

Pienalue-estimointi (78189), Kevät 2013

Harjoitustyö (Päivitetty versio)

Aineopinnot: Tehtävät 1, 2 ja 3 (Vapaaehtoinen harjoitustyö)

Aineopinnoissa tehtävä 4 on JOKERI josta saa bonusta!

Syventävät opinnot: Tehtävät 1, 2, 3 ja 4 (HUOM: Syventävissä harjoitustyö on pakollinen)

Harjoitustyö tehdään yksittäistyönä tai 2 hengen ryhmissä. Työn hyväksytty suoritus: 2 op.

Lopputuotteena on noin 10 sivun raportti liitteineen, esimerkiksi rakenteella:

Kansilehti (otsikko, tekijä(t), kurssi, päiväys, merkintä Aineopinnot tai Syventävät opinnot), Tekstiosa (jaa soveltuviin lukuihin ja alilukuihin), Viitteet (kirjallisuus), Liitteet (valitut ohjelmakoodit ja tulostusotteet).

Työn palautus: **30.4.2013 mennessä** liitetiedostona (PDF) risto.lehtonen@helsinki.fi tai postitse:

Risto Lehtonen, HY, Matematiikan ja tilastotieteen laitos, PL 68 (Gustaf Hällströmin katu 2b), 00014 Helsingin yliopisto. Työn voi myös jättää ML-laitoksen toimiston postilaatikkoon (Exactum 3 krs.) aukioloaikoina.

Perusjoukko: otos.sas7bdat, $N = 966$ alkiota (demojen aineisto, linkki löytyy kurssin kotisivulta)

Osajoukkorakenne: $D = 10$ osajoukkoa (domains)

Estimoitavat parametrit: Osajoukkojen totaalit muuttujalle y

Apuinformaatio: Alkiotasoinen apumuuttuja x

Laskenta: SAS ja SAS-makro EBLUPGREG. Laskennan voi tehdä haluttaessa myös SPSS-, Stata- tai R-kielellä.

Tehtävä 1. Kehikkoperusjoukon kuvailevat tarkastelut ja otosten poiminta

Suorita mielestäsi soveltuvat perusjoukon ja osajoukkojen tilastolliset kuvailut (perusjoukon parametrien ja eri tunnuslukujen laskenta ja taulukointi, graafiset kuvailut).

Tehtävä 2. Suunniteltujen osajoukkojen tilanne (ositettu otanta)

a) Poimi perusjoukosta 10 % otos ($\text{samprate}=.10$) ositetulla yksinkertaisella satunnaisotannalla (STR-SRSWOR) niin, että ositteina ovat osajoukot ($D = 10$ ositetta). Käytä suhteellista kiintiöintiä. HUOM: SAS:n SURVEYSELECT-proseduurissa käytä $\text{samprate} = \text{optiota}$. Valitse oma siemenluku SEED.

b) Estimoi osajoukkojen totaalit HT-estimoinnin avulla. Raportoi käyttämäsi totaali- ja varianssiestimaattorit, ohjelmakoodit ja tulokset taulukoituna (totaaliestimaatit, keskivirhe-estimaatit, variaatiokertoimet).

c) Estimoi osajoukkojen totaalit GREG-estimoinnin avulla käyttämällä apuna soveltuvaa tilastollista mallia ja lisäinformaatiota. Raportoi käyttämäsi totaali- ja varianssiestimaattorit, tilastolliset mallit, ohjelmakoodit ja tulokset taulukoituna (totaaliestimaatit, keskivirhe-estimaatit, variaatiokertoimet).

d) Vertaile tehtävien 2b ja 2c estimointituloksia ja tee perustellut johtopäätökset.

e) Suorita kohdat a) ja b) poimintasusteella $\text{samprate}=.25$ (25 % otos) ja vertaa HT-estimointituloksia 10 % otoksesta saattuihin tuloksiin.

Tehtävä 3. Ei-suunniteltujen osajoukkojen tilanne

a) Poimi perusjoukosta otos yksinkertaisella satunnaisotannalla (SRSWOR, ei ositusta) niin, että otoskoko on $n = 100$ alkiota. HUOM: Valitse tässäkin oma siemenluku SEED.

b) Estimoi osajoukkojen totaalit HT-estimoinnin avulla. Raportoi käyttämäsi totaali- ja varianssiestimaattorit, ohjelmakoodit ja tulokset taulukoituna (totaaliestimaatit, keskivirhe-estimaatit, variaatiokertoimet).

c) Estimoi osajoukkojen totaalit GREG-estimoinnin avulla käyttämällä apuna soveltuvaa tilastollista mallia ja lisäinformaatiota. Raportoi käyttämäsi totaali- ja varianssiestimaattorit, tilastolliset mallit, ohjelmakoodit ja tulokset taulukoituna (totaaliestimaatit, keskivirhe-estimaatit, variaatiokertoimet).

d) Vertaile tehtävien 3b ja 3c estimointituloksia ja tee perustellut johtopäätökset.

Tehtävä 4. SAS-makro EBLUPGREG

a) Poimi SURVEYSELECT-proseduurilla perusjoukosta SRSWOR-otokset:

Otos 1: otoskoko $n = 50$

Otos 2: otoskoko $n = 100$

Otos 3: otoskoko $n = 250$

Kyseessä on siis ei-suunniteltujen osajoukkojen tilanne.

b) GREG-estimointi SAS-makrolla EBLUPGREG

Estimoi y -muuttujan osajoukkototaalit ja keskivirheet GREG-menetelmällä otoksille 1, 2 ja 3. Taulukoi tulokset, vertaa niitä ja raportoi päätelmät.

c) EBLUP-estimointi SAS-makrolla EBLUPGREG

Estimoi y -muuttujan osajoukkototaalit ja keskivirheet (MSE:n neliöjuuret) EBLUP-menetelmällä otoksille 1, 2 ja 3. Taulukoi tulokset, vertaa niitä ja raportoi päätelmät.

d) Vertaa GREG-estimoinnin ja EBLUP-estimoinnin tuloksia ja raportoi päätelmät.