
Otantamenetelmät (78143)

Syksy 2008

OSA 3: Erityiskysymyksiä

Risto Lehtonen
risto.lehtonen@helsinki.fi

Otantamethodiikan 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

- Uudelleenpainotus ja imputointi. Perusteita.
Prof. Risto Lehtonen, Helsingin yliopisto, 2.12.2008

■ Tilastolliset ohjelmistot

Otanta-asetelmat (8)

■ (1) Alkiotasoinen otanta

(*element sampling*)

- Otantayksikkönä on perusjoukon alkio (esim. henkilö).
- Otos poimitaan valitulla otantamenetelmällä suoraan perusjoukon alkioiden muodostamasta kehikkoperusjoukosta
 - Väestörekisteri, toimipaikkarekisteri jne.

■ (2) Ryväotanta

(*cluster sampling*)

- Otantayksikkönä on perusjoukon alkioiden muodostama luonnollinen ryhmä eli **ryväs** (*cluster*)

□ Esim:

- Kunta, terveyskeskuspiiri
 - Terveys 2000
- Koulu, opetusryhmä
 - PISA

□ **Esimerkkejä ryväyksiköistä omalta toiminta-alueeltasi?**

Sisäkorrelaatio ryväotannassa

■ Ryväotanta-asetelmat

- Samaan rypäeseen kuulumisella on taipumus samankaltaistaa alkioita tutkittavien ilmiöiden suhteen
 - Koulututkimukset – rypäänä opetusryhmä
 - Oppimistulokset
 - Työolotutkimukset – rypäänä työpaikka
 - Työolot

Sisäkorrelaatio $\hat{\rho}_{int}$

Intra - cluster correlation

Likimääräinen kaava

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

missä \bar{m} on keskimääräinen ryväskoko

Rypäät ovat usein

positiivisesti sisäkorreloituneita

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

Lisäksi on voimassa:

$$deff(\hat{\rho}) = 1 + (\bar{m} - 1)\hat{\rho}_{int}$$

Esimerkkiaineisto: Työterveyshuoltotutkimus OHC

- **Otanta-asetelma**
 - Ositettu yksi- ja kaksiasteinen ryväotanta
 - Toimipaikat ryppäinä
 - Ositus ryppään 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)
 - $n = 7841$ 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

Kuinka jatkaa eteenpäin?

- Yhteiskuntatilastotiede, kevät 2009
 - Imputointimenetelmät
 - Painotusmenetelmät surveyssä
 - Pienalue-estimointi
 - Survey-seminaari
- Pro gradu yhteiskuntatilastotieteen alalta
 - Tilastokeskus
 - Kelan tutkimusosasto
 - Muu valtionhallinnon tutkimuslaitos tai virasto