Sisällysluettelo

ESIPUHE 4

ALKUSANAT E-KIRJA –VERSIOON 5

SISÄLLYSLUETTELO 6

1. JOHDANTO TILASTOLLISEEN PÄÄTTELYYN 8
   1.1 INDUKTIO JA DEDUKTIO 9
   1.2 SYYT JA VAIKUTUKSET 11
   TEHTÄVIÄ 13
2. TODENNÄKÖISYYS 14
   2.1 MERKINTÖJÄ 15
   2.2 KLASSINEN TODENNÄKÖISYYS 16
   2.3 TODENNÄKÖISYYDEN FREKVENSSITULKINTA 17
   2.4 TODENNÄKÖISYYDEN OMINAISUUKSIA 17
   2.5 TODENNÄKÖISYYS JA P-ARVO 20
   2.6 TODENNÄKÖISYYS JA PINTA-ALATULKINTA 21
   TEHTÄVIÄ 22
3. OTANTAJAKAUMAT JA NIIDEN KÄYTTÖ 23
   3.1 EPÄJATKUVIA JAKAUMIA 24
   3.2 NORMAALIJAKAUMA 25
   3.3 T-JAKAUMA JA χ^2 -JAKAUMA 28
4. HYPOTEESIEN TESTAAMINEN 30
   4.1 TESTATTAVA HYPOTEESI 30
   4.2 HYPOTEESI MATEMAATTISESSA MUODOSSA 31
   4.3 PÄÄTÖSTEORIA JA TESTISUURE 33
   4.4 KIINTEÄN P-ARVON KÄYTTÖ 36
   4.5 P-ARVON ETSIMINEN TAULUKOSTA JA TIETOKONEELLA 38
      Kiinteä p-arvo taulukossa 38
      Havaittu p-arvo laskentataulukko-ohjelmassa 39
   4.6 MERKITTÄVYYS JA MERKITSEVYYS 41
   TEHTÄVIÄ 42
5. JAKAUMIA HYÖDYNTÄVIÄ TESTEJÄ 43
   5.1 χ^2 –RIIPPUMATTOMUUSTESTIN TULKINTA 43
   5.2 KORRELAATION MERKITSEVYYS 45
   5.3 KAHDEN KORRELAATION ERON MERKITSEVYYS 46
   5.4 T-TESTIN TULKINTA 48
   5.4 PIENI OTOSKOKO JA JAKAUMAOLETUS 50
   TEHTÄVIÄ 51
6. LUOTTAMUSVÄLI JA SEN LASKEMINEN 52

6.1 LUOTTAMUSVÄLIN LASKEMINEN 52
6.2 ESIMERKKI LUOTTAMUSVÄLIN LASKEMISESTA 54
6.3 CRONBACHIN ALFA JA SEN LUOTTAMUSVÄLIN LASKEMINEN 54
Alfan luottamusvälin laskemisen tekniikka 56
Kriittiset pisteet ja luottamusväli SPSS -ohjelmistolla 57
Kriittiset pisteet ja luottamusväli Excel-laskentataulukolla 60
Kriittiset pisteet ja luottamusväli SURVO 98C –ohjelmistolla 62
Alfan luottamusväli ja erilaiset otoskoot 64
Johtopäätöksiä alfan luottamusvälistä 64
TEHTÄVIÄ 65

7. TULOSTEN TARKASTELU JA POHDINTA 66

7.1 MATEMATIIKKA JA TULKINNAT 66
7.2 TULOKSET JA TULKINNAT 67
7.3 SUOSITUKSET JA TULKINNAT 67

8. LOPUKSI 69

LÄHTEET 70

ASIA- JA HENKILÖHAKEMISTO 71

TAULUKOT 73

4.5 p-arvon etsiminen taulukosta ja tietokoneella

Tilastotiedon parhaissa menetelmäkirjoissa on taulukoitu edellä mainittujen kolmen merkitsevyystason mukaisia kriittisiä arvoja. Esimerkiksi Ranta-Rita-Koukissa (1986) on huomattava määrä erilaisia tilastollisia taulukoita. Toisaalta nykyisillä taulukkolaskentaohjelmilla – kuten esimerkiksi Excelillä – on mahdollista laskea hyvin yksinkertaisesti samat kriittiset arvot, mutta myös havaittuja p-arvoja. Tässä luvussa käydään nopeasti läpi molemmat mahdollisuudet löytää oman tutkimuksen kannalta oleelliset vertailuarvot nollahypoteesin hylkäämisen tai hyväksymisen ratkaisemiseksi.

Kiinteä p-arvo taulukossa

Jos tiedämme testisuureen arvon – sanokaamme vaikka χ^2-testiarvon – voimme valmiiksi taulukoitujen arvojen perusteella päätellä, pitäisikö nollahypoteesi hylätä vai saako nollahypoteesi tukea. Kiinteää p-arvoa etsittäessä taulukosta tarvitaan kaksi tietoa:

1. **millä riskitasolla päätelmä tehdään, eli mikä on valittu p-arvo**
2. **mikä on vapausaste eli *degrees of freedom*, lyhyemmin *df***

Jos otamme esimerkiksi *Tilastollisen kuvauksen perusteet* –kirjassa esitetyn ristiintaulukon analyysin yhteydessä saadun χ^2–testiarvon 2.727, tiedämme samalla myös vapausasteet. Ristiintaulukkomme oli 2x2 –taulukko, joten tarvitsemamme vapausaste on 1, joka merkitään df=1. Tämä yleensä ilmaistaan myös itse χ^2–testiarvon merkinnässä, joka tarkalleen ottaen pitäisi merkitä $\chi^2(1)$, jossa suluissa oleva 1 viittaa vapausasteisiin. Valitsemme kaksisuuntaiseksi riskitasoksi 5 % eli 0.05. Näin ollen χ^2 –taulukon avulla voimme päätellä edustaako saamamme testisuureen arvo 2.727 niin harvinaista tapausta jakaumassa, että H_o joudutaan hylkäämään. Kriittinen piste luetaan taulukosta siitä risteyspisteestä, jossa kohtaavat yhden vapausasteen rivi ja 0.05 arvon sarake.

p-arvo = 0.05

Vapausaste=1

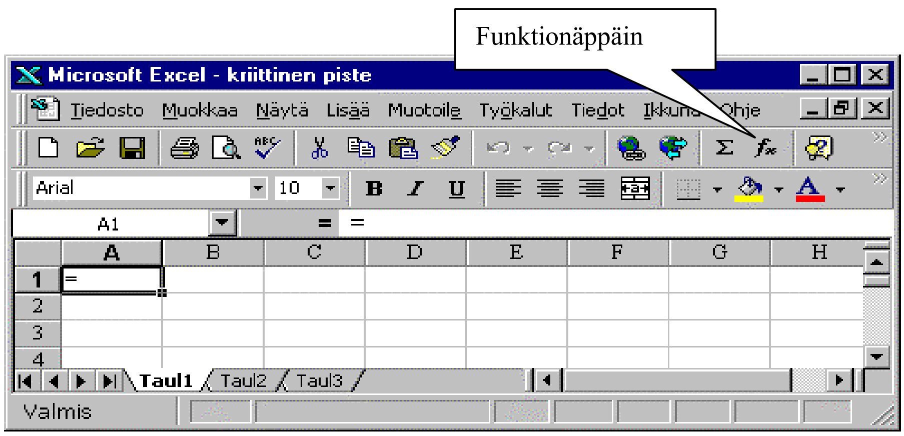
df	0.1	0.05	0.01	0.001
1	2.706	3.841	6.635	10.828
2	4.605	5.991	9.210	13.816
3	6.251	7.815	11.345	16.266
4	7.779	9.488	13.277	18.467
5	9.236	11.070	15.086	20.515
6	10.645	12.592	16.812	22.458
7	12.017	14.067	18.475	24.322

Kriittinen piste on 3.841. Tätä arvoa suurempia arvoja (mukaan lukien itse arvo) on siis tuon muotoisessa χ^2 –jakaumassa 5 %. Nyt saamamme χ^2 –arvo 2.727 on pienempi kuin ko. kriittinen piste, joten emme voi hylätä nollahypoteesia.

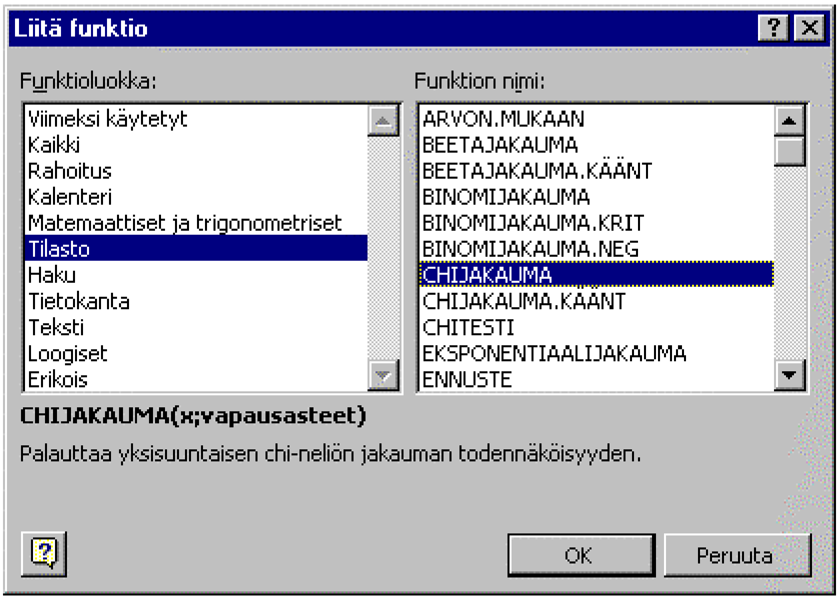
Havaittu p-arvo laskentataulukko-ohjelmassa

Nykyisin lienee helpompi etsiä havaittuja p-arvoja tietokoneella kuin taulukkokirjoista selaillen. Havaittu p-arvo on tarkempi mitta kuin kiinteä p-arvo. Jo aiemmin olemme todenneet, ettei mikään järkeilyperuste puolla kiinteän p-arvon käyttöä kuin erikoistilanteissa.

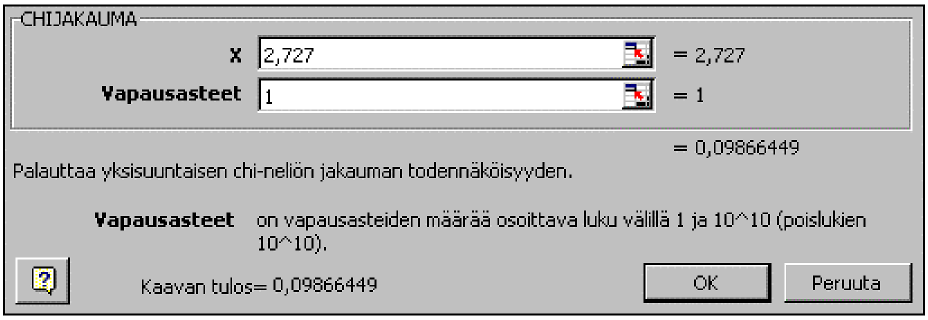
Excel-taulukkolaskentaohjelmassa havaittu p-arvo haetaan seuraavasti. Missä hyvänsä taulukon ruudussa kaksoisklikataan ruutua ja kirjoitetaan = -merkki ja etsitään funktionäppäin painike tai *Lisää Funktio* –komento.



Funktionäppäimen painamisen jälkeen avautuu ikkuna, josta valitaan sopiva jakauma: χ^2 –jakauma löytyy nimellä CHIJAKAUMA, t-jakauma löytyy nimellä TJAKAUMA ja f-jakauma nimellä FJAKAUMA.



Ikkunan alareunassa kerrotaan että tarvitaan kaksi arvoa: **x** ja **vapausasteet**. Luvuista **x** viittaa testisuureen arvoon (2.727) ja toinen nimensä mukaiseen vapausasteeseen (1). OK-näppäintä painamalla pääsemme antamaan kyseiset arvot ja saamme khiin neliö -jakauman todennäköisyyden.



Taulukko kertookin meille jo tässä vaiheessa, että ko. pisteen ja sitä suurempien arvojen todennäköisyys on 0.0987. Huomattakoon, että luvussa 2.727 on käytettävä pilkkua, jotta Excel ymmärtäisi sen numeroksi. OK-näppäintä painamalla saamme havaitun p-arvon Excel taulukkoon: