Logistinen regressio
Harjoitukset 5.4.2001

9.1. Tarkastellaan luentomonisteessa esitettyä (7.2 Suklaaesimerkki) suklaamakeisten valintaa koskevaa esimerkkiä. Analysoi koetulos, kun käytät selittäjinä vain muuttujia `DARK` ja `SOFT`.

(a) Estimoi valintamalli ja tulkitse tulos.

(b) Laske eri vaihtoehtojen valintatodennäköisyydet.

9.2. Oletetaan kolmen vaihtoehdon valintatilanne, jossa vaihtoehtoja luonnehtii yksi muuttuja x. Koska henkilön $i,\ i = 1, 2, \dots, n$ valintatodennäköisyydet toteuttavat logitmallin (7.3.1), niin

$$logit[\gamma_j(x_{ij})] = \alpha_i + \beta x_{ij},\ j = 1, 2, 3.$$

Osoita, että ehdolliset todennäköisyydet $P(Y_{ij} = 1|S_i = 1)$ ovat muotoa (7.3.3), missä Y_{ij} on valintaa osoittava muuttuja ja S_i on i. henkilölle sallittu valintojen lukumäärä.

9.3. Olkoon edellisen tehtävän valintamallin tilanteessa $x_1 = 3, x_2 = 1$ ja $x_3 = 5$. Nämä arvot ovat samat kaikilla henkilöillä. Laske kaikkien valintojen a ja b suhteelliset mahdollisuudet $\frac{\pi_a}{\pi_b}$.

9.4. Muodosta Taulukon 8.1 (Luentomoniste) vastinpariaineistosta logitmalliin perustuva uskottavuusfunktio ja logaritmoitu uskottavuusfunktio. Laske siitä suurimman uskottavuuden estimaatti.

9.5. Estimoi Taulukon 8.3 Diabetes-aineistosta (Luentomoniste) logitmalliin (8.2.2) perustuva parametri β ehdollisella suurimman uskottavuuden menetelmällä. Tulkitse estimaatti.