STAT 703/J703
January 30th, 2007
-Lecture 5-

Instructor: Brian Habing
Department of Statistics
LeConte 203
Telephone: 803-777-3578
E-mail: habing@stat.sc.edu

8.5 – Maximum Likelihood

The idea behind maximum likelihood estimation is to find the parameters that seem most likely to have resulted in the observed statistic.

In the case of more than one observation, the likelihood is:

$\text{lik}(\theta|x_1,\ldots x_n)=f(x_1,\ldots x_n|\theta)=\Pi f(x_i|\theta)$

It is unpleasant to find the maximum of a product though…

An easier function to work with is the log likelihood

$L(q)=\log(\text{lik}(\theta))$

$=\log(\prod f(x_i|\theta))$

$=\sum \log(f(x_i|\theta))$

STAT 702/J702 B.Habing Univ. of S.C. 4

MLE Example 1 (Section 8.5.1)

Consider a multinomial experiment, where each of the n observations has probability p_i of falling in cell i=1…m.

$$f(x_1,\dots x_m \mid p_1\dots,p_m)=\frac{n!}{\prod_{i=1}^{m} x_i!}\prod_{i=1}^{m} p_i^{x_i}$$

STAT 702/J702 B.Habing Univ. of S.C. 5

Often there is some restriction on the probabilities however. Consider the case in genetics where a gene occurs with frequence θ in the population.

STAT 702/J702 B.Habing Univ. of S.C. 6

Say we observe 342 ++, 500 +- and 187 --. What is our estimate of θ?

MLE example 2 – Logistic Regression

In linear regression we try to predict one continuous variable from another.

Under a few basic assumptions this results in fairly easy parameter estimation and use of t and F distributions.

Consider the case of attempting to predict a 0,1 variable from a continuous variable.

$$P[Y_i = 1 \mid x_i] = \frac{1}{1 + e^{-(\alpha + \beta x_i)}}$$

8.5.2 – Large Sample Properties

Theorem A: The MLE is Consistent (under appropriate regularity conditions)

Information function

$$I(\theta) = E\left[\frac{\partial}{\partial\theta}\log f(X \mid \theta)\right]^2$$

$$= -E\left[\frac{\partial^2}{\partial\theta^2}\log f(X \mid \theta)\right]^2$$

Theorem B: Under appropriate regularity conditions the MLE is asymptotically normal with mean θ and variance $\frac{1}{nI(\theta)}$.