

Otanta-aineistojen analyysi, kevät 2010

Aineopinnot (vapaaehtoinen harjoitustyö, 2 op)

HARJOITUSTYÖ

Harjoitustyö tehdään henkilökohtaisena työnä.

Lopputuotteena on noin 10 sivun raportti liitteineen, esimerkiksi rakenteella:

Kansilehti (otsikko, tekijä, kurssi, päiväys)

Tekstiosa (jaa soveltuviin lukuihin ja alilukuihin)

Viitteet (kirjallisuus)

Liitteet (valitut ohjelmakoodit ja tulostusotteet).

Työssä voi käyttää SAS-versiota 9.1.3 tai 9.2.

Työn palautus: **31.3.2010 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

Tehtävä 1.

Data-analyysia. Olkoon kiinnostuksen kohteena psyykinen rasittuneisuus (OHC-aineiston tulosmuuttajat PSYCH ja PSYCH2). Halutaan selvittää, millainen yhteys ilmiöllä on työoloihin ja sairastavuuteen. Tätä varten tulosmuuttujia mallinnetaan soveltuvan tilastollisen mallin sekä selittävien muuttujien sukupuoli (SEX), ikä (AGE), työn fyysiset haitat (PHYS) ja pitkäaikaissairastavuus (CHRON) avulla. Erityisesti kiinnostuksen kohteena olevat selittäjät ovat PHYS ja CHRON, joiden ikä-sukupuolivakioitu yhteys psyykkiseen rasittuneisuuteen halutaan selvittää.

Suorita seuraavat työvaiheet:

- a) Kuvaa lyhyesti analyysiaselman ja siihen liittyvän aineiston ja otanta-asetelman erityispiirteet. Mitä asioita tässä yhteydessä tulee ottaa huomioon, kun analyyttistä lähestymistapaa ja analyysityökaluja valitaan?
- b) Valitse analyysitilanteeseen soveltuva tilastollinen malli ja soveltuva SAS- ohjelmiston analyysiproseduuri molemmille tulosmuuttujille (jatkuva PSYCH ja binäärinen PSYCH2). Perustele, miksi päädyit valitsemaasi malliin ja analyysiproseduurin kummassakin tapauksessa.
- c) Sovita molemmille tulosmuuttujille malli asetelmaperusteisesti valitsemallasi analyysiproseduurilla. Etsi tulkittavaksi kelvollinen ”lopullinen” malli. Raportoi mallin valintaproseduurin tulokset taulukkomuodossa ja verbaalisesti. Raportoi lopullisen mallin parametristimaatit, keskivirheet, asetelmakertoimet, testaustulokset ja muut valitsemasi relevantit tunnusluvut. Pohdi diagnostisesti ”lopullisen” mallin hyvyttä.
- d) Suorita kohdan c) operaatiot molemmille tulosmuuttujille malliperusteisesti (SRS-oletus) soveltuvalla SAS-ohjelmiston analyysiproseduurilla. Vertaa tuloksia asetelmaperusteisiin tuloksiin ja raportoi vertailutulokset sekä johtopäätökset.