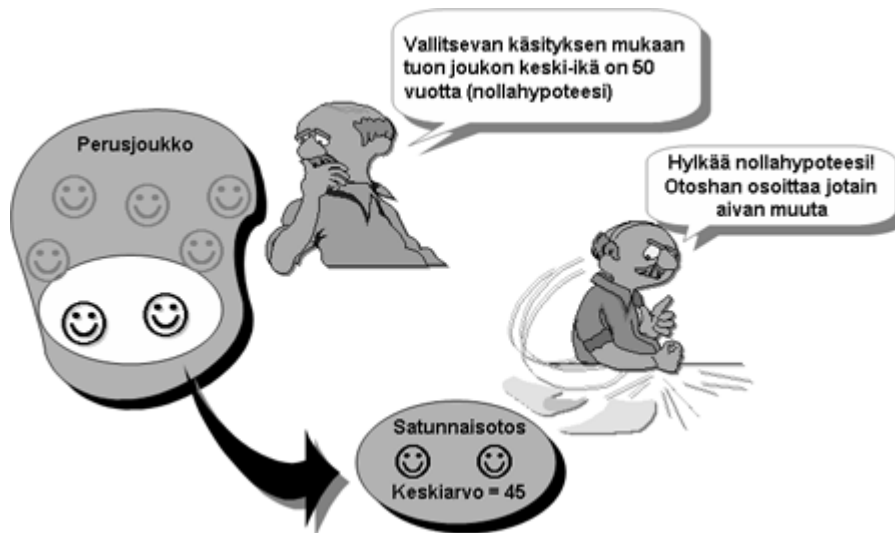
## Virhemarginaali/Luottamusväli

Jos haluat tietää perusjoukon tunnusluvun arvon ja käytössäsi on otos perusjoukosta, niin paras arvaus perusjoukon tunnusluvun arvoksi on otoksesta laskettu tunnusluku. Otantavirheen aiheuttaman epävarmuuden voit ilmaista virhemarginaalina. Yleensä ilmoitetaan 95 % virhemarginaali. Luottamusväliksi kutsutaan otoksesta lasketun tunnusluvun ympärille muodostettua väliä tunnusluku  $\pm$  virhemarginaali.

Esimerkki. Mielipidekyselyn mukaan 51 % suomalaisista kannattaa uuden ydinvoimalan rakentamista. Virhemarginaali on 3 prosenttiyksikköä. 95 % luottamusväli on siis 48 % - 54 %. Tämä tarkoittaa sitä, että 95 % varmuudella todellinen ydinvoiman kannatus on välillä 48 % - 54 %.

## Hypoteesin testaus

Hypoteesin testauksen lähtökohtana on nollahypoteesi, joka oletetaan oikeaksi, ellei otoksesta löydy todisteita sitä vastaan. Nollahypoteesi voi koskea esimerkiksi keskiarvon tai prosenttiluvun suuruutta. Tällöin nollahypoteesin lähteenä voi olla vallitseva käsitys, teoria, aikaisempi tutkimus, valmistajan ilmoitus jne. Tavallisimmin hypoteesi koskee ryhmien välistä eroa tai muuttujien välistä riippuvuutta. Tällöin nollahypoteesina on 'ei eroa' tai 'ei riippuvuutta'. Jos otos antaa riittävät todisteet nollahypoteesia vastaan, niin vaihtoehtoinen hypoteesi astuu voimaan.



Koska hypoteesin testaus perustuu otokseen, niin virhepäätelmän mahdollisuus on läsnä. Hypoteesin testauksessa toteutuu yksi seuraavista neljästä vaihtoehdosta:

- Nollahypoteesi on oikeasti totta ja testauksen tuloksena nollahypoteesi jää voimaan (oikea päätös).
- Nollahypoteesi ei oikeasti ole totta, mutta testauksen tuloksena nollahypoteesi jää voimaan (hyväksymisvirhe).
- Nollahypoteesi on oikeasti totta, mutta testauksen tuloksena nollahypoteesi päätetään hylätä (hylkäämisvirhe).
- Nollahypoteesi ei ole oikeasti totta ja testauksen tuloksena nollahypoteesi päätetään hylätä (oikea päätös).

	<b>Todellinen tilanne</b>	
<b>Testauksen tulos</b>	<b>Nollahypoteesi totta</b>	<b>Nollahypoteesi ei ole totta</b>
<b>Hyväksy nollahypoteesi</b>	Oikea päätös	Hyväksymisvirhe
<b>Hylkää nollahypoteesi</b>	Hylkäämisvirhe	Oikea päätös

Nollahypoteesi on perusolettamus ja se on syytä jättää voimaan ellei ole riittäviä todisteita sitä vastaan. Tämän vuoksi hylkäämisvirhettä (nollahypoteesi on oikeasti totta, mutta testauksen tuloksena se päätetään hylätä) pidetään vakavana virheenä, jota ei mielellään tehdä. Hylkäämisvirheen mahdollisuus on seurausta otantavirheestä ja hylkäämisvirheen todennäköisyys voidaan laskea. Hylkäämisvirheen todennäköisyyttä kutsutaan **p-arvoksi** tai havaituksi merkitsevyystasoksi.

Toinen tapa tulkita p-arvo on seuraava: p-arvo on todennäköisyys sille, että havaittu poikkeama nollahypoteesista on sattuman (otantavirheen) aiheuttama.

Yleisesti käytetty päättelysääntö on seuraavanlainen:

**Jos p-arvo on alle 0,05 (5 %), niin nollahypoteesi hylätään. Muussa tapauksessa nollahypoteesi jää voimaan.**

Hyväksymisvirheen todennäköisyyden laskeminen on vaikeampaa kuin hylkäämisvirheen todennäköisyyden. Kannattaa kuitenkin pitää mielessä, että mitä pienempää hylkäämisvirheen todennäköisyyttä vaaditaan nollahypoteesin hylkäämiseksi sitä suuremmaksi kasvaa hyväksymisvirheen todennäköisyys. Päättelysäännössä käytetty 5 % raja onkin kompromissi hylkäämisvirheen ja hyväksymisvirheen välillä. Käytännön tilanteesta riippuen voi-

daan raja asettaa muunkin suuruiseksi. Jos hylkäämisvirhe koetaan erityisen kohtalokkaaksi, niin rajaksi voidaan asettaa esimerkiksi 1 % tai 0,1 %. Jos taas halutaan helpommin erottaa poikkeamia nollahypoteesista, niin rajaksi voidaan asettaa esimerkiksi 10 %.

Jos sekä hylkäämisvirheen että hyväksymisvirheen todennäköisyyttä halutaan yhtä aikaa pienemmäksi, niin on turvauduttava isompaan otokseen.

Hypoteesin testaukseen liittyy tyypillisesti seuraavia vaiheita:

- Muotoile nollahypoteesi ja vaihtoehtoinen hypoteesi.
- Kerää havainnot (satunnaisesti valittu otos!).
- Laske hylkäämisvirheen todennäköisyys eli p-arvo.
- Päätteleysääntö: Hylkää nollahypoteesi, jos p-arvo on pienempi kuin 0,05 (5 %). Muussa tapauksessa nollahypoteesi jää voimaan. Ilmoita p-arvo perusteluna.

Esimerkki. Laakerien valmistaja vastaanottaa alihankkijalta ison erän laakerinkuulia, joiden halkaisijan pitäisi olla 5,30 millimetriä. Laakerien valmistaja haluaa tarkistaa, että vastaanotetut laakerinkuulat ovat sopivan kokoisia. Tätä varten asetetaan hypoteesit:

**Nollahypoteesi:** Kuulien halkaisijan keskiarvo on 5,30 millimetriä.

**Vaihtoehtoinen hypoteesi:** Kuulien halkaisijan keskiarvo on eri kuin 5,30 millimetriä.

Vastaanotetusta erästä valitaan 100 kappaleen otos. Otoksesta lasketaan halkaisijan keskiarvoksi 5,31 millimetriä ja keskihajonnaksi 0,10 millimetriä. Hylkäämisvirheen todennäköisyydeksi eli p-arvoksi saadaan noin 0,32. Näin suuri p-arvo (suurempi kuin 0,05) merkitsee sitä, että nollahypoteesi jää voimaan ja vastaanotettu laakerinkuulaerä voidaan hyväksyä.

Hylkäämisvirheen todennäköisyyden eli p-arvon laskentatapa vaihtelee testattavan hypoteesin mukaan. Tämän monisteen esimerkkien laskemiseen löydät laskentapohjia ja ohjeita Excel-työkirjasta <http://myy.haaga-helia.fi/~taaak/p/p.xlsx>.

## 1.1 Tiekartta

Seuraavassa taulukossa on esitetty menetelmät, joita käsitellään tässä monisteessa. Mukaan on valittu vain muutamia yleisimmin käytettyjä menetelmiä.

Tarkoitus	Muuttujan mitta-asteikko	
	Määrällinen	Kategorinen
Yhtä muuttujaa koskeva päättely	Keskiarvon luottamusväli Keskiarvon testaus (t-testi)	Prosenttiluvun luottamusväli Prosenttiluvun testaus (z-testi)
Kahden ryhmän vertailu	Kaksi riippumatonta otosta: -Riippumattomien otosten t-testi  Kaksi riippuvaa otosta: -Riippuvien otosten t-testi	Khiin neliö -riippumattomuustesti
Kahden muuttujan välinen riippuvuus	Korrelaatiokertoimen testaus	Khiin neliö -riippumattomuustesti

## 2 YHTÄ MUUTTUJAA KOSKEVA PÄÄTTELY

### 2.1 Keskiarvon luottamusväli

Keskiarvon virhemarginaalia ja luottamusväliä laskettaessa edellytetään, että otoskeskiarvo noudattaa likimain normaalijakaumaa. Jos otoskoko on yli 30, niin normaalijakautuneisuutta ei tarvitse erikseen tarkistaa.

Jos käytössä ei ole muuta tietoa kuin otoksesta laskettu keskiarvo, niin se on paras arvaus perusjoukon keskiarvoksi. Kun perusjoukon keskiarvo arvioidaan otoskeskiarvon suuruisiksi, niin arvioon liittyy epävarmuus. Epävarmuus ilmoitetaan virhemarginaalina. Yleensä ilmoitetaan 95 % virhemarginaali.

Esimerkki. Moottorien valmistaja vastaanottaa alihankkijalta ison erän moottorien osia, joiden pituuden pitäisi olla 156,0 millimetriä. Moottorien valmistajan laadunvalvontaosasto ottaa saapuneesta erästä 50 kappaleen otoksen. Otoksesta laskettu pituuden keskiarvo on 156,30 millimetriä ja keskihajonta 0,34 millimetriä.

Laskemalla saadaan 95 % virhemarginaaliksi 0,10 millimetriä. Moottorien osien pituuden keskiarvon 95 % luottamusväli on siis 156,2 mm – 156,4 mm. Otoksen perusteella voidaan todeta, että moottorien osien pituus ei ole tavoiteltu 156,0 mm.

Voit laskea virhemarginaalin helposti laskimella. Käytännön sovelluksissa saat 95 % virhemarginaalin riittävällä tarkkuudella laskemalla ( $\sigma$  on perusjoukon keskihajonta,  $n$  on otoskoko):

$$2 \times \frac{\sigma}{\sqrt{n}}$$

Jos perusjoukon keskihajontaa  $\sigma$  ei tunneta (useimmiten ei tunneta), niin sen tilalla voidaan käyttää otoksesta laskettua keskihajontaa. Tällöin yllä olevan kaavan käyttö edellyttää, että otoskoko on yli 30.

Esimerkki. Lampun kestoian keskiarvoksi saadaan 100 lampun otoksesta 2500 tuntia. Keskihajonnaksi saadaan 150 tuntia. Laskemalla saadaan 95 % virhemarginaaliksi noin 30 tuntia. Kyseisen lampputyypin kestoian keskiarvon 95 % luottamusväli on siis 2470 tuntia - 2530 tuntia.

Esimerkki. Annostelukoneen pitäisi pussittaa 500 gramman pusseja. Pussin painon keskiarvo 20 pussin otoksessa on 480,3 grammaa ja keskihajonta 20,0 grammaa. Laskemalla saadaan 95 % virhemarginaaliksi noin 9,4 grammaa. Tavoitearvo 500 grammaa ei mahdu luottamusvälin 471 g - 490 g sisään, joten annostelukone on luultavasti väärin säädetty.

Edellä annettu likimääräinen laskentakaava ei sovellu tähän esimerkkiin, koska otoskoko on vain 20.

Otos on huomattava osa perusjoukosta

Jos otoskoko on yli 5 % perusjoukon koosta, niin voit kertoa virhemarginaalin äärellisen perusjoukon korjauskertoimella ( $N$  = perusjoukon koko,  $n$  on otoskoko):

$$\sqrt{\frac{N-n}{N-1}}$$

Korjauskertoimen käyttö tuottaa pienemmän virhemarginaalin.

## 2.2 Keskiarvon testaus (t-testi)

Keskiarvon testauksessa edellytetään, että otoskeskiarvo noudattaa likimain normaalijakaumaa. Jos otoskoko on yli 30, niin normaalijakautuneisuutta ei tarvitse erikseen tarkistaa.

### Hypoteesit

Jos käytössä on ennakko-oletus, nollahypoteesi, perusjoukon keskiarvosta, niin otoksen keskiarvoa voidaan verrata nollahypoteesin mukaiseen arvoon. Nollahypoteesi voi pohjautua esimerkiksi vallitsevaan käsitykseen, teoriaan, aikaisempaan tutkimukseen, valmistajan ilmoitukseen jne. Nollahypoteesin rinnalle asetetaan vaihtoehtoinen hypoteesi. Keskiarvon kaksisuuntaisessa testauksessa asetetaan hypoteesit ( $x_0$  on jokin luku):

**Nollahypoteesi:** Perusjoukon keskiarvo on yhtä suuri kuin  $x_0$ .

**Vaihtoehtoinen hypoteesi:** Perusjoukon keskiarvo on eri suuri kuin  $x_0$ .

Jos ollaan kiinnostuneita vain poikkeamasta jompaankumpaan suuntaan, niin käytetään yksisuuntaista testiä. Tällöin vaihtoehtoinen hypoteesi muotoillaan tilanteen mukaisesti. Esimerkiksi:

**Vaihtoehtoinen hypoteesi:** Perusjoukon keskiarvo on pienempi kuin  $x_0$ .

Esimerkki. Pullotuskone pitäisi olla säädetty siten, että se pullottaa 1/3 litran pulloja. Laadun valvoja testaa toistuvilla otoksilla hypoteeseja:

**Nollahypoteesi:** Pullojen sisällön keskiarvo 1/3 litraa.

**Vaihtoehtoinen hypoteesi:** Pullojen sisällön keskiarvo eri suuri kuin 1/3 litraa.

Laadun valvoja ottaa pullotuslinjalta 15 pullon otoksen ja saa sisältöjen keskiarvoksi 0,3420 litraa ja keskihajonnaksi 0,0115 litraa. Kaksisuuntaisen t-testin p-arvoksi saadaan noin 0,011. Nollahypoteesi hylätään, koska p-arvo 0,011 on pienempi kuin 0,050. Tämä on myös riski sille, että mahdollinen pullotuslinjan pysäyttäminen ja säätöjen korjaaminen tehdään turhaan.

Esimerkki. Leipomo ilmoittaa leivän suolapitoisuudeksi 1,3 %. Kuluttajia edustavan järjestön tutkija asettaa hypoteesit:

**Nollahypoteesi:** Leipien keskimääräinen suolapitoisuus on 1,3 %

**Vaihtoehtoinen hypoteesi:** Leipien keskimääräinen suolapitoisuus on suurempi kuin 1,3 %.

Tutkija valitsee satunnaisesti 20 leivän otoksen. Laboratoriotutkimuksen perusteella saadaan suolapitoisuuden keskiarvoksi 1,5 % ja keskihajonnaksi 0,3 prosenttiyksikköä. Yksisuuntaisen t-testin p-arvoksi saadaan noin 0,004, joka on pienempi kuin 0,050. Näin ollen nollahypoteesi hylätään.

Kannattaa huomioida suuri keskihajonta. Leipien suolapitoisuus voi siis vaihdella suuresti leivästä toiseen. Näin suuri vaihtelu ei anna hyvää kuvaa leipomon työtavoista.

## 2.3 Prosenttiluvun luottamusväli

Prosenttiluvun virhemarginaalia ja luottamusväliä laskettaessa edellytetään, että  $np$  ja  $n(1-p)$  ovat molemmat suuruudeltaan vähintään 10 ( $n$ =otoskoko,  $p$ =prosenttiluku).

Jos käytössä ei ole muuta tietoa kuin otoksesta laskettu prosenttiluku, niin se on paras arvaus perusjoukon prosenttiluvuksi. Kun perusjoukon prosenttiluku arvioidaan otoksesta lasketun prosenttiluvun suuruiseksi, niin arvioon liittyy epävarmuus. Epävarmuus ilmoitetaan virhemarginaalina. Yleensä ilmoitetaan 95 % virhemarginaali.

Esimerkki. Otoksesta ( $n=1800$ ) laskettu viallisten tuotteiden osuus on 5,0 % ja virhemarginaali 1,0 prosenttiyksikköä. 95 % luottamusväli viallisten osuudelle on 4,0 % - 6,0 %.

Voit laskea virhemarginaalin helposti laskimella. Käytännön sovelluksissa saat 95 % virhemarginaalin riittävällä tarkkuudella laskemalla ( $p$  on otoksesta laskettu prosenttiluku desimaalimuodossa,  $n$  on otoskoko):

$$2 \times \sqrt{\frac{p(1-p)}{n}}$$

Esimerkki. Kyselytutkimuksen otoskoko  $n=1000$  henkilöä. Otoksesta 51 % ( $p = 0,51$ ) on uuden ydinvoimalan kannalla. Laskemalla saadaan virhemarginaaliksi noin 3 prosenttiyksikköä. Näin ollen 95 % luottamusväli ydinvoiman kannattajien osuudelle on 48 % - 54 %.

Otos on huomattava osa perusjoukosta

Jos otoskoko on yli 5 % perusjoukon koosta, niin voit kertoa virhemarginaalin äärellisen perusjoukon korjauskertoimella ( $N$  on perusjoukon koko,  $n$  on otoskoko):

$$\sqrt{\frac{N-n}{N-1}}$$

Korjauskertoimen käyttö tuottaa pienemmän virhemarginaalin.

## 2.4 Prosenttiluvun testaus (z-testi)

Prosenttiluvun z-testissä edellytetään, että  $np$  ja  $n(1-p)$  ovat molemmat suuruudeltaan vähintään 10 ( $n$ =otoskoko,  $p$ =prosenttiluku).

Hypoteesit

Jos käytössä on ennakko-oletus, nollahypoteesi, perusjoukon prosenttiluvusta, niin otoksen prosenttilukua voidaan verrata nollahypoteesin mukaiseen arvoon. Nollahypoteesi voi pohjautua esimerkiksi olemassa olevaan teoriaan, vallitsevaan käsitykseen, aikaisempaan tutkimukseen, valmistajan ilmoitukseen jne. Nollahypoteesin rinnalle asetetaan vaihtoehtoinen hypoteesi. Prosenttiluvun kaksisuuntaisessa testauksessa asetetaan hypoteesit ( $P_0$  on luku väliltä 0-100):

**Nollahypoteesi:** Perusjoukon prosenttiluku on yhtä suuri kuin  $P_0$  %.

**Vaihtoehtoinen hypoteesi:** Perusjoukon prosenttiluku on eri suuri kuin  $P_0$  %.

Jos ollaan kiinnostuneita vain poikkeamasta jompaankumpaan suuntaan, niin käytetään yksisuuntaista testiä. Tällöin vaihtoehtoinen hypoteesi muotoillaan tilanteen mukaisesti.

Esimerkiksi:

**Vaihtoehtoinen hypoteesi:** Perusjoukon prosenttiluku on pienempi  $P_0$  %.

Esimerkki. Puolueen kannatus oli aiemmin 22,8 %. Tutkimuslaitos asetti seuraavat hypoteesit:

**Nollahypoteesi:** Puolueen kannatus on 22,8 %.

**Vaihtoehtoinen hypoteesi:** Puolueen kannatus on laskenut aiemmasta (pienempi kuin 22,8 %).

Satunnaisesti valitussa 800 henkilön otoksessa puolueen kannattajia oli 166. Yksisuuntaisen testin p-arvoksi saadaan 0,076, joka on suurempi kuin 0,050. Nollahypoteesi jää voimaan.

On kuitenkin syytä huomata, että p-arvo 0,076 on melko lähellä hylkäämisrajaa 0,050. Tulosta voisi pitää suuntaa antavana sen puolesta, että kannatus on laskenut aiemmasta. Yleisemminkin p-arvoja 0,050 ja 0,100 välillä voidaan pitää suuntaa antavina.

## 3 KAHDEN RYHMÄN VERTAILU

Riippumattomat vai riippuvat otokset?

Jos otetaan satunnaisotokset kahdesta eri perusjoukosta, niin kyseessä on toisistaan riippumattomat otokset.

**Esimerkki.** Jos halutaan verrata kahdella eri menetelmällä A ja B valmistettujen lamppujen kestoikää, niin voidaan ottaa otos menetelmällä A valmistettuja lamppuja ja toinen otos menetelmällä B valmistettuja lamppuja.

Myös saman satunnaisotoksen sisällä olevia ryhmiä voidaan pitää riippumattomina.

**Esimerkki.** Jos yrityksen työntekijöistä otetaan satunnaisotos, niin voimme pitää otokseen sisältyviä naisia ja miehiä toisistaan riippumattomina otoksina (otos naisista ja otos miehistä).

Jos toistetaan mittaus samoille tutkittaville, niin mittauskerrat muodostavat toistaan riippuvat otokset.

**Esimerkki.** Jos mitataan samojen kuluttajien asennetta tuotteeseen ennen ja jälkeen tuote-esittelyn, niin kyseessä ovat toisistaan riippuvat otokset.

Toisistaan riippuvat otokset voidaan muodostaa myös käyttämällä toisiaan vastaavia pareja.

**Esimerkki.** Verrataan kahden akkutyypin kestoja matkapuhelimissa. Testiin valitaan useita matkapuhelinmalleja, kaksi kutakin. Kustakin matkapuhelinmallista muodostetaan pari, jotta päästään testaamaan kumpaakin akkutyyppeä kyseisessä matkapuhelinmallissa. Akkutyyppeihin liittyvät otokset ovat toisistaan riippuvat.

### 3.1 Riippumattomien otosten t-testi

Riippumattomien otosten t-testiä käytettäessä edellytetään, että molempien ryhmien otoskeskiarvot noudattavat normaalijakaumaa. Jos ryhmistä otetut otokset ovat molemmat suuruudeltaan yli 30, niin normaalijakautuneisuutta ei tarvitse erikseen tarkistaa.

**Hypoteesit**

Kaksisuuntaisessa testissä asetetaan hypoteesit:

**Nollahypoteesi:** Ryhmien keskiarvot ovat samat.

**Vaihtoehtoinen hypoteesi:** Ryhmien keskiarvot poikkeavat toisistaan.

Jos ollaan kiinnostuneita vain poikkeamasta tiettyyn suuntaan, niin käytetään yksisuuntaista testiä. Tällöin vaihtoehtoisessa hypoteesissa ilmaistaan minkä suuntaisesta erosta ollaan kiinnostuneita. Esimerkiksi:

**Vaihtoehtoinen hypoteesi:** Toisen ryhmän keskiarvo on suurempi.

**Esimerkki.** Lamppujen valmistaja valmistaa samantyyppisiä lamppuja kahdella eri menetelmällä. Tutkimus- ja kehittämisosasto valitsee 40 lampun otoksen kummastakin menetelmästä ja mittaa lamppujen kestoajat. Hypoteeseina ovat:

**Nollahypoteesi:** Kestoiän keskiarvo on sama molemmissa menetelmissä

**Vaihtoehtoinen hypoteesi:** Kestoiän keskiarvot ovat erisuuret eri menetelmissä.

Riippumattomien otosten kaksisuuntaisen t-testin p-arvoksi saadaan 0,006, joka on pienempi kuin 0,050. Tämän perusteella nollahypoteesi hylätään.

## 3.2 Riippuvien otosten t-testi

Riippuvien otosten t-testiä käytettäessä edellytetään, että parien välisten erojen otoskeskiarvo noudattaa normaalijakaumaa. Jos ryhmistä ototut otokset ovat kooltaan yli 30, niin normaalijakautuneisuutta ei tarvitse erikseen tarkistaa.

**Esimerkki.** Oletetaan, että henkilöiden reaktioajat mitataan selvänä ja yhden promillen humalassa. Tällöin käytössä on kaksi otosta: selvät ja promillen humalassa olevat. Otokset ovat toisistaan riippuvat, koska molemmissa otoksissa on mukana samat henkilöt.

### Hypoteesit

Kaksisuuntaisessa testissä asetetaan hypoteesit:

**Nollahypoteesi:** Ryhmien keskiarvot ovat samat.

**Vaihtoehtoinen hypoteesi:** Ryhmien keskiarvot poikkeavat toisistaan.

Jos tutkija on kiinnostunut vain poikkeamasta tiettyyn suuntaan, niin käytetään yksisuuntaista testiä. Tällöin vaihtoehtoisessa hypoteesissa ilmaistaan minkä suuntaisesta erosta ollaan kiinnostuneita. Esimerkiksi:

**Vaihtoehtoinen hypoteesi:** Toisen ryhmän keskiarvo on suurempi.

**Esimerkki.** Lääkäri testaa erikoisruokavaliota potilaille, joiden suvussa esiintyy perinnöllistä taipumusta sydänsairauksiin. Erityisruokavalion tarkoituksena on alentaa painoa ja sydänsairauksien kannalta haitallisten triglyseridien määrää elimistössä. Potilaiden paino ja triglyseridi arvot tutkitaan ennen ja jälkeen erityisruokavalion. Lääkäri asettaa hypoteesit:

**Nollahypoteesi:** Ei eroa painon keskiarvoissa.

**Vaihtoehtoinen hypoteesi:** Painon keskiarvo on ruokavalion jälkeen pienempi.

**Nollahypoteesi:** Ei eroa triglyseridin keskiarvoissa.

**Vaihtoehtoinen hypoteesi:** Triglyseridin keskiarvo on ruokavalion jälkeen pienempi.

## 3.3 Khiin neliö -riippumattomuustesti

Khiin neliö -riippumattomuustestillä voidaan verrata kahden ryhmän prosenttilukuja. Ryhmien vertailuun käytettävä testi on laskennallisesti täsmälleen sama kuin luvussa 4.2 esiteltävä khiin neliö -riippumattomuustesti.

# 4 KAHDEN MUUTTUJAN VÄLINEN RIIPPUVUUS

## 4.1 Korrelaatiokertoimen testaus

Pearsonin korrelaatiokertoimen kohdalla oletetaan kummankin muuttujan noudattavan likimain normaalijakaumaa millä tahansa toisen muuttujan arvolla. Usein oletuksesta käytetään lievempää versiota, jonka mukaan kumpikin muuttuja noudattaa normaalijakaumaa. Jos otoskoko on yli 30, niin normaalijakaumaoletusta ei tarvitse erikseen tarkistaa.

### Hypoteesit

Korrelaatiokertoimen kaksisuuntaisessa testauksessa asetetaan hypoteesit:

**Nollahypoteesi:** Perusjoukon korrelaatiokerroin on nolla.

**Vaihtoehtoinen hypoteesi:** Perusjoukon korrelaatiokerroin on nolosta poikkeava.

Jos ollaan kiinnostuneita tietyn merkkisestä (+ tai -) korrelaatiosta, niin käytetään yksisuuntaista testiä. Tällöin vaihtoehtoinen hypoteesi muotoillaan tilanteen mukaisesti. Esimerkiksi:

**Vaihtoehtoinen hypoteesi:** Perusjoukon korrelaatiokerroin on positiivinen.

Esimerkki. Tutkittiin asiakkaiden käsityksiä eri osa-alueista mielipideasteikolla 0 - 10. Erittymisenä kiinnostuksen kohteina oli eri osa-alueiden yhteys asiakkaan yleiseen tyytyväisyyteen. Esimerkiksi testattiin korrelaatiota muuttujien ”asiakkaan mielikuva toimitusnopeudesta” ja ”yleinen tyytyväisyys” välillä. Korrelaatiokertoimeksi saatiin 100 asiakkaan otoksesta 0,651.

**Nollahypoteesi:** Asiakkaiden mielikuva toimitusnopeudesta ei ole yhteydessä yleiseen tyytyväisyyteen (korrelaatiokerroin on 0).

**Vaihtoehtoinen hypoteesi:** Asiakkaiden mielikuva toimitusnopeudesta on positiivisessa yhteydessä yleiseen tyytyväisyyteen (korrelaatiokerroin on suurempi kuin 0).

Korrelaatiokertoimen ja otoskoon avulla lasketuksi p-arvoksi saatiin 0,000. Näin ollen nollahypoteesi hylätään ja voidaan todeta, että kyseessä on tilastollisesti merkitsevä korrelaatio.

## 4.2 Khiin neliö -riippumattomuustesti

Khiin neliö -riippumattomuustestillä voidaan testata kahden muuttujan välistä riippuvuutta. Khiin neliö -riippumattomuustestin käyttöedellytyksenä on,

- korkeintaan 20 % nollahypoteesin mukaisen jakauman lukumäärästä on pienempiä kuin 5.
- nollahypoteesin mukaisen jakauman lukumäärät ovat suuruudeltaan vähintään 1.

Nollahypoteesin mukainen jakauma tarkoittaa teoreettista jakaumaa, jossa riippuvuutta ei esiinny.

### Hypoteesit

Khiin neliö -riippumattomuustestin hypoteesit ovat:

**Nollahypoteesi:** Muuttujien välillä ei ole riippuvuutta.

**Vaihtoehtoinen hypoteesi:** Muuttujien välillä on riippuvuutta.

Esimerkki. Työntekijöistä otettiin satunnainen otos ja suoritettiin kyselytutkimus. Kyselyssä selvitettiin mm. vastaajan sukupuoli ja tyytyväisyys johtoon.

**Nollahypoteesi:** Sukupuolen ja tyytyväisyyden välillä ei ole riippuvuutta.

**Vaihtoehtoinen hypoteesi:** Sukupuolen ja tyytyväisyyden välillä on riippuvuutta.

Khiin neliö -testin p-arvo 0,017 on pienempi kuin 0,050, joten nollahypoteesi hylätään.

Huomaa, että asiaa voidaan tarkastella myös ryhmien välisenä erona (miesten ja naisten tyytyväisyydessä ei ole eroa). Testi lasketaan samalla tavalla riippumatta siitä onko näkökulmana sukupuolen ja mielipiteen riippuvuus vai miesten ja naisten mielipiteiden ero.

# 5 TÄRKEITÄ HUOMIOITA

## 5.1 p-arvon tulkinta

Nollahypoteesin hylkäämisraja (esimerkiksi 5 %) on mielivaltainen. Ei ole käytännössä juurikaan eroa jos p-arvo on 4,5 % tai 5,5 %. Kuitenkin edellä esitetyn mekaanisen (ja mielivaltaisen) päättelysäännön mukaan p-arvo 4,5 % johtaa nollahypoteesin hylkäämiseen ja 5,5 % ei johda nollahypoteesin hylkäämiseen. Asia kannattaakin ymmärtää seuraavasti:

**Mitä pienempi p-arvo sitä enemmän vaihtoehtoinen hypoteesi saa tukea.**

Käytännön tilanne, johon testaamien liittyy, täytyy myös aina huomioida.

Esimerkki. Puolueen kannatuksen muutoksia selvitetessä voidaan hyvinkin arvioida kannatuksen muuttuneen vaikka p-arvo olisikin suurempi kuin 5 %. Jos p-arvo on 5 % ja 10 % välillä, niin tulosta voidaan hyvinkin pitää suuntaa antavana sen puolesta, että kannatus on muuttunut.

Esimerkki. On tavallista, että valmistettujen tuotteiden laatua seurataan jatkuvasti otosten avulla. Tällöin nollahypoteesina on, että valmistusprosessi toimii kuten pitääkin ja tuotteet ovat ominaisuuksiltaan tavoitearvojen mukaisia. Jos otos antaa todisteita nollahypoteesia vastaan, niin riippuu toimintaympäristöstä miten tähän suhtaudutaan. Kyseessä on itse asiassa tärkeä päätöksentekotilanne, jonka vaihtoehtoina ovat:

<b>Testauksen tulos</b>	<b>Todellinen tilanne</b>	
	<b>Prosessi OK</b>	<b>Prosessissa jotain vialla</b>
<b>Jatka tuotantoa</b>	Oikea päätös	Hyväksymisvirhe
<b>Pysäytä tuotanto</b>	Hylkäämisvirhe	Oikea päätös

Jos tehdään hyväksymisvirhe, niin virheellinen tuotanto saa jatkua, mistä on tietenkin haitallisia seurauksia.

Jos tehdään hylkäämisvirhe, niin tuotanto pysäytetään turhaan vian etsimistä varten ja tämä maksaa rahaa.

Päätöksentekijän täytyy löytää toimintaympäristöön sopiva p-arvo, jonka alittamien johtaa tuotannon pysäyttämiseen. Mitä kalliimpi hylkäämisvirhe on verrattuna hyväksymisvirheeseen sitä pienempää p-arvoa edellytetään nollahypoteesin hylkäämiseksi.

## 5.2 Tilastollinen merkitsevyys ja käytännön merkitsevyys

Hypoteesin testauksessa on tapana puhua tilastollisesta merkitsevyydestä. Yhden muuttujan testeissä kyse on esimerkiksi keskiarvon tai prosenttiluvun merkitsevästä erosta nollahypoteesiin verrattuna. Ryhmien vertailussa kyse on ryhmien välisten erojen merkitsevyydestä. Riippuvuuden testaamisesta kyse on riippuvuuden merkitsevyydestä. Tilastollisen merkitsevyyden ohella on syytä miettiä myös käytännön merkitsevyyttä.

Esimerkki. Oletetaan, että älykkyystestin maksimipistemäärä on 200.

**Nollahypoteesi:** Miehillä ja naisilla sama keskiarvo.

Valitaan satunnaisesti otos miehiä ja otos naisia suorittamaan kyseinen älykkyystesti. Miesten ja naisten keskiarvopistemäärän eroksi saadaan 0,5 pistettä ja p-arvoksi saadaan 3 %. Tällöin nollahypoteesi hylätään ja miesten ja naisten ero on näin osoitettu tilastollisesti

merkitseväksi. Voidaan kuitenkin oikeutetusti kysyä, onko 0,5 pisteen ero tässä asiassa käytännössä millään tavalla merkityksellinen?

Erityisesti isojen otosten kohdalla tilastollinen merkitsevyys saadaan usein osoitettua vaikka käytännön merkitsevyys on kyseenalainen. Myös toisin päin voi käydä. Käytännöllinen merkitsevyys voi vaikuttaa ilmeiseltä vaikka tilastollista merkitsevyyttä ei saada osoitettua esimerkiksi pienen otoskoon takia. Järkeä täytyy aina säilyttää päässä ja lopullisen tilanearvion ja päätöksen tekee ihminen.

### 5.3 Normaalijakautuneisuus ja otoskoko 30

Monissa testeissä edellytetään otoskeskiarvojen normaalijakautuneisuutta. Pienillä otoksilla normaalijakautuneisuus on syytä tarkistaa. Otoskeskiarvojen jakaumaa ei päästä suoraan tarkastelemaan. Sen sijaan testataan otoksen perusteella muuttujan arvojen normaalijakautuneisuutta. Tarkistuksen voi tehdä laatimalla muuttujan arvoista histogrammin. Jos muuttujan arvot noudattavat likimain normaalijakaumaa niin myös otoskeskiarvojen voidaan olettaa noudattavan normaalijakaumaa.

Edellä on esitetty, että otoskoosta 30 ylöspäin testin käyttöedellytyksiin kuuluvaa normaalijakautuneisuutta ei tarvitse erikseen tarkistaa. Rajana pidetty otoskoko 30 ei ole mikään maaginen raja, vaan tilanteen mukaista harkintaa kannattaa käyttää.

Keskeisen raja-arvolauseen mukaan eri otoksista saatavien otoskeskiarvojen jakauma lähennee normaalijakaumaa otoskoon kasvaessa, riippumatta siitä minkälainen jakauma muuttujalla on perusjoukossa. Käytännössä on havaittu, että otoskoosta 30 ylöspäin ollaan useimmissa tapauksissa jo riittävän lähellä normaalijakaumaa. Jos muuttujan jakauma perusjoukossa on epätavallinen (erittäin vino, monihuippuinen, jne.), niin tarvitaan isompi otos normaalijakautuneisuuden takaamiseksi.

## 6 EXCEL JA TILASTOLLINEN PÄÄTTELY

Yhtä muuttujaa koskeva päättely


Yhteen muuttujaan liittyen Excelillä voidaan helposti laskea keskiarvoon ja prosenttilukuun liittyvät virhemarginaalit ja p-arvot. Laskentapohjia ja esimerkkejä löydät työkirjasta <http://myy.haaga-helia.fi/~taaak/p/p.xlsx>. Esimerkiksi keskiarvon virhemarginaalin laskentapohjaan voit kirjoittaa otoskoon ja keskihajonnan ja lukea sen jälkeen virhemarginaalin.

Keskiarvon virhemarginaali	
Syötä lähtötiedot kehystettyihin soluihin. Älä koske muihin tietoihin.	
Otoskoko	100
Keskihajonta	0,3000
Luottamustaso	95 % (yleensä 95 %)
<i>t-jakauman kriittinen arvo</i>	1,9842
<b>Virhemarginaali</b>	<b>0,0595</b>

Huomaa, että virhemarginaalien laskentaan käytetyt laskentapohjat antavat hieman tässä monisteessa annetuista laskentakaavoista poikkeavia arvoja. Tässä monisteessa annetut laskentakaavat antavat likimain oikean virhemarginaalin ja kelpaavat useimpiin käytännön sovelluksiin.

Kahden ryhmän vertailu

Excelin **TTEST** (**TTESTI**) -funktiolla voit kätevästi laskea sekä riippumattomien että riippuvien otosten vertailutestien p-arvot. Funktion uusi nimi Excel 2010 -versiossa on **T.TEST** (**T.TESTI**). Vanha funktion nimi toimii edelleen ja sitä täytyy käyttää, jos haluat säilyttää yhteensopivuuden vanhempien Excel-versioiden kanssa.

Jos lisää funktion lisätoiminnolla , niin vanhat funktiot löytyvät **Compatibility**-kategoriasta (**Yhteensopivuusfunktiot**).

Esimerkkejä **TTEST**-funktion käytöstä <http://myy.haaga-helia.fi/~taaak/p/p.xlsx>.

Kahden muuttujan välinen riippuvuus

Työkirjassa <http://myy.haaga-helia.fi/~taaak/p/p.xlsx> on laskentapohja korrelaatiokertoimeen liittyvän p-arvon laskemiseksi ja esimerkki khiin neliö -riippumattomuustestin p-arvon laskemisesta.