

Otantamenetelmät (78143), Syksy 2012

AINEOPINNOT

Harjoitustyö

Harjoitustyö tehdään 1-2 hengen ryhmissä. Työn hyväksytyt suoritus: 2 op.
Lopputuotteena on noin 10 sivun raportti liitteineen, esimerkiksi rakenteella:
Kansilehti (otsikko, tekijä(t), kurssi, päiväys, Aineopinnot)
Tekstiosa (jaa soveltuviin lukuihin ja alilukuihin)
Viitteet (kirjallisuus)
Liitteet (valitut ohjelmakoodit ja tulostusotteet).

Työn palautus: **31.1.2013 mennessä** liitetiedostona (PDF) sähköpostitse: risto.lehtonen@helsinki.fi
tai postitse:

Risto Lehtonen, Helsingin yliopisto, Matematiikan ja tilastotieteen laitos, PL 68 (Gustaf Hällströmin katu 2b),
00014 Helsingin yliopisto

Estimoitavana parametrina on työttömien kokonaismäärä Keski-Suomen läänissä (tulosmuuttuja UE91). Kehikkoperusjoukkona on Province91-populaatio. Tehtävissä voit käyttää ja kehittää edelleen harjoituksissa käsitellyjä SAS-ohjelmakoodeja. Voit käyttää myös SPSS:n Complex Samples -modulia tai R-funktioita.

Tehtävä 1. Lisäinformaation käyttö otanta-asetelmassa

a) Tutki graafisesti tulosmuuttujan UE91 jakaumaa perusjoukossa. Raportoi havaintosi.

b) Poimi perusjoukosta yksinkertainen satunnaisotos palauttamatta (SRSWOR) Bernoulli-otannalla niin, että odotettu otoskoko on 8 kuntaa (valitse oma siemenluku SEED). Estimoit totaali sekä totaaliestimaattorin asetelmavarianssi, keskivirhe ja variaatiokerroin. Raportoi käyttämäsi estimaattorit ja tulokset. Mistä syystä yksinkertainen satunnaisotanta ei ole hyvä strategiavalinta tässä perusjoukossa? Perustele vastauksesi.

c) Poimi perusjoukosta $n = 8$ kunnan otos ositetulla otannalla muuttujan UE91 totaalin estimointia varten. Valitse itse ositusmuuttuja, kiintiöintimenetelmä ja otantamenetelmä ositteiden sisällä (pyri mahdollisimman tehokkaaseen ositetun otannan asetelmaan). Perustele valintasi. Estimoit laatimasi ositetun otannan asetelman perusteella totaali sekä totaaliestimaattorin asetelmavarianssi, keskivirhe ja variaatiokerroin. Estimoit lisäksi totaaliestimaattorin asetelmakerroin. Raportoi otanta-asetelma, käyttämäsi estimaattorit ja numeeriset tulokset.

d) Vertaa kohtien 1b ja 1c strategioiden tehokkuutta.

Tehtävä 2. Lisäinformaation käyttö estimointiasetelmassa

a) Poimi perusjoukosta $n = 8$ kunnan SRSWOR-otos SAS-proseduurilla SURVEYSELECT (valitse oma siemenluku SEED) ja estimoit otoksen perusteella totaali sekä totaaliestimaattorin asetelmavarianssi, keskivirhe ja variaatiokerroin. Raportoi otanta-asetelma, käyttämäsi estimaattorit ja numeeriset tulokset.

b) Suunnittele ja toteuta mahdollisimman tehokas strategia totaalin estimointia varten, kun käytettävissä on kohdassa 2a poimimasi SRSWOR-otos sekä kehikkoperusjoukossa oleva lisäinformaatio. Perustele valintasi. Estimoit laatimasi strategian perusteella totaali sekä totaaliestimaattorin asetelmavarianssi, keskivirhe ja variaatiokerroin. Estimoit lisäksi totaaliestimaattorin asetelmakerroin. Raportoi strategia, käyttämäsi estimaattorit ja numeeriset tulokset.

c) Vertaile tehtävien 1c ja 2b estimointituloksia ja tee perustellut johtopäätökset.

Table 2.1 The Province'91 population. Percentage unemployment (%UE) and totals of unemployed persons (UE91), labour force (LAB91), population in 1991 (POP91) and number of households (HOU85) by municipality in the province of Central Finland in 1985.

ID	LABEL	STR	CLU	%UE	UE91	LAB91	POP91	HOU85
Urban				12.67	8022	63 314	129 460	49 842
1	Jyväskylä	1	1	12.20	4123	33786	67 200	26 881
2	Jämsä	1	2	11.07	666	6016	12 907	4663
3	Jämsänkoski	1	2	13.83	528	3818	8118	3019
4	Keuruu	1	2	12.84	760	5919	12 707	4896
5	Saarijärvi	1	3	14.62	721	4930	10 774	3730
6	Suolahti	1	5	15.12	457	3022	6159	2389
7	Äänekoski	1	3	13.17	767	5823	11 595	4264
Rural				12.63	7076	56 011	125 124	41 911
8	Hankasalmi	2	5	15.07	391	2594	6080	2179
9	Joutsa	2	6	9.38	194	2069	4594	1823
10	Jyväskylän mlk.	2	7	11.82	1623	13727	29 349	9230
11	Kannonkoski	2	4	18.64	153	821	1919	726
12	Karstula	2	4	13.53	341	2521	5594	1868
13	Kinnula	2	8	13.92	129	927	2324	675
14	Kivijärvi	2	8	15.63	128	819	1972	634
15	Konginkangas	2	3	21.04	142	675	1636	556
16	Konnevesi	2	5	12.91	201	1557	3453	1215
17	Korpilahti	2	1	11.15	239	2144	5181	1793
18	Kuhmoinen	2	2	12.91	187	1448	3357	1463
19	Kyyjärvi	2	4	11.31	94	831	1977	672
20	Laukaa	2	5	12.11	874	7218	16 042	4952
21	Leivonmäki	2	6	10.65	61	573	1370	545
22	Luhanka	2	6	10.34	54	522	1153	435
23	Multia	2	7	11.24	119	1059	2375	925
24	Muurame	2	1	9.79	296	3024	6830	1853
25	Petäjävesi	2	7	15.08	262	1737	3800	1352
26	Pihtipudas	2	8	13.02	331	2543	5654	1946
27	Pykkönmäki	2	4	17.98	98	545	1266	473
28	Sumiainen	2	3	12.80	79	617	1426	485
29	Säynätsalo	2	1	10.28	166	1615	3628	1226
30	Toivakka	2	6	11.72	127	1084	2499	834
31	Uurainen	2	7	16.47	219	1330	3004	932
32	Viitasaari	2	8	14.16	568	4011	8641	3119
Whole province				12.65	15 098	119 325	254 584	91 753

Sources: Statistics Finland: Population Census 1985. Statistics Finland (1992): Statistical Yearbook of Finland, Volume 87. Ministry of Labour of Finland (1991): Employment Service Statistics, November 30, 1991.