

Tilastotieteen harjoitustyö 2018

Samuli Virtanen

spevir@utu.fi

op. nro. 511178

1. Numeeristen vastemuuttujien mallinnus

Elinolo2018 (Tilastokeskuksen elinolotutkimuksen aineisto, N=2199)

Kukin opiskelija puolestaan poimii 600 kokoisen satunnaisotoksen kyseisestä datasta seuraavalla tavalla:

```
Transform-Random Number Generators...
Set Starting Point-Fixed Value-annetaan oma opiskelijanumero-0k
DATA>SELECT CASES>RANDOM SAMPLE OF CASES>
EXACTLY 600 CASES FROM THE FIRST 2199 CASES
CONTINUE
UNSELECTED CASES ARE FILTERED>
```

1.1. Varianssianalyysi

Tutki, onko sukupuolella ja asumisahtaudella yhteyttä asunnon pinta-alaan.

1.1.1. Suunnitelma

Käytetään kaksisuuntaista varianssianalyysiä Normaalijakaumaoletus

Ensimmäiseksi tutkitaan jakaumia sukupuolittain asumisahtauden perusteella. Shapiro-Wilk-testin tulosten perusteella yksikään ryhmä ei ole normaalisti jakautunut. Havaintojen suuresta määrästä johtuen voidaan kuitenkin käyttää parametrista testiä.

1.1.2. Kaksisuuntainen varianssianalyysi

Tehdään kaksisuuntainen varianssianalyysi SPSS:n Univariate-työkalulla.

Tuloksista nähdään että selittävien muuttujien päävaikutukset ovat merkitseviä merkitsevyystasolla $p < 0,05$, mutta yhdysvaikutus ei ole merkitsevä.

Tukeyn testi puolestaan osoittaa että asumisahtauden luokkien väliset erot ovat kaikki merkitseviä ($p < 0,05$).

Koska yhdysvaikutus ei ole merkitsevä, voidaan tulkita että miehillä on keskimäärin suuremmat asunnot kuin naisilla, ja asunnon pinta-ala kasvaa koetun asumisahtauden mukaan järjestyksessä ahdas -> normaali -> tilava (pienestä pinta-alasta suurempaan).

1.2. Regressiomalli

Tutki, onko kotitalouden kuluttajayksiköiden lukumäärällä, asumismenoilla yhteensä ja alueella asumisaikalla yhteyttä asunnon pinta-alaan.

Tutkitaan sirontakuvioita.

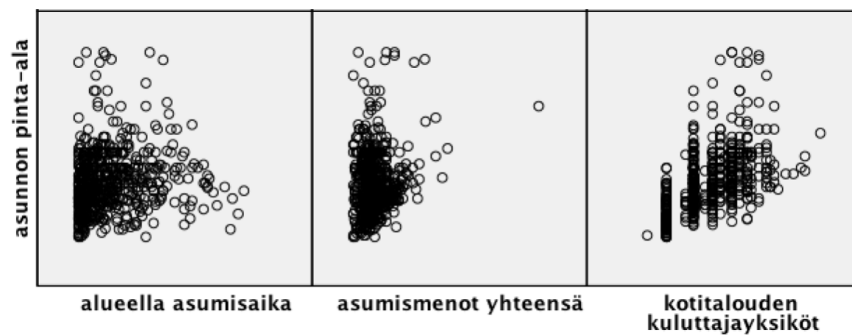


Figure 1: Scatter plots

Kovin selkeitä korrelaatioita ei esiinny. Näyttäisi kuitenkin hieman siltä että kodin kuluttajayksiköiden määrän kasvaessa myös asunnon pinta-ala kasvaa ja toisaalta samoin käy myös asumismenojen kasvaessa. Yksi outlier asumismenoissa hankaloittaa hieman kuvan tulkintaa.

1.3. Toistomittausmalli

Tutkijalla on hypoteesi, että potilaan mielestä saatu ohjaus leikkauksen jälkeen toiminnallista seikoista (Functional_M2) on ollut vähäisempää kuin odotettu ennen leikkausta (Functional_M1). Eli keskiarvo toisessa mittauksessa on matalampi. Lisäksi kiinnostaa se, onko tuo ero mittauksen välillä erilainen sukupuolittain.

Tutki saavatko nämä tutkimushypoteesit tukea mallittamalla aineisto toistettujen mittausten varianssianalyysillä.

Huom. Numeeristen vastemuuttujien mallituksessa on varianssianalyysien osalta tehtävä tarvittava kuvaileva tarkastelu ja regressiomallissa yhteyksien suoraviivaisuuksien tarkastelu ja jäännöstarkastelu.