

Otantamenetelmät / Syksy 2008

Loppuputenti 16.12.2008

SYVENTÄVÄT OPINNOT

Valitse seuraavista neljä tehtävää.

1. Määrittele käsitteet (sanallisesti ja/tai kaavojen avulla)

- Sisäkorrelaatio (*intra-class correlation*) systemaattisen otannan yhteydessä.
- Totaaliestimaattorin keskineliövirhe (*Mean Squared Error, MSE*).
- Osite (*Stratum*).

2. Ositettu otanta (*Stratified sampling*)

- Selosta alkiotasoisien ositetun otannan tavoite ja työvaiheet. Mitä lisäinformaatiota tarvitaan perusjoukon alkioista ositettua otantaa varten?
- Millaisissa tilanteissa voidaan odottaa, että alkiotasoinen ositettu otanta on totaalini estimoinnissa tehokkaampaa kuin yksinkertainen satunnaisotanta palauttamatta (SRSWOR)? Miksi? Perustele vastauksesi.
- Selosta suhteellisen kiintiöinnin (*proportional allocation*) ja tasakiintiöinnin (*equal allocation*) periaate ja toteuttaminen alkiotasoisessa ositetussa otannassa. Mitä lisäinformaatiota tarvitaan perusjoukosta suhteellista kiintiöintiä ja tasakiintiöintiä varten?

3. PPS-otanta (*Sampling with Probabilities Proportional to Size*)

- Selosta Poisson-tyyppisen PPS-otannan poimintaproseduuri. Mitä lisäinformaatiota tarvitaan perusjoukon alkioista Poisson-otantaa varten?
- Selosta toimenpiteet PPS-otannan tilanteessa, jossa kahden perusjoukon alkion kokomuuttujan arvot tuottavat sisällysmistodennäköisyyden, joka on ykköstä suurempi.
- Muotoile sanallisesti ja/tai kaavojen avulla ehdot, joiden voimassa ollessa PPS-otanta on tehokkaampaa kuin yksinkertainen satunnaisotanta palauttamatta (SRSWOR).

4. Malliavusteinen estimointi (*Model-assisted estimation*)

- Selosta regressioestimoinnin tavoite, toimintaperiaate ja työvaiheet totaalini estimoinnin yhteydessä. Mitä lisäinformaatiota tarvitaan perusjoukosta regressioestimointia varten?
- Regressioestimaattorin \hat{t}_{reg} varianssiestimaattori voidaan SRSWOR-tilanteessa kirjoittaa muotoon $\hat{v}(\hat{t}_{reg}) = N^2(1 - n/N)(1/n)\hat{s}_e^2$, missä e viittaa jäännöksiin. Vastaavan SRSWOR-perusteisen totaaliestimaattorin \hat{t} SRSWOR-varienssiestimaattori on $\hat{v}(\hat{t}) = N^2(1 - n/N)(1/n)\hat{s}_y^2$, missä y viittaa tulosmuuttujaan. Päätele kaavojen perusteella, miksi regressioestimoinnilla on taipumus tuottaa pienempi varianssiestimaatti kuin vastaavalla SRSWOR-perusteisella estimoinnilla. Perustele vastauksesi.

5. Regressioestimointi

- Olkkoon regressioestimaattori muotoa $\hat{t}_{reg} = \hat{b}_0N + \hat{b}_1T_{z_1} + \hat{b}_2T_{z_2}$. Sijoita liitteenä olevista tulostusotteista saatavat tunnusluvut sopivasti regressioestimaattorin kaavaan.
- Olkkoon regressioestimaattori muotoa $\hat{t}_{reg} = \hat{t}_{HT} + \hat{b}_1(T_{z_1} - \hat{t}_{z_1}) + \hat{b}_2(T_{z_2} - \hat{t}_{z_2})$. Sijoita liitteenä olevista tulostusotteista saatavat tunnusluvut sopivasti regressioestimaattorin kaavaan.

Helsingin yliopisto
Matematiikan ja tilastotieteen laitos

LIITE Otantamenetelmät / Syventävät opinnot / Tentti 16.12.2008

(1) The SURVEYMEANS Procedure

Data Summary

Number of Observations	10
Sum of Weights	32

Statistics

Variable	Sum	Std Dev
UE91	8444.800000	1863.811512
HOU85	49069	9508.130131
URB85	3.200000	2.653300

(2) The SURVEYREG Procedure

Regression Analysis for Dependent Variable UE91

Data Summary

Number of Observations	10
Sum of Weights	32.00000
Weighted Mean of UE91	263.90000
Weighted Sum of UE91	8444.8

Estimated Regression Coefficients

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	-20.919727	15.4206499	-1.36	0.2080
HOU85	0.181513	0.0094228	19.26	<.0001
URB85	64.876014	25.1279025	2.58	0.0296

Coefficients of Estimate "UE91 Total"

Effect	Row 1
Intercept	32
HOU85	91753
URB85	7

Analysis of Estimable Functions

Parameter	Estimate	Standard Error	t Value	Pr > t
UE91 Total	16439.0686	401.119397	40.98	<.0001