Frekvenssiaineistojen analyysi
Harjoitukset 1.11.2001

7.1. Tarkastellaan luennoilla käsiteltyä ikääntyviä autoilijoita (65-84-vuotiaita) koskevaa aineistoa, johon kuului 16262 henkilöä. Jokaista autoilijaa seurattiin 0-4 vuotta. Naisten seuranta-aika oli yhteensä 17300 vuotta ja miesten seuranta-aika 21400 vuotta. Naisille sattui seuranta-aikana yhteensä 175 onnettomuutta ja miehille 320. Onnettomuuksien lukumäärät noudattavat toisistaan riippumatta Poissonin jakaumaa.

(a) Mitä on uskottavuusfunktion maksimi?

(b) Uskottavuussuhteen arvo testattaessa hypoteesia, että miesten ja naisten onnettomuusaste on sama.

(c) Valkoisille kuljettajille sattui 348 onnettomuutta 29400 vuoden seuranta-aikana ja mustille 147 onnettomuutta 9300 vuoden seuranta-aikana. Testaa, poikkeavatko eri rotujen onnettomuusasteet merkitsevästi toisistaan. Laske 95%:n luottamusväli koko populaation todelliselle onnettomuusasteelle.

7.2. Tarkastellaan pykälässä 5.3.2 esitettyä rapuaineistoa (`crabs`). Aineistossa on muuttujat väri, selkäpiikin kunto, pituus, seuralaisten lukumäärä ja paino (`color, spine, width, satell, weight`). Kopioi aineisto verkosta.

(a) Selitä seuralaisten lukumäärää rapujen painolla. Sovita aineistoon Poissonin logaritmisesti lineaarinen malli $\log \mu(x) = \alpha + \beta x$ ja estimoi 2.44 kg painavien rapunaaraiden seuralaisten lukumäärä.

(b) Laske $\hat{\beta}$:n luottamusväli. Kuinka paljon painon tulee kasvaa, jotta seuralaisten lukumäärä kasvaa keskimäärin yhdellä?

7.3. Tarkastellaan edelleen Poissonin logaritmisesti lineaarista mallia $\log \mu(x) = \alpha + \beta x$, kun selitetään seuralaisten lukumäärää rapujen painolla.

(a) Mitä ovat uskottavuusfunktion ja logaritmoidun uskottavuusfunktion maksimit?

(b) Uskottavuusfunktion maksimi, jos hypoteesi "$H_0 : \beta = 0$" pitää paikkansa?

(c) Mikä on uskottavuussuhteen logaritmin arvo?

7.4. Tarkastellaan edelleen rapuaineistoa.

(a) Selitetään seuralaisten lukumäärää painolla ja värillä (nominaaliasteikollinen). Tulkitse parametrien estimaatit.

(b) Estimoi seuralaisten lukumäärä, kun paino on 2.44 kg ja rapu tumma (color=4).

(c) Testaa, onko väri merkitsevä selittäjä.

7.5. Oletetaan että uudet melanomatapaukset Taulukon 5.5 eri luokissa noudattavat toisistaan riippumatta Poissonin jakaumaa (Gail 1978 ja Koch, Imrey, et al. 1985). Taulukon väkiluku on riskin alaisen väestön määrä. Selitetään melanoman esiintymisastetta (tapaukset/väkiluku) iällä ja asuinalueella. Sovitetaan aineistoon logaritmisesti lineaarinen malli.

Taulukko 5.5 Uudet melanomatapaukset valkoihoisessa miesväestössä kahdella eri alueella vuosina 1969-1971

Ikäryhmä	Alue	Tapauksia	Väkiluku
35-44	etelä	75	220407
45-54	etelä	58	198119
55-64	etelä	63	134084
65-74	etelä	45	70708
75+	etelä	27	34233
<35	etelä	64	1074246
35-44	pohjoinen	76	564535
45-54	pohjoinen	98	592983
55-64	pohjoinen	104	450740
65-74	pohjoinen	63	270908
75+	pohjoinen	80	161850
<35	pohjoinen	61	2880262