Econometrics – Maximum Likelihood Estimation

William M Volckmann II

January 3, 2017

1 MLE Redux

We already know a little bit about maximum likelihood estimation (mle). We've seen the **likelihood function**,

$$L(\theta; x) = \prod_{i=1}^{n} f(x_i; \theta),$$

where we treat L as a function of θ, hence often writing $L(\theta)$. We've used the **log likelihood function**,

$$l(\theta) = \ln\big(L(\theta)\big) = \sum_{i=1}^{n} \ln\big(f(x_i; \theta)\big),$$

which is often easier to work with than the likelihood function itself.

The maximizer of $L(\theta)$, which is also the maximizer of $l(\theta)$, is our point estimator for θ, which we denote $\widehat{\theta}$ and call it the **maximum likelihood estimator** of θ. In these notes, we are going to give theoretical justification for considering the mle. Because the mle might not have an analytical solution; it might not be unique; hell, it might not even exist. So we need to tighten up our understanding of the mle.

2 Newton's Method of Iteration

If there is no analytical solution for the mle, then we can use numerical analysis by way of **Newton's Method of Iteration** to approximate it. We make an initial guess at the solution, $\widehat{\theta}_0$. Take the second-order Taylor expansion of $l(\theta)$ around $\widehat{\theta}_0$,

$$l(\theta) \approx l(\widehat{\theta}_0) + l'(\widehat{\theta}_0)(\theta - \widehat{\theta}_0) + \frac{1}{2} l''(\widehat{\theta}_0)(\theta - \widehat{\theta}_0)^2.$$

Take the derivative with respect to θ, set it equal to zero, and solve for θ to get the subsequent iteration,

$$l'(\widehat{\theta}_0) + l''(\widehat{\theta}_0)(\theta - \widehat{\theta}_0) := 0 \quad \Longrightarrow \quad \widehat{\theta}_1 = \widehat{\theta}_0 - \frac{l'(\widehat{\theta}_0)}{l''(\widehat{\theta}_0)}.$$

Proceed with the same chain of logic but with the new guess $\widehat{\theta}_1$. Continue with the series of iterative guesses

$$\widehat{\theta}_1 = \widehat{\theta}_0 - \frac{l'(\widehat{\theta}_0)}{l''(\widehat{\theta}_0)} \quad \Longrightarrow \quad \widehat{\theta}_2 = \widehat{\theta}_1 - \frac{l'(\widehat{\theta}_1)}{l''(\widehat{\theta}_1)} \quad \Longrightarrow \quad \ldots$$

and so on. In general, the sth iteration will be

$$\widehat{\theta}_s = \widehat{\theta}_{s-1} - \frac{l'(\widehat{\theta}_{s-1})}{l''(\widehat{\theta}_{s-1})}.$$

Iteration continues until the change in each iteration and the change in $l(\widehat{\theta}_s)$ is deemed sufficiently small.

3 Asymptotic Results

We start with the following **regularity conditions**.

(a) The pdfs are distinct for each θ. That is, if $\theta \neq \theta'$, then $f(x_i;\theta) \neq f(x_i;\theta')$.

(b) The pdfs have common support for all θ.

(c) The **true value** of θ, denoted θ_0, is an interior point of the sample space. Doing so gives a new $\widehat{\theta}_2$,