

JYVÄSKYLÄN YLIOPISTO  
Matematiikan ja tilastotieteen laitos  
Johdatus tilastolliseen mallintamiseen

Harjoitustyö 1  
Viimeinen palautuspäivä 14.11.2011

**Kirjoita siisti raportti jollakin tekstinkäsittelyllä. Esitä raporttissasi vain mielestäsi tarpeelliset taulukot ja kuvat, joista helposti käy ilmi mitä ne sisältävät. Viiden sivun raportti on riittävä. Liitä käyttämäsi koodi raporttisi loppuun. Koodi mukaan luettuna sivumäärä voi olla enemmän kuin viisi.**

**Jokainen kirjoittaa raporttinsa itsenäisesti.**

**Palautus paperisena luennoilla, harjoituksissa tai työhuoneeni (MaD321) vieressä olevaan lokeriin. Nido paperit yhteen!**

Aineisto on yomat. Muuttujat aid1, aid2, englantia, ruotsi, fys, kemia, mat ovat peruskoulun päättötodistuksen arvosanoja; sp on sukupuoli, 0 = nainen, 1 = mies; koulu on koodattu 1,2,3,4 kuten luento-esimerkeissä; vaste on pmat = matematiikan yo-kirjoitusten pitkän matematiikan pistemäärä v. 1994.

1. Ottamatta huomioon muita mahdollisia selittäjiä tutki poikkeako koulujen keskimääräinen menestys vuoden 94 kirjoituksissa. Jos poikkeaa, laske Tukeyn monivertailumenetelmällä parittaiset samanaikaiset 95 %:n luottamusvälit. Kerro sanoin, millaisia eroja havaitset.
2. Lisää malliin sukupuoli päävaikutuksena. Tulkitse kertoimet monisteen luvun 2 tapaan sopivien ryhmien keskiarvojen erotuksina. Päättele tämän analyysin perusteella poikkeako miesten ja naisten menestys yo-kirjoitusten pitkässä matematiikassa. Piirrä interaktiokuvaaja. Mitä havaitset? Miten näkemäsi mahdollisesti vaikuttaa edellä esittämääsi sukupuolten eroon.
3. Lisää malliin sukupuolen ja koulun interaktio. Testaa sen merkitsevyyttä. Laske sukupuolten keskiarvojen ero kouluittain. Kommentoi laskujesi tulosta. Laske koulujen erot erikseen miehille ja naisille. Kommentoi laskujen tuloksia.
4. Lisää malliin peruskoulun matematiikan numero sellaisen muunnoksen jälkeen, että voit mallin vakion tulkita järkevästi. Tulkitse kertoimet monisteen luvun 2 tapaan sopivien ryhmien keskiarvojen erotuksina.
5. Tutki sopivalla testillä, onko malliin syytä lisätä a) matematiikan numeron ja sukupuolen interaktio ja b) matematiikan numeron ja koulun interaktio.
6. Tutki parantaako jokin muu tiedostossa oleva peruskoulun arvosana tai jokin niiden yhdistelmä mallin ennustavuutta. Esitä kommenttisi lopullisen mallisi ennustekyvystä.