Logistinen regressio
Harjoitukset 15.2.2001

2.1. Selitä logistisella regressiomallilla aineistosta `chddata` (`http://mtl.uta.fi/tilasto/logreg/Datat/`) mahdollisuutta sairastua sepelvaltimotautiin, kun selittäjänä on ikä (vrt. tehtävä 1.5). Estimoi malli ja tulkitse tulos.

2.2. Laske edellisessä tehtävässä estimoidusta mallista vedonlyöntisuhde siten, että vertaillaan mahdollisuuksia sairastua, kun ikäero on 10 vuotta.

2.3. Estimoi aineistosta `penalty` (luentoesimerkki) GENMOD proseduurilla malli, jossa selitetään kuolemanrangaistuksen mahdollisuutta, kun selittäjänä on syytetyn rotu, uhrin rotu, teon törkeys (culp) sekä syytetyn rodun ja uhrin rodun välinen yhdysvaikutus. Kirjoita malli logistisena regressiomallina ja laske kaikilla selittäjien arvokombinaatioilla todennäköisyydet saada kuolemanrangaistus.

2.4. Oletetaan logitmalli

$$\log(\frac{\pi_i}{1-\pi_i}) = \beta\ i,\ \ i = 1, 2, 3.$$

Olkoot $n_1 = n_2 = 3$ ja $n_3 = 4$ havaintojen lukumäärät eri i:n arvoilla (havaintoja yhteensä $n_1 + n_2 + n_3 = 10$. Merkitään "onnistumisten" eli ykkösten lukumäärää y_i, $i = 1, 2, 3$ ja olkoon $y_1 = y_2 = 1$ ja $y_3 = 2$.

(a) Kirjoita uskottavuusfunktio ja

(b) logaritmoitu uskottavuusfunktio.

2.5. Lausu edellisen tehtävän tapauksessa

(a) parametrin β derivaatta $L'(\beta)$ ja

(b) suurimman uskottavuuden estimointiyhtälö.

2.6. Estimoi tehtävän 2.4 logitmalli SAS:illa (NOINT optio MODEL lausees- sa antaa mallin, jossa ei ole vakiotermiä).