Frekvenssiaineistojen analyysi
Harjoitukset 15.11.2001

9.1. Selitetään ravun pituudella ja värillä todennäköisyyttä, että rapunaaraalla on seuralaisia. Väri (`color`) on luokittelumuuttuja (1=vaalea, 2=keskivärinen, 3=melko tumma, 4=tumma). Estimoi malli, jossa `color`-muuttujaa käsitellään kvantitatiivisena. Olkoot muuttujan arvot $\{1, 2, 3, 4\}$, kuten alkuperäisessä aineistossa. Vertaa saatua mallia malliin, jossa `color` on luokittelumuuttuja. Kumpi malli on parempi?

9.2. Tarkastellaan Taulukon 6.7 aineistoa, jossa on tutkimustuloksia AZT-lääkkeen hidastavaa vaikutusta AIDS:in oireiden ilmaantumiseen. Tutkimukseen osallistuneista 338:sta tartunnan saaneesta henkilöstä osa arvottiin hoitoryhmään (AZT). Vertailuryhmään kuuluneet saivat lääkettä jos oireita ilmaantui.

Taulukko 6.7 AIDSin oireiden kehittyminen, AZTn käyttö ja rotu.

Rotu	AZT	Oireita	Ei oireita
Valkoinen	Kyllä	14	93
	Ei	32	81
Musta	Kyllä	11	52
	Ei	12	43

Estimoi aineistosta logit-malli

$$logit(\pi) = \alpha + \beta_i^X + \beta_k^Z, \qquad (6.8.1)$$

joka esittää AZT:n (X) vaikutusta oireiden esiintymiseen parametrien β_1^X ja β_2^X avulla sekä rodun (Z) vaikutusta parametrien β_1^Z ja β_2^Z avulla. Mitä ovat parametrien estimaatit, kun

(a) $\beta_2^X = \beta_1^Z = 0$,

(b) $\beta_1^X = \beta_1^Z = 0$ ja

(c) $\beta_1^X = -\beta_2^X$ ja $\beta_1^Z = -\beta_2^Z$?

9.3. (Jatkoa edelliseen tehtävään.) Testaa logit-mallin (6.8.1) avulla, ovatko

(a) X ja Y ehdollisesti riippumattomia, kun Z on annettu ja

(b) vallitseeko homogeeninen riippuvuus $\theta_{XY(Valkoinen)} = \theta_{XY(Musta)}$.

(c) Vertaa tuloksia vastaaviin CMH-testin ja BD-testin tuloksiin (4. luku).

9.4. Taulukossa (6.8) on raportoitu passiivisen tupakoinnin vaikutuksia keuhkosyövän esiintymiseen (Tehtävä 5.4.). Testaa mallin (6.8.1) avulla hypoteesi, että

(a) keuhkosyöpä on riippumaton passiivisesta tupakoinnista, kun maa on kontrollimuuttuja ($k = 1, 2, 3$).

(b) $H_0 : \theta_{XY(Japani)} = \theta_{XY(I-B)} = \theta_{XY(USA)}$.

Taulukko 6.8 Passiivinen tupakointi ja keuhkosyöpä.

Maa	Puoliso polttaa	Syöpätapaukset	Verrokit
Japani	Ei	21	82
	Kyllä	73	188
Iso-Britania	Ei	5	16
	Kyllä	19	38
Yhdysvallat	Ei	71	249
	Kyllä	137	363

9.5. Tarkastellaan 100 pyöräilijän otoksessa pyörän tyypin (maastopyörä, muu) ja kypärän käytön välistä riippuvuutta (Taulukko 6.9). Sovita aineistoon logaritmisesti lineaarinen riippumattomuusmalli sekä täysi (kyllästetty) malli.

Taulukko 6.9 Pyörän tyyppi ja kypärän käyttö.

	Kypärä	
Pyörän tyyppi	Kyllä	Ei
Maastopyörä	34	32
Muu	10	24