

[MTTTA1] TILASTOMENETELMIEN PERUSTEET, KEVÄT 2019

<https://coursepages.uta.fi/mttta1/kevat-2019/>

HARJOITUS 2 viikko 5

RYHMÄT:

ke	12.15–13.45	Is. C6	Leppälä
to	08.30–10.00	Is. C6	Korhonen
to	12.15–13.45	Is. C6	Korhonen
to	14.15–15.45	Is. C8	Leppälä

Aiheet: Varianssianalyysi, χ^2 - yhteensopivuustesti

1. Tarkastele sivulla <https://coursepages.uta.fi/mttt1/esimerkkiaineistoja/> olevaa aineistoa http://www.sis.uta.fi/tilasto/tiltp_aineistoja/Tre_myydyt_kaksiot_2016.sav, jossa on tietoja Tampereella myydyistä kaksioista. Anna aineistosta esimerkki tilanteesta, jossa voisit käyttää kaksisuuntaista varianssianalyysiä. Piirrä myös esimerkkiisi liittyvä graafinen esitys, josta voit alustavasti arvioida riippuvuutta. Tee riippuvuuden arvioiti grafiikkasi perusteella.

2. Tutustu χ^2 - jakauman tiheysfunktion kuvaajiin sivun <http://vassarstats.net/vsdist.html> kautta. Vertaa jakauman kuvaajia vapausastein 1, 5, 10 ja 20.

3. Kasvitieteilijä olettaa, että erään kasvin kukista $\frac{3}{4}$ on punaisia ja loput valkoisia. Hän kasvatti 200 kasvia, joista 170 oli punakukkaista. Onko kasvitieteilijän syytä muuttaa käsitystään? Tuki asiaa laskemalla tilanteeseen sopiva Z -testisuure sekä χ^2 - testisuure. Mitä huomaat, kun vertaat testisuureiden arvoja? Määritä molempiin testeihin liittyvät p-arvot. Käytä p-arvojen määrittämisessä esim. laskentaohjelmaa <http://onlinestatbook.com/2/calculators/calculators.html> .

4. Yritys valmistaa tehtaassaan erästä tuotetta. Epäillään, että virheellisten tuotteiden lukumäärä kasvaisi työpäivän loppua kohden. Viimeisen kuukauden aikana virheellisten tuotteiden lukumäärät työtunneittain olivat 2, 4, 3, 3, 3, 7, 8, 10. Vaikuttaako ajankohta virheellisten tuotteiden määrään?

5. Tarkasteltiin kaksilapsisia perheitä. Saatiin frekvenssijakauma:

Tyttöjen lukumäärä perheessä	0	1	2
Perheitä	60	100	40

Halutaan tutkia, syntyykö tyttöjä ja poikia saman verran. Aseta tilanteeseen liittyvät hypoteesit. Millä testillä voit testata asettamaasi nollahypoteesia? Mitä jakaumaa testisuuresi noudattaa nollahypoteesisi ollessa tosi? Laske testisuureen arvo. Suorita testaus 5 %:n riskitasolla. Määritä (tai arvioi) pienin riskitaso, jolla nollahypoteesisi voidaan hylätä.

6. Eräs tehdas tekee kivihiilestä koksia (<http://fi.wikipedia.org/wiki/Koksi>) . Kirjattiin päivittäin 260 päivän ajan saadun koksen määrä (prosentteina käytetystä kivihiilestä). Saatiin frekvenssijakauma:

Koksi %	Päivien lukumäärä
66,5-67,5	7
67,5-68,5	23
68,5-69,5	61
69,5-70,5	87
70,5-71,5	56
71,5-72,5	19
72,5-73,5	7

Jakauman perusteella saadaan (arviot) keskiarvoksi 69,95 ja keskihajonnaksi 1,26. Miten nämä arviot on pystytty laskemaan? Halutaan tutkia, voidaanko vaihtelua koksiprosentissa pitää normaalijakautuneena. Laske luokan 71,5 – 72,5 teoreettinen frekvenssi. Kun lasketaan kaikkien luokkien teoreettiset frekvenssit, voidaan laskea χ^2 -testisuure. Testisuureen arvoksi saadaan 1,78. Tee johtopäätelmä.