Asymptotic Distribution of MLE

Maura Mezzetti

Department of Economics and Finance
Università Tor Vergata

Outline

Score Function and Fisher Information

Score Function

$$u(\theta) = \frac{\partial logL(\theta|x)}{\partial\theta}$$

The first derivative of the log-likelihood function is called Fisher's score functio

Score Function and Fisher Information

Score Function

$$u(\theta) = \frac{\partial logL(\theta|x)}{\partial\theta}$$

The first derivative of the log-likelihood function is called Fisher's score functio

The score vector

Important results for the score vector:

Let $\hat{\theta}_{MLE}$ be the MLE for θ

$$u(\hat{\theta}_{MLE}) = 0$$

Under some regular conditions (see Casella Berger):

$$E[u(\theta)] = 0$$

The score vector

Important results for the score vector:

Let $\hat{\theta}_{MLE}$ be the MLE for θ

$$u(\hat{\theta}_{MLE}) = 0$$

Under some regular conditions (see Casella Berger):

$$E[u(\theta)] = 0$$

under some regular conditions (see Casella Berger):

$$E[u^2(\theta)] = E\left[\left(\frac{\partial logL(\theta|x)}{\partial\theta}\right)^2\right] = -E\left[\frac{\partial^2 logL(\theta|x)}{\partial\theta^2}\right]$$

The score vector

Important results for the score vector:

Let $\hat{\theta}_{MLE}$ be the MLE for θ

$$u(\hat{\theta}_{MLE}) = 0$$

Under some regular conditions (see Casella Berger):

$$E[u(\theta)] = 0$$

under some regular conditions (see Casella Berger):

$$E[u^2(\theta)] = E\left[\left(\frac{\partial logL(\theta|x)}{\partial\theta}\right)^2\right] = -E\left[\frac{\partial^2 logL(\theta|x)}{\partial\theta^2}\right]$$

Score Function and Fisher Information

Score Function

$$u(\theta) = \frac{\partial logL(\theta|x)}{\partial\theta}$$

The first derivative of the log-likelihood function is called Fisher's score functio

Expected Fisher Information $I_n(\theta) = E_\theta\left(-\frac{\delta^2 logL(\theta)}{\delta^2\theta}\right)$

$$I_n(\theta) = E\left(\left(\frac{\partial}{\partial\theta}log(L(x;\theta)\right)^2\right)$$

$$I_n(\theta) = var\left(u(\theta)\right)$$

Score Function and Fisher Information

Score Function

$$u(\theta) = \frac{\partial logL(\theta|x)}{\partial\theta}$$

The first derivative of the log-likelihood function is called Fisher's score functio

Expected Fisher Information $I_n(\theta) = E_\theta\left(-\frac{\delta^2 logL(\theta)}{\delta^2\theta}\right)$

$$I_n(\theta) = E\left(\frac{\partial}{\partial\theta} log(L(x;\theta))^2\right)$$

$$I_n(\theta) = var\left(u(\theta)\right)$$

Score Function and Fisher Information

Score Function

$$u(\theta) = \frac{\partial log L(\theta|x)}{\partial \theta}$$

The first derivative of the log-likelihood function is called Fisher's score functio

Expected Fisher Information $I_n(\theta) = E_\theta\left(-\frac{\delta^2 log L(\theta)}{\delta^2 \theta}\right)$

$$I_n(\theta) = E_\theta\left(-\frac{\delta^2 log L(\theta)}{\delta^2 \theta}\right) == E\left(\left(\frac{\partial}{\partial \theta} log(L(x;\theta)\right)^2\right) = var\left(u(\theta)\right)$$

$$I(\theta) = E\left(\left(\frac{\partial}{\partial \theta} log(f(x;\theta)\right)^2\right)$$

$$I_n(\theta) =$$

Score Function and Fisher Information

Score Function

$$u(\theta) = \frac{\partial log L(\theta|x)}{\partial \theta}$$

The first derivative of the log-likelihood function is called Fisher's score functio

Expected Fisher Information $I_n(\theta) = E_\theta\left(-\frac{\delta^2 log L(\theta)}{\delta^2 \theta}\right)$

$$I_n(\theta) = E_\theta\left(-\frac{\delta^2 log L(\theta)}{\delta^2 \theta}\right) == E\left(\left(\frac{\partial}{\partial \theta} log(L(x;\theta)\right)^2\right) = var\left(u(\theta)\right)$$

$$I(\theta) = E\left(\left(\frac{\partial}{\partial \theta} log(f(x;\theta)\right)^2\right)$$

$$I_n(\theta) =$$

Central Limit Theorem

Theorem Let $X_1, X_2, \ldots$ be a sequence of IID r.v. with expectation μ and variance σ^2. Then, the random variable:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\sqrt{n}}{\sigma}\left(M_n - \mu\right)$$

converge in law to a standard normal $N(0,1)$.
The Law of Large Numbers and the Central Limit Theorem require specific integrability assumptions: finite expectation for the Law of Large Numbers, and finite variance for the Central Limit Theorem. When these assumptions are not satisfied, the conclusions may not be true.

Reminder: Taylor Expansion

Taylors Theorem

If f is a function continuous and n times differentiable in an interval $[x, x+h]$, then there exists some point in this interval, denoted by $x + \lambda h$ for some $\lambda \in [0,1]$, such that

$$\begin{aligned} f(x+h) &= f(x) + hf'(x) + \frac{h}{2} f''(x) + \ldots + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(x) + \\ &+ \ldots + \frac{h^n}{n!} f^{(n)}(x + \lambda x) \end{aligned}$$

Bernoulli distribution

$$
\begin{aligned}
f(x|\theta) &= \theta^x(1-\theta)^{1-x} \qquad x \in 0,1 \\
f(x_1,\ldots,x_n|\theta) &= \theta^{\sum x_i}(1-\theta)^{n-\sum_x i} \qquad x \in 0,1 \\
L(\theta|x) &= \theta^{\sum x_i}(1-\theta)^{n-\sum_x i} \qquad x \in 0,1 \\
log\,(L(\theta|x)) &= \sum_i x_i log\theta + (n - \sum_i x_i) log(1-\theta) \\
\frac{dl(x,\theta)}{d\theta} &= \frac{x}{\theta} - \frac{1-x}{1-\theta} \\
\frac{d^2 l(x,\theta)}{d\theta^2} &= -\frac{x}{\theta^2} - \frac{1-x}{(1-\theta)^2} \\
I(\theta) &= \frac{\theta}{\theta^2} + \frac{1-\theta}{(1-\theta)^2} = \frac{1}{\theta(1-\theta)} \\
I_n(\theta) &= \frac{n}{\theta(1-\theta)}
\end{aligned}
$$