Chapter 2

Likelihood, Estimation and Testing

2.1 Likelihood and log-likelihood.

In this and the following section, we review briefly the basic ideas and results of likelihood inference: details may be found in any standard mathematical statistics text for beginning graduate students. A vector of data random variables, $\mathbf{Y}$, whose value $\mathbf{y}$ is observed, has one of a family of probability distributions $\{P_\theta; \theta \in \Theta\}$, indexed by a *parameter* θ in *parameter space* Θ. The goals of estimation are to make inferences about which P_θ gave rise to the observed $\mathbf{y}$, and to assess the uncertainty associated with this inference.

The *likelihood* is $L_{\mathbf{y}}(\theta) = P_\theta(\mathbf{y})$, a function of θ. The likelihood provides the connection between the data $\mathbf{y}$ and the probability model P_θ. A *statistic* is a function of the data random variables $\mathbf{Y}$, an *estimator* $T = T(\mathbf{Y})$ is a statistic taking values in Θ, while the *estimate* is $T(\mathbf{y})$, the value taken by the estimator that is used to estimate θ.

For example, suppose Y_i, $i = 1, ..., n$ are independent identically distributed Bernoulli random variables, $B(1, \theta)$, the indicators of success in n independent trials, each with success probability θ. Then $P_\theta(y) = \theta^y(1-\theta)^{1-y}$ $(y = 0, 1)$ for each trial, and $L(\theta) = \prod_1^n(\theta^{y_i}(1-\theta)^{1-y_i})$. The log-likelihood is

$$\ell(\theta) = \log L(\theta) = (\sum_1^n y_i)\log(\theta) + (n - \sum_1^n y_i)\log(1-\theta). \tag{2.1}$$

Note that the (log)-likelihood depends only on the value of $T = \sum_1^n Y_i$, the total number of successes, which has a binomial $B(n, \theta)$ distribution. The likelihood based on the binomial probability of the observed value t of T is

$$\begin{aligned} L(\theta) &= P_\theta(T = t) = \frac{n!}{k!(n-k)!}\theta^t(1-\theta)^{n-t} \\ \ell(\theta) &= \log L(\theta) = \text{const} + t\log(\theta) + (n-t)\log(1-\theta). \end{aligned} \tag{2.2}$$