Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Faglig kontakt under eksamen:
Nikolai Ushakov 918 04 616

EXAM IN TMA4315 GENERALIZED LINEAR MODELS

Monday 30 November 2009
Time: 09:00–13:00

Tillatte hjelpemidler:
Tabeller og formler i statistikk, Tapir forlag,
K.Rottmann, Matematisk formelsamling,
Ett gult A4-ark med IMF-stempel med egne håndskrevne formler og notater,
Godkjent enkel kalkulator.

Sensur: 14 Desember 2009

Oppgave 1

N ($N > 100$) independent observations $Y_1, ..., Y_N$ have normal distributions with the same variance σ^2.

$$EY_1 = \mu_1 = \beta_1 + \beta_2,$$
$$EY_2 = EY_3 = ... = EY_N = \mu = \beta_1,$$

where β_1 and β_2 are parameters of interest.

a) What is the design matrix X?

b) Using the simple rule of thumb $h_{ii} > 2p/N$ show that the first observation (and only it) is highly influential (here $H = [h_{ij}]$ is the hat matrix, and p is the number of parameters, $p = 2$).

c) Find the maximum likelihood estimator of $\beta = (\beta_1, \beta_2)^T$.

Recall:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Oppgave 2

X_1 and X_2 are two independent random variables having the same distribution with the probability density function

$$f_X(x;\theta) = \theta x^{\theta-1} I_{(0,1)}(x),\ \ \theta > 0.$$

Let $Y = \max\{X_1, X_2\}$.

a) Does the distribution of Y belong to the exponential family?

b) Show that

$$E \ln Y = -\frac{1}{2\theta}.$$

Oppgave 3

Consider the following GLM: $Y_1, ..., Y_N$ are independent, $N = 2n$; Y_i ~binomial(n_i, π_i), i=1,...,N (n_i are known); $\beta = (\beta_1, \beta_2)^T$ are parameters of interest;

$$\ln \frac{\pi_i}{1-\pi_i} = x_i^T \beta,$$

where the design matrix has form

$$X = \begin{bmatrix} a_1 & 0 \\ & \cdots \\ a_n & 0 \\ 0 & c_1 \\ & \cdots \\ 0 & c_n \end{bmatrix}.$$

a) Find the score vector (in terms of Y_i, n_i, π_i, a_i, c_i) and show that b_1 does not depend of $Y_{n+1}, ..., Y_N$ while b_2 does not depend of $Y_1, ..., Y_n$, where $b = (b_1, b_2)^T$ is the maximum likelihood estimator of $\beta = (\beta_1, \beta_2)^T$.

b) Find the information matrix and show that the method of scoring for b has the form

$$b_1^{(m)} = b_1^{(m-1)} + \frac{\sum_{i=1}^{n}(Y_i - n_i\pi_i^{(m-1)})a_i}{\sum_{i=1}^{n} n_i\pi_i^{(m-1)}(1-\pi_i^{(m-1)})a_i^2},$$

$$b_2^{(m)} = b_2^{(m-1)} + \frac{\sum_{i=n+1}^{N}(Y_i - n_i\pi_i^{(m-1)})c_{i-n}}{\sum_{i=n+1}^{N} n_i\pi_i^{(m-1)}(1-\pi_i^{(m-1)})c_{i-n}^2}.$$