Imperial College London

Lecture 05: Asymptotic Properties II

Statistical Modelling I

Dr. Riccardo Passeggeri

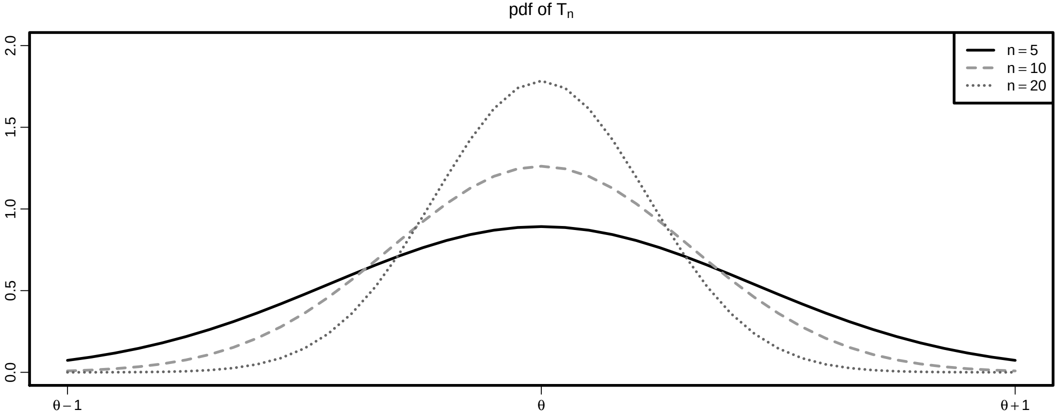
Outline

Introduction

Sampling distributions are important for statistical inference

- Consistency is only a minimal requirement on an estimator T_n
- To derive valid hypothesis tests and confidence intervals we need the sampling distribution of T_n

Example: $Y_1, Y_2, \cdots \sim N(\theta, 1)$



We have shown $T_n \sim N(\theta, \frac{1}{n})$. Centering and scaling, we see that

$$\sqrt{n}(T_n - \theta) \sim N(0, 1)$$

Asymptotically normal estimator

The distribution of many estimators cannot be computed as nicely as in the above example. However, the distribution of many estimators can be approximated by a normal distribution.

Definition

A sequence T_n of estimators for $\theta \in \mathbb{R}$ is called *asymptotically normal* if, for some $\sigma^2(\theta)$

$$\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \sigma^2(\theta))$$

The Central Limit Theorem (CLT)

Theorem (Central Limit Theorem)

Let $Y_1, \ldots, Y_n$ be iid random variables with $E(Y_i) = \mu$ and $Var(Y_i) = \sigma^2 < \infty$. Then the sequence $\sqrt{n}(\bar{Y} - \mu)$ converges in distribution to a $N(0, \sigma^2)$ distribution.

Example: $Y_1, Y_2, \ldots, Y_n$ iid Bernoulli(θ)

Consider the estimator $T_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ for θ

Standard error

Under mild regularity conditions, the standard error of an asymptotically normal estimator T_n can be estimated by $\mathrm{SE}_\theta(T_n) \approx \sigma(T_n)/\sqrt{n}$.

Example: sample mean

Summary

- Asymptotically normal estimators allow us to approximate the sampling distribution of T_n regardless of the data distribution
- The CLT tells us that sample averages are asymptotically normal
- What about other estimators?
- Next: We want to show that estimators besides sample averages are asymptotically normal.

Large sample results

Slutsky's lemma

Theorem (Slutsky's lemma)

Let X_n, X and Y_n be random variables (or vectors). If $X_n \rightarrow_d X$ and $Y_n \rightarrow_p c$ for a constant c, then

(i) $X_n + Y_n \rightarrow_d X + c$;

(ii) $Y_n X_n \rightarrow_d cX$;

(iii) $Y_n^{-1} X_n \rightarrow_d c^{-1} X$ *provided* $c \neq 0$.

Example: $X \sim \text{Binomial}(n, p)$

The estimator sequence $T_n = \frac{X+1}{n+2}$ is asymptotically normal

The delta method

If T_n is asymptotically normal, what about $g(T_n)$?

Theorem (Delta Method)

Suppose that T_n is an asymptotically normal estimator of θ with

$$\sqrt{n}(T_n - \theta) \to_d N(0, \sigma^2(\theta)).$$

Let $g : \Theta \to \mathbb{R}$ be a differentiable function with $g'(\theta) \neq 0$. Then,

$$\sqrt{n}(g(T_n) - g(\theta)) \to_d N(0, g'(\theta)^2 \sigma^2(\theta)).$$

Example: $Y_1, \ldots, Y_n$ iid Bernoulli(p)

The odds of an event A are $P(A)/(1 - P(A))$. Consider estimating the odds that $Y_i = 1$.

Continuous mapping

One last useful result is that stochastic convergence, like convergence in metric spaces is preserved under continuous mappings

Theorem (Continuous mapping theorem)

Let $g : \mathbb{R}^k \to \mathbb{R}^m$ be continuous at every point of a set C such that $P(X \in C) = 1$.

1. *If $X_n \to_d X$, then $g(X_n) \to_d g(X)$*
2. *If $X_n \to_p X$, then $g(X_n) \to_p g(X)$*
3. *If $X_n \to_{as} X$, then $g(X_n) \to_{as} g(X)$*

Summary

The following results are often used to derive the asymptotic distribution of a wide variety of statistics

- The CLT
- Slutsky's lemma
- The delta method
- The continuous mapping theorem