

# SAS-ohjelmiston perusteet

## Harjoitus 5

ke/to 10-11.12.08 klo 16-18 ML42

Datamatriisit löytyvät netistä <http://mtl.uta.fi/tilasto/sas/syky08/dataEG/>

1. (0-2 p) Datamatriisi **satunnaisia**: Ovatko muuttujat X ja Y identtiset? Jos eivät ole, niin millä rivillä tai riveillä ne eroavat toisistaan? Entä onko niin, että muuttujien X (tai Y) ja Z arvot ovat tosiasiansa samat, vain järjestys on eri?
2. (0-2 p) Tee 20 alkion satunnaisotos (Data → Random Sample) aineistosta **satunnaisia** ja testaa t-testillä hypoteesia, että muuttujan X otoskeskiarvo on sama kuin sen keskiarvo koko aineistossa. Toteuta otoksen otto niin että pystyt toistamaan sen tarvittaessa (siis niin, että saadaan sama otos).
3. (0-2 p) Luokittele datamatriisin **satunnaisia** muuttujan X arvot 5 luokkaan niin, että kussakin luokassa olisi suurin piirtein yhtä monta havaintoa. Havainnollista jaon onnistumista sopivalla kuviolla (esim. Graph → Bar Chart). Vertaa EG:n automaattisesti tekemää jako viiteen luokkaan saman kuvion yhteydessä.
4. (0-2 p) Monisteen [http://www.sascommunity.org/mwiki/images/6/63/Ex\\_for\\_SUSEG.pdf](http://www.sascommunity.org/mwiki/images/6/63/Ex_for_SUSEG.pdf) tehtävä 6.3. s. 37. Datamatriisi **exer06\_03**.

**Bonustehtävä** Jatkoa tehtävään 4: havainnollista  $R^2$  (Coefficient of Determination) merkitystä 1-suuntaisessa varianssianalyysissä konstruoimalla aineisto, jossa sen arvo on 0, ja aineisto, jossa sen arvo on 1.

Lisäksi esitellään harjoitustöitä. Jotta harjoitustyöstä saisi ylipäättään pisteitä, se on palautettava viimeistään lauantaina 20.12. klo 10 mennessä.

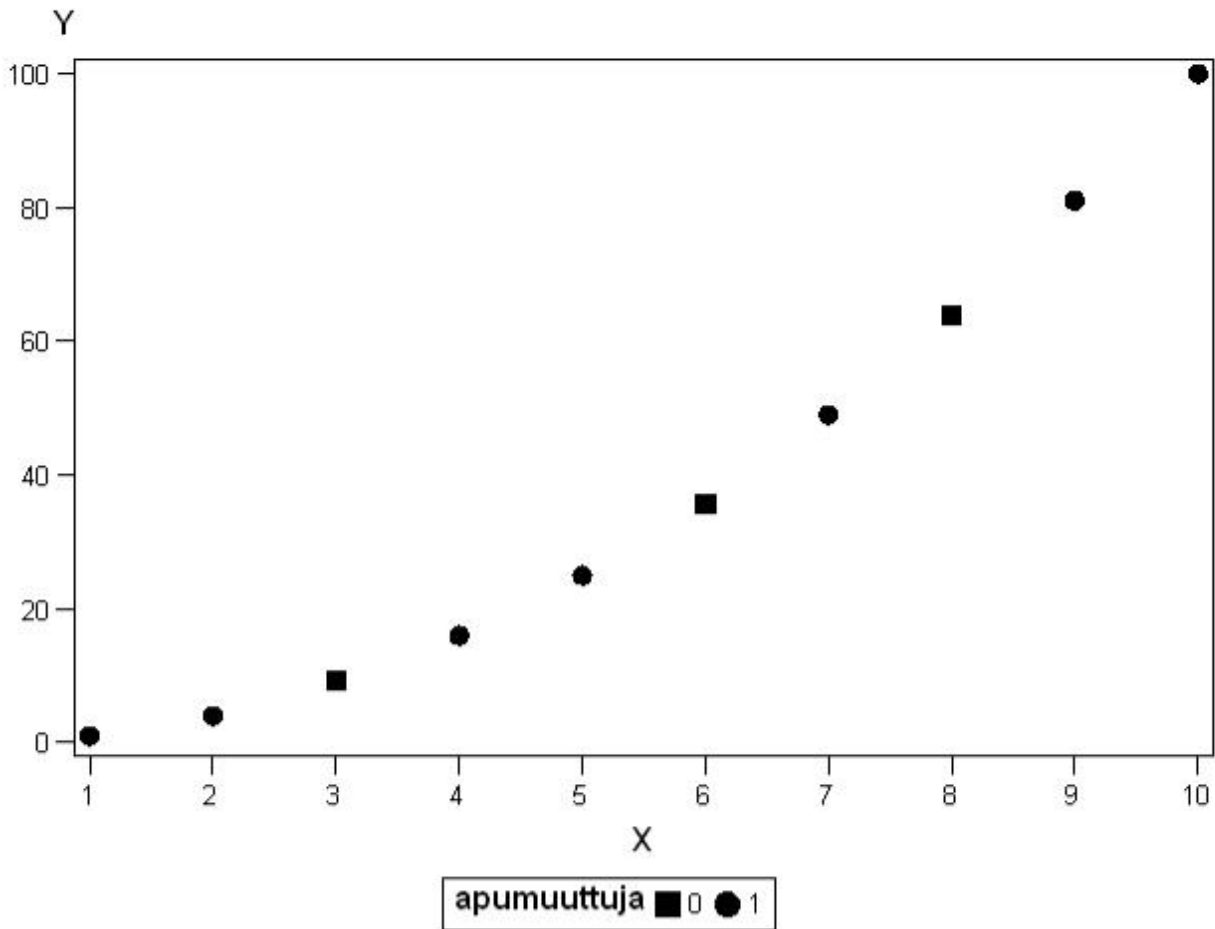
## Harjoitus 6

Palautus kirjallisesti lauantaina 20.12. klo 10 mennessä mennessä (koskee myös sähköistä palautusta). HUOMAA pidennetty aika.

5. Anna sellainen esimerkki makromuuttujan käyttämisestä 1-suuntaisen varianssianalyysin yhteydessä, jossa makromuuttujan arvoksi annetaan selitettävä muuttuja (datamatriisissäsi on siis oltava vähintään kaksi muuttujaa, joita voidaan tarkastella saman luokittelevan muuttujan luokissa).
6. Anna sellainen esimerkki makromuuttujan käyttämisestä Filter-Query -komennon yhteydessä, jossa makromuuttujina on alaraja ja yläraja ja kyselyssä valitaan ehdon 'alaraja < X < yläraja' mukainen osa-aineisto. (Riittää että datamatriisissäsi on muuttuja X.)

7. Datamatriisissa **pisteet** on muuttuja X ja Y, jossa yleensä  $Y = X^2$  paitsi kolmessa poikkeustapauksessa. Tehtävänäsi on piirtää EG:llä sellainen pistekuvio pisteistä (X,Y), joissa kolme poikkeavaa tapausta erottuvat. Alla esimerkkinä Graph-Line Plot -valinnalla tuotettu ratkaisu (huomaa vihjeet kuviossa):

### Line Plot



8. Havainnollista makromuuttujan käyttämisestä SAS-koodissa käymällä läpi Hurmeen monisteen esimerkki s. 113-114. Koodin kirjoittamisen lisäksi suorita myös se. (Tee tarvittava muutos dataviittaukseen.)

**Bonustehtävä.** Datamatriisissa **polynomi** on muuttujat x ja y, jossa

$$y = a_0 + a_1 x + a_1 x^2 + \dots + a_9 x^9 + a_{10} x^{10}.$$

Selvitä **EG:n** avulla polynomin kertoimet  $a_0, a_1, \dots, a_{10}$  (voivat olla myös nollia).