



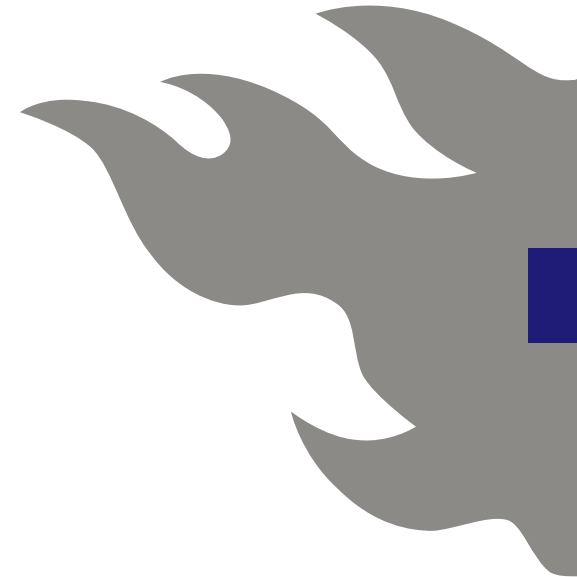
HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Otantamenetelmät (78143) Syksy 2011

TEEMAT 3 & 4

Risto Lehtonen

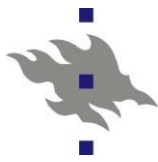
risto.lehtonen@helsinki.fi





Teema 3

ERITYISKYSYMYKSIÄ



Otannan erityiskysymyksiä

■ Ryväsotanta

- Survey sampling reference guidelines (2008 edition)
Introduction to sample design and estimation techniques
Section 3.2.5 Cluster sampling

■ Otoskoon määrääminen

- Survey sampling reference guidelines (2008 edition)
Introduction to sample design and estimation techniques
Section 3.3 Sample size determination

■ Vastauskadon hallinnan perusteita

- Lehtonen-Pahkinen (2004) Chapter 3
- VLISS Training Key 114, 117, 123



Sisäkorrelaatio ryväsotannassa

- Samaan rypääseen kuulumisella on taipumus samankaltaistaa alkioita tutkittavien ilmiöiden suhteen
- Koulututkimukset
 - Rypäänä opetusryhmä
 - Oppimistulokset, esim. [PISA](#)
- Työolotutkimukset
 - Rypäänä työpaikka
 - Työolot, esim. Kelan työterveyshuoltotutkimus
 - OHC data, [VLISS](#)



Sisäkorrelaatio

Sisäkorrelaatio $\hat{\rho}_{\text{int}}$ *Intra - cluster correlation*

Likimääräinen kaava $\hat{\rho}_{\text{int}} = (deff - 1) / (\bar{m} - 1)$

missä \bar{m} on keskimääräinen ryväskoko

Rypäät ovat usein positiivisesti sisäkorreloituneita

eli $\hat{\rho}_{\text{int}} > 0$ kun $deff > 1$

Lisäksi on voimassa: $deff(\hat{\rho}) = 1 + (\bar{m} - 1)\hat{\rho}_{\text{int}}$

missä $deff(\hat{\rho})$ on asetelmakerroin

(design effect)



Esimerkkiaineisto: Työterveyshuoltotutkimus OHC

■ Otanta-asetelma

- Ositettu yksi- ja kaksiasteinen ryväotanta
- Toimipaikat rypäinä
- Ositus rypään koon ja toimialan mukaan
 - Pienet toimipaikat: Yksiasteinen otanta
 - Suuret toimipaikat: Kaksiasteinen otanta

- Henkilötasolla likimain **itsepainottuva** (*self-weighting*) otos

■ Demonstraatioaineisto SAS-data OHC

■ Rajaus:

- Toimipaikat, joissa vähintään 10 työntekijää
- $H = 5$ ositetta (*strata*)
- $m = 250$ toimipaikkaa (ryvästä, *clusters*)
- $n = 7841$ henkilöä

- Vaihteleva määrä otosrypäitä per osite

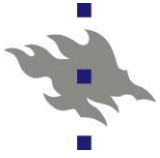


- **Deff-estimaatit OHC**
- **(Lehtonen&Pahkinen 2004)**

Table 5.8

Averages of design-effect estimates of proportion estimates of selected groups of binary response variables in the OHC Survey data set (number of variables in parentheses).

Study variable	Mean deff
Physical working conditions (12)	6.5
Psycho-social working conditions (11)	3.3
Psychosomatic symptoms (8)	2.0
Psychic symptoms (9)	1.8



Teema 4

OHJELMISTO



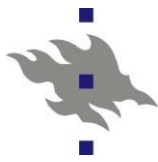
Ohjelmisto

- SAS-ohjelmisto
 - SAS-proseduurit
 - SAS-makrot
- SPSS module Complex Samples
- Stata:n svy-ohjelmat
- Erikoisalue: Pienalue-estimointi
 - SAS-makro EBLUPGREG
 - Ohjelma Domest
 - R-kieliset ohjelmat
- Software



SAS-ohjelmisto, kuvailu

- Survey-proseduurit, joilla otanta-asetelma (ositus, ryvästyminen, painotus) voidaan ottaa huomioon estimoinnissa
- SURVEYMEANS
 - Keskiarvot ja kokonaismäärät
- SURVEYREG
 - Regressioestimointi
- SURVEYFREQ
 - Ristiintaulukointi ja perustestit



SAS-ohjelmisto, analyysi

■ Survey-proseduurit

<http://support.sas.com/rnd/app/da/new/dasurvey.html#>

■ SURVEYFREQ

- Monipuolinen valikoima tilastollisia testejä

■ SURVEYREG

- Lineaariset mallit

■ SURVEYLOGISTIC

- Logistiset mallit

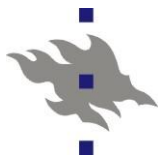
■ SURVEYPHREG

- Elinaikamallit



SAS-proseduuri SURVEYMEANS

- Asetelmaperusteinen estimointi
- Kokonaismäärien, keskiarvojen ja osuuksien estimointi koko aineistossa ja osajoukoissa
- Osajoukkoestimointi (domains)
 - BY statement
 - Ositekohtainen estimointi (*planned domains*)
 - DOMAIN statement
 - Estimointi muuntyyppisille osajoukoille (*unplanned domains*)



SAS-makrot, pienalue-estimointi

- EURAREA Project

- Asetelmaperusteinen estimointi
 - GREG-estimaattori
 - Lineaarinen kiinteiden tekijöiden malli
 - Varianssiestimaattori

- Malliperusteinen estimointi
 - EBLUP-estimaattori
 - Lineaarinen sekamalli
 - MSE:n estimaattori



Kuinka jatkaa eteenpäin?

- Yhteiskuntatilastotiede, kevät 2012
 - Otanta-aineistojen analyysi
- Pro gradu yhteiskuntatilastotieteen alalta
 - Tilastokeskus
 - Kelan tutkimusosasto
 - Muu valtionhallinnon tutkimuslaitos tai virasto