Imperial College
London

# Lecture 05: Asymptotic Properties II

Statistical Modelling I

Dr. Riccardo Passeggeri

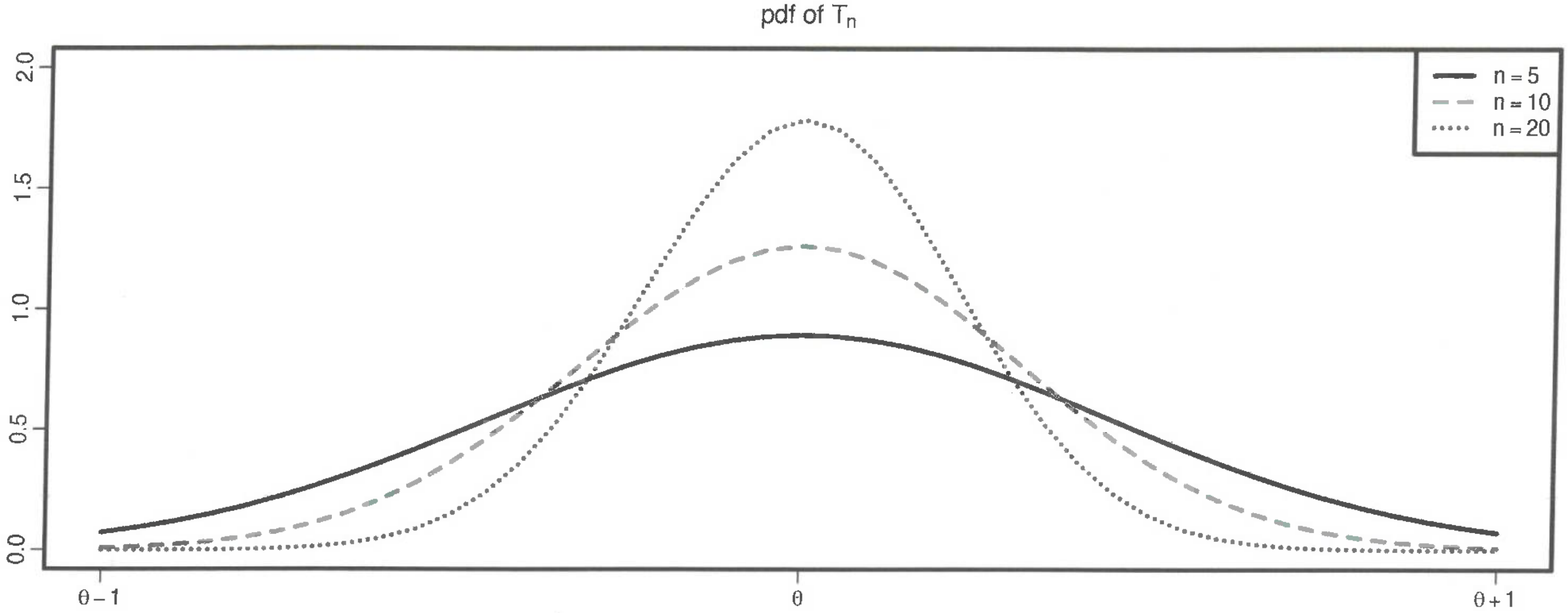
# Outline

# 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

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

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

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

# 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.

[redacted]

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

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

$$\overline{\phantom{xxxxxxxxxxxx}}$$
$$\sqrt{\sigma^2}$$

$$\frac{\sqrt{n}\,(T_n - \theta)}{\sqrt{\sigma^2}} \xrightarrow{d} N(0, 1)$$

# The Central Limit Theorem (CLT)

[redacted]

*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($\theta$)

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

$$\sqrt{n}\,(T_n - \theta) \xrightarrow{d} N(0, \theta(1-\theta)) \quad , \text{as } n \to \infty$$

# 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** AND THE PARAMETER IS THE MEAN

$$SE(T_n) = \frac{\sigma(T_n)}{\sqrt{n}} \iff Var(T_n) = \frac{\sigma^2(T_n)}{n}$$

$$Var(T_n) = E[(T_n-\theta)^2] = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} Y_i - \theta\right)^2\right] =$$

$$= \frac{1}{n^2}\, n\, E[(Y_1-\theta)^2] = \frac{1}{n^2}\, n\, \sigma^2(\theta) = \frac{\sigma^2(\theta)}{n} \approx \frac{\sigma^2(T_n)}{n}$$

$T_n \xrightarrow{P} \theta$ and if $\sigma^2(\cdot)$ IS CONTINUOUS THEN

$\sigma^2(T_n) \xrightarrow{P} \sigma^2(\theta)$ (BY THE CONTINUOUS MAPPING THEOREM)

$$\implies \frac{\sigma^2(\theta)}{n} \approx \frac{\sigma^2(T_n)}{n}$$

## 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

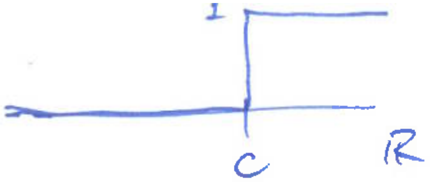
[redacted]

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

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

(ii) $Y_n X_n \to_d cX$;

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




$$\frac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{P} p = E[X_1]$$

## Example: $X \sim \text{Binomial}(n, p)$, $X = \sum_{i=1}^{n} X_i$  $X_i \overset{iid}{\sim} \text{Bernoulli}(p)$

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

$$T_n = \frac{X+1}{n+2} = \frac{n}{n+2}\frac{X}{n} + \frac{1}{n+2} = \frac{n}{n+2}\frac{X}{n} + \frac{X}{n} - \frac{X}{n} + \frac{1}{n+2}$$

$$\sqrt{n}(T_n - p) = \sqrt{n}\left(\frac{n}{n+2}\frac{X}{n} + \frac{X}{n} - \frac{X}{n} + \frac{1}{n+2} - p\right) =$$

$$= \underbrace{\sqrt{n}\left(\frac{X}{n}\left(\frac{n}{n+2} - 1\right)\right)}_{= \frac{X}{n}\sqrt{n}\left(\frac{-2}{n+2}\right) \xrightarrow[\text{a.s.}]{P} 0} + \underbrace{\sqrt{n}\left(\frac{1}{n+2}\right)}_{\xrightarrow[\text{a.s.}]{P} 0} + \underbrace{\sqrt{n}\left(\frac{X}{n} - p\right)}_{= \sqrt{n}\left(\frac{\sum_{i=1}^{n} X_i}{n} - p\right) \xrightarrow{d} N(0, p(1-p))}$$

by SLUTSKY LEMMA

$$\sqrt{n}(T_n - p) \xrightarrow{d} N(0, p(1-p))$$

# The delta method

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

[redacted]

*Suppose that $T_n$ is an asymptotically normal estimator of $\theta$ 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,*

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

$\mathbb{1}$ or $I$ $\quad \mathbb{1}_B(x) = \mathbb{1}_{(x\in B)} = \mathbb{1}_{\{x\in B\}} = \mathbb{1}(x\in B) = \begin{cases} 1 & \text{if } x\in B \\ 0 & \text{if } x\notin B \end{cases}$



# Example: $Y_1, \ldots, Y_n$ iid Bernoulli($p$), $p\in(0,1)$

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

$P(Y_i=1)/(1-P(Y_i=1)) = P(Y_1=1)/(1-P(Y_1=1))$

TO ESTIMATE $P(Y_1=1) = E[\mathbb{1}_{\{Y_1=1\}}]$

$\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{Y_i=1\}} \xrightarrow{P} P(Y_1=1) = p \qquad \mathbb{1}_{\{Y_i=1\}} = Y_i$

$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{Y_i=1\}} - p\right) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i - p\right) \xrightarrow{d} N(0, p(1-p))$

~~(X HERE IS A CONSTANT)~~

$g(x) = \frac{x}{1-x}$ $\quad g$ IS DIFFERENTIABLE IN $(0,1)$ AND $g'(x) = \frac{1}{(1-x)^2} \neq 0 \quad \forall x\in(0,1)$

$\sqrt{n}\left(\frac{\frac{1}{n}\sum_{i=1}^{n} Y_i}{1-\frac{1}{n}\sum_{i=1}^{n} Y_i} - \frac{p}{1-p}\right) \xrightarrow[\text{BY DELTA METHOD}]{d} N\left(0, p(1-p)\left(\frac{1}{(1-p)^2}\right)^2\right) = N\left(0, \frac{p}{(1-p)^3}\right)$

$\to$ CHECKING THE CONDITIONS OF THE DELTA METHOD THEOREM

# Continuous mapping

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

[redacted]

*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