

Asymptotic Distribution of MLE

Maura Mezzetti

Department of Economics and Finance
Università Tor Vergata

Outline

- 1 Prerequisites
- 2 Example Score Function and Fisher Information

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 function

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 function

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):

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

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):

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

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 function

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

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

$$I_n(\theta) = \text{var}(u(\theta))$$

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 function

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

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

$$I_n(\theta) = \text{var}(u(\theta))$$

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 function

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

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

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

$$I_n(\theta) = nI(\theta)$$

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 function

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

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

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

$$I_n(\theta) = nI(\theta)$$

Central Limit Theorem

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

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

converges 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

Taylor's 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

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

Asymptotic Normality of the MLE

Theorem (Asymptotic normality of the MLE)

Let $X_1, \dots, X_n$ be i.i.d. with density $f(x|\theta)$, and let $\hat{\theta}_n$ be the MLE of θ . Under regularity conditions (see Casella & Berger, Sec. 10.1),

$$\sqrt{n} \left(\hat{\theta}_n - \theta \right) \longrightarrow_d N \left(0, \frac{1}{I(\theta)} \right),$$

where $I(\theta)$ is the Fisher information of a single observation.

Hence, for large n ,

$$\hat{\theta}_n \approx N \left(\theta, \frac{1}{I_n(\theta)} \right), \quad \text{with } I_n(\theta) = n I(\theta).$$

Bernoulli distribution

$$f(x|\theta) = \theta^x(1-\theta)^{1-x} \quad x \in \{0, 1\}$$

$$f(x_1, \dots, x_n|\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \quad x \in \{0, 1\}$$

$$L(\theta|x) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \quad x \in \{0, 1\}$$

$$\log(L(\theta|x)) = \sum_i x_i \log \theta + (n - \sum_i x_i) \log(1 - \theta)$$

$$\frac{d \log L(\theta; x)}{d\theta} = \frac{\sum_i x_i}{\theta} - \frac{n - \sum_i x_i}{1 - \theta}$$

$$\frac{d^2 \log L(\theta; x)}{d\theta^2} = -\frac{\sum_i x_i}{\theta^2} - \frac{n - \sum_i x_i}{(1 - \theta)^2}$$

$$l(\theta) = \frac{\theta}{\theta^2} + \frac{(1 - \theta)}{(1 - \theta)^2} = \frac{1}{\theta(1 - \theta)}$$

$$I_n(\theta) = \frac{n}{\theta(1 - \theta)}$$

Bernoulli distribution

Score Function

$$\frac{\sum_i x_i}{\theta} - \frac{n - \sum_i x_i}{1 - \theta}$$

Fisher Information

$$\frac{n}{\theta(1 - \theta)}$$

Bernoulli distribution

Score Function

$$\frac{\sum_i x_i}{\theta} - \frac{n - \sum_i x_i}{1 - \theta}$$

Fisher Information

$$\frac{n}{\theta(1 - \theta)}$$

Asymptotic Distribution

$$\bar{X} \approx N\left(\theta, \frac{\theta(1 - \theta)}{n}\right)$$

Bernoulli distribution

Score Function

$$\frac{\sum_i x_i}{\theta} - \frac{n - \sum_i x_i}{1 - \theta}$$

Fisher Information

$$\frac{n}{\theta(1 - \theta)}$$

Asymptotic Distribution

$$\bar{X} \approx N\left(\theta, \frac{\theta(1 - \theta)}{n}\right)$$

Geometric Distribution: $p(x|\pi) = \pi(1 - \pi)^x$

Score Function

$$\frac{n}{\pi} - \frac{\sum_i x_i}{1 - \pi}$$

Fisher Information

$$\frac{d^2 \log L(\pi|x)}{d\pi^2} = -\frac{n}{\pi^2} - \frac{\sum_i x_i}{(1-\pi)^2}, \text{ Since } E(x) = \frac{1-\pi}{\pi}$$

$$I_n(\pi) = \frac{n}{\pi^2} + \frac{n}{\pi(1-\pi)} = \frac{n}{\pi^2(1-\pi)}$$

Geometric Distribution: $p(x|\pi) = \pi(1 - \pi)^x$

Score Function

$$\frac{n}{\pi} - \frac{\sum_i x_i}{1 - \pi}$$

Fisher Information

$$\frac{d^2 \log L(\pi|x)}{d\pi^2} = -\frac{n}{\pi^2} - \frac{\sum_i x_i}{(1-\pi)^2}, \text{ Since } E(x) = \frac{1-\pi}{\pi}$$

$$I_n(\pi) = \frac{n}{\pi^2} + \frac{n}{\pi(1-\pi)} = \frac{n}{\pi^2(1-\pi)}$$

Asymptotic Distribution

$$\frac{n}{n + \sum X_i} \approx N\left(\pi, \frac{\pi^2(1-\pi)}{n}\right)$$

Geometric Distribution: $p(x|\pi) = \pi(1 - \pi)^x$

Score Function

$$\frac{n}{\pi} - \frac{\sum_i x_i}{1 - \pi}$$

Fisher Information

$$\frac{d^2 \log L(\pi|x)}{d\pi^2} = -\frac{n}{\pi^2} - \frac{\sum_i x_i}{(1-\pi)^2}, \text{ Since } E(x) = \frac{1-\pi}{\pi}$$

$$I_n(\pi) = \frac{n}{\pi^2} + \frac{n}{\pi(1-\pi)} = \frac{n}{\pi^2(1-\pi)}$$

Asymptotic Distribution

$$\frac{n}{n + \sum X_i} \approx N\left(\pi, \frac{\pi^2(1-\pi)}{n}\right)$$

Gaussian Distribution

Score Function

$$\frac{\sum_i (x_i - \mu)}{\sigma^2}$$

Fisher Information

$$\frac{n}{\sigma^2}$$

Gaussian Distribution

Score Function

$$\frac{\sum_i (x_i - \mu)}{\sigma^2}$$

Fisher Information

$$\frac{n}{\sigma^2}$$

Asymptotic Distribution

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Gaussian Distribution

Score Function

$$\frac{\sum_i (x_i - \mu)}{\sigma^2}$$

Fisher Information

$$\frac{n}{\sigma^2}$$

Asymptotic Distribution

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$