Logistinen regressio
Harjoitukset 22.3.2001

7.1. Selitetään logistista regressiomallia käyttäen keskosen syntymisen todennäköisyyttä äidin ominaisuuksilla ja raskaudenaikaisilla tapahtumilla (Aineisto on `lowbwt`). Hae mallinvalintakriteerien AIC ja SC mielessä parhaat mallit. Tarkista Hosmerin ja Lemeshowin testillä mallin yhteensopivuus aineiston kanssa.

7.2. Kokeile, mitä malleja etenevä valinta, askeltava valinta ja taaksepäin eliminointi ehdottavat. Vertaile myös näiden mallien antamia mallinvalintakriteerien AIC ja SC arvoja sekä yhteensopivuutta Hosmerin ja Lemeshowin testisuureen mielessä.

7.3. Tarkastellaan suhteellisten mahdollisuuksien mallia (6.3.1). Oleta, että vastemuuttujan lukkien lukumäärä $J = 4, \beta = 1$ ja $\alpha_j = j,\ j = 1, 2, 3$. Tarkastele, miten todennäköisyydet $P(Y \leq j),\ j = 1, 2, 3$ käyttäytyvät? Entä miten käyttäytyvät todennäköisyydet

$$P(Y = j) = P(Y \leq j) - P(Y \leq j - 1)?$$

7.4. Oleta `wallet` aineistossa muuttuja `wallet` järjestysasteikolliseksi. Käytä suhteellisten mahdollisuuksien mallia, kun selitettävänä on `wallet` ja selittäjinä muut aineiston muuttujat. Lausu mallin yhtälöt parametrien estimaatteja käyttäen. Mitä on

(a) P(`wallet`=1), kun `business`=1, `punish`=1, `explain`=1?

(b) P(`wallet`=3), kun `business`=1, `punish`=0, `explain`=1?

(c) P(`wallet`=2), kun `business`=0, `punish`=3, `explain`=0?

7.5. Tarkastellaan luentoesimerkin aineistoa `happy` ja kahta eri logitmallia, joista ensimmäisessä on selitettävänä `lesshap` ja toisessa `nottoo`. Perustele ensin yleisesti, miksi kummankin mallin kertoimien pitäisi olla samoja, jos suhteellisten mahdollisuuksien malli pitää paikkansa. Tarkastele sitten, pitääkö suhteellisten mahdollisuuksien malli paikkansa tässä aineistossa.