Imperial College London

Lecture 07: Maximum Likelihood Estimation (Asymptotic Results)

Statistical Modelling I

Dr. Riccardo Passeggeri

Outline

Introduction

Large sample properties

Let $X_1, X_2, \ldots$ be iid observations with pdf (or pmf) $f_\theta(x)$, where $\theta \in \Theta$ and Θ is an open interval. Let $\theta_0 \in \Theta$ denote the true parameter. Under regularity conditions (e.g. $\{x : f_\theta(x) > 0\}$ does not depend on θ), the following holds:

(i) There exists a **consistent** sequence $(\hat{\theta}_n)_{n\in\mathbb{N}}$ of maximum likelihood estimators. [$\hat{\theta}_n$ is an MLE based on $X_1, \ldots, X_n$].

(ii) Suppose $(\hat{\theta}_n)_{n\in\mathbb{N}}$ is a consistent sequence of MLEs. Then

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, (I_f(\theta_0))^{-1}),$$

where $I_f(\theta) = E_\theta[(\frac{\partial}{\partial\theta} \log f_\theta(X))^2]$ is the **Fisher Information** of a sample of size 1.

Consistency

Sketch of proof (1/3)

There exists a **consistent** sequence $(\hat{\theta}_n)_{n\in\mathbb{N}}$ of maximum likelihood estimators.

Sketch of proof (2/3)

Sketch of proof (3/3)

Asymptotic normality

Sketch of proof (1/4)

Suppose $(\hat{\theta}_n)_{n\in\mathbb{N}}$ is a consistent sequence of MLEs. Then

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, (I_f(\theta_0))^{-1}),$$

where $I_f(\theta) = E_\theta[(\frac{\partial}{\partial\theta} \log f_\theta(X))^2]$ is the **Fisher Information** of a sample of size 1.

Sketch of proof (2/4)

Sketch of proof (3/4)

Sketch of proof (4/4)

Summary

These and similar arguments are frequently used in asymptotic statistics.

Asymptotic normality will be used in the next lecture to derive large sample confidence intervals.