2.5 Maximum Likelihood Methods

Ming Lu

2022-04-02

Let be a random variable for which we have sample observations , and let be a vector of parameters, pertaining to some statistical model for , and which we want to estimate. Assuming that the observations are made randomly and independently of each other, we can consider their joint distribution as the product of their conditional distributions (Eq. (2.36)). Then, if the distribution has density function , we define the likelihood of observing the values of , given , as

The value of that maximizes this likelihood function is the maximum-likelihood estimate (MLE) of . This MLE might not be unique, and might even not exists. In the case that it exists, but is not unique, any of the points where the maximum is attained is an MLE of . A useful mathematical trick for simplifying the computation of an MLE is to find instead a value of that maximizes the logarithm of , called the log likelihood function:

Maximizing is the same as maximizing , because is a monotone function. The optimization of is often obtain only by numerical methods. In this can be done with the function , or by direct use of the optimization method optim().

A further simplification to the likelihood computation is to assume that all random observations are made independently. In this case Eq. (2.48) and Eq. (2.49) turn into

Although the independence assumption is quite strong for time series, it simplifies likelihood computation to analytical form. As an example of the MLE method we show how to obtain the best estimate of a constant variance from sample returns. Example Let be a sequence of independently observed log returns. We assume these returns have some constant variance , which we want to estimate under the normal distribution model, and a constant mean which we know. By the independence of the observations the log likelihood function is as in Eq. (2.50), with and fixed, and the probability density functionf is given by the Gaussian density function (Eq. (2.27)). Therefore, the log likelihood of observing is

By the maximum likelihood method the value of that maximizes (2.51) is the best estimate (or more likely) value of . Differentiating with respect to and setting the resulting derivative to zero, we obtain that the maximum likelihood estimate of is given by

Note that this estimated value is biased (cf. Remark 2.2), since

However, the bias of tends to 0 as , so the maximum likelihood estimate is asymptotically unbiased for . Thus, taking sufficiently large we have a good estimator of the variance of log returns, sampled in a given interval of time, and which in fact, by construction, it is the most likely value that can be assigned to that variance.