StAN Exercise Sheet 3

Richard D. Gill
Mathematical Institute, University of Leiden, Netherlands
http://www.math.leidenuniv.nl/~gill

18 October, 2012

1 Fisher information

1.1 Poisson distribution

Suppose X is a single observation from the Poisson distribution with mean λ: thus $\mathrm{P}(X = x) = p(x|\lambda) = \lambda^x e^{-\lambda}/x!$, $x = 0, 1, \ldots$. The Fisher information is defined by $I(\lambda) = \mathrm{E}_\lambda((\mathrm{d}\log p(X|\lambda)/\mathrm{d}\lambda)^2)$.

Compute the Fisher information.

Verify that $\mathrm{E}_\lambda(\mathrm{d}\log p(X|\lambda)/\mathrm{d}\lambda) = 0$ (expected score equals zero).

Verify that $\mathrm{E}_\lambda(-\mathrm{d}^2\log p(X|\lambda)/\mathrm{d}\lambda^2) = I(\lambda)$.

Compute the maximum likelihood estimator of λ based on an i.i.d. sample of size n from this distribution.

What are mean, variance and mean square error of the m.l.e.? Verify that both mean square error and variance behave like C/n for $n \to \infty$, same C (i.e., n times these quantities converge to $C > 0$ as $n \to \infty$.

Let $\ell(\lambda)$ denote the log likelihood for λ based on these n observations. Show that the negative inverse of $\mathrm{d}^2\ell/\mathrm{d}\lambda^2$, evaluated at λ equal to the m.l.e., also behaves like C/n for $n \to \infty$, same C.

1.2 Normal distribution

Repeat the previous exercise for the case of the normal distribution $N(\mu, 1)$ with unknown mean μ.

2 Least squares as m.l.e. under normality

2.1 Simple linear regression

Suppose that $x_1, \ldots, x_n$ are known constants. Suppose that for unknown parameters α, β and σ, observations Y_i are built up as follows: $Y_i = \alpha + \beta x_i + \epsilon_i$, where the ϵ_i are independent (unobserved) errors drawn from the $N(0, \sigma^2)$ distribution.

Compute the m.l.e.'s of α, β and σ.

Hint: do this first in the situation that $\sum_i x_i = 0$. Derive the result for the general situation by writing $\alpha + \beta x_i = (\alpha + \beta \overline{x}) + \beta(x_i - \overline{x}) = \alpha' + \beta' x_i'$ where $\alpha' = \alpha + \beta \overline{x}$, $\beta' = \beta$, $x_i' = x_i - \overline{x}$, $\overline{x} = \sum_i x_i / n$, a known constant. Note that the two statistical models for the data $Y_1, \ldots, Y_n$ – model 1, $Y_i = \alpha + \beta x_i + \epsilon_i$, and model 2, $Y_i = \alpha' + \beta' x_i' + \epsilon_i$ – describe exactly the same family of possible distributions of the data Y_i as the parameters (α, β, σ) and $(\alpha', \beta', \sigma)$ vary freely.