STA 410/2102, Fall 2000 — Assignment #2

Due in class on November 14. Worth 16% of the final mark.

When a species is introduced into a favourable new environment, its population will at first increase exponentially. Eventually, however the population will approach the maximum that the environment can support, and population growth will level off. One theoretical model for this process is based on the "logistic" curve. The number of organisms present at time t is modeled as

$$C(t) \;=\; \frac{M}{1+e^{-\kappa(t-\theta)}}$$

where M, κ, and θ are parameters of the curve. M is the maximum size of the population, θ is the time when half this maximum size is reached, and κ controls how rapidly the population grows.

Suppose that biologists perform an experiment in which a small number of insects of a certain species are introduced to an isolated island. At n times, t_i, following this introduction, the biologists set out traps, and see how many organisms of this species they collect. (The number collected is a negligible part of the total population, so it doesn't affect the population growth.) The biologists believe that their traps operate such that the number trapped has a Poisson distribution with mean equal to the population size times an unknown constant, γ. That is, the probability of trapping c_i insects at time t_i is

$$P(c_i) \;=\; \frac{\lambda(t_i)^{c_i}}{c_i!}e^{-\lambda(t_i)}$$

where $\lambda(t_i) = \gamma C(t_i)$.

Because γ is unknown, the biologists cannot hope to determine M — only the product γM is identifiable. To avoid possible problems with negative means, we'll work with $\eta = \log(\gamma M)$ as one of the parameters, with κ and θ being the others.

The likelihood function from the n counts of trapped insects is therefore

$$L(\eta,\kappa,\theta) \;=\; \prod_{i=1}^{n} \frac{\lambda(t_i)^{c_i}}{c_i!}e^{-\lambda(t_i)}$$

with

$$\lambda(t) \;=\; \frac{e^{\eta}}{1+e^{-\kappa(t-\theta)}}$$

Your task is to write an S/R program that uses Newton-Raphson iteration to find the maximum likelihood estimates for η, κ, and θ for a given data set, $(t_1, c_1), \ldots, (t_n, c_n)$, where the times t_i are positive real numbers and the counts c_i are non-negative integers. Also find an estimate for the the variance-covariance matrix of the maximum likelihood estimator, based on the observed information matrix.

You should investigate whether or not Newton-Raphson iteration always converges for this problem, and come up with some way of selecting reasonable starting values for the iteration.

Test your program on the data in the file `/u/radford/ass2.data` on utstat or CQUEST (it's also available from the web page). You can read this data into a data frame with the command:

```
> data <- read.table("/u/radford/ass2.data",head=T)
```