Imperial College London

Lecture 04: Asymptotic Properties I

Statistical Modelling I

Dr. Riccardo Passeggeri

Outline

Introduction

Evaluating estimators

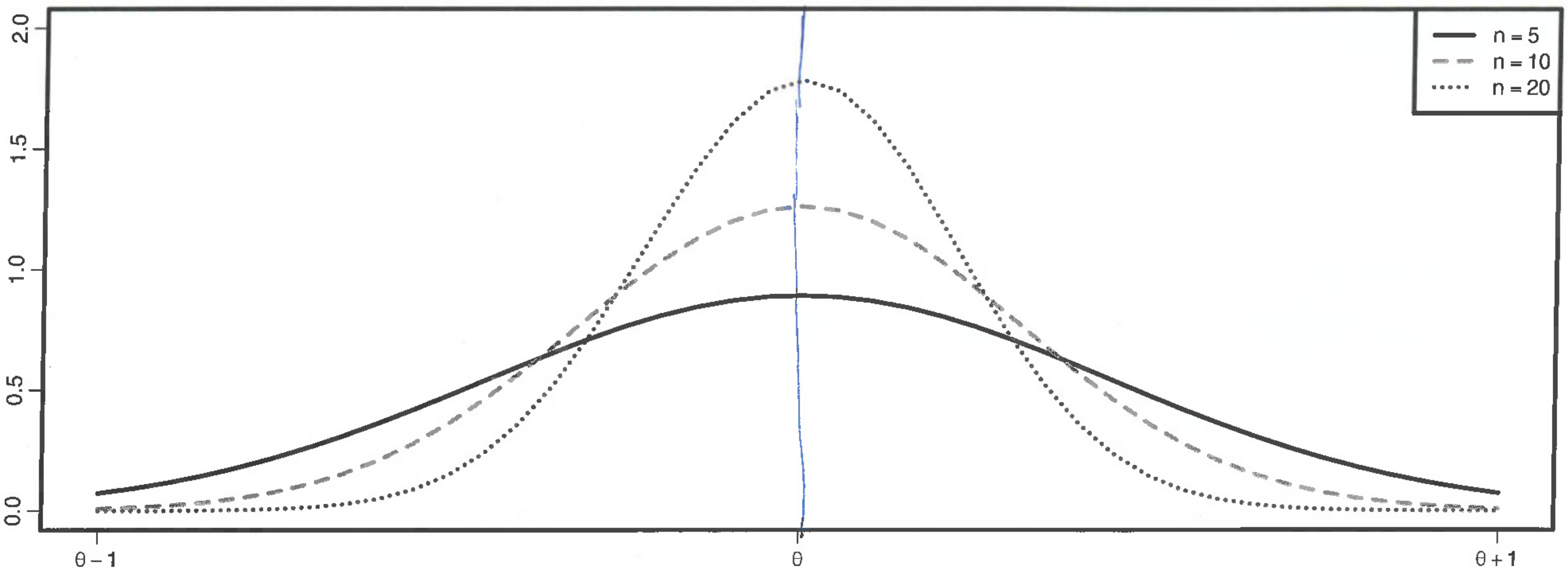
Evaluating an estimator $T = T(X_1, \ldots, X_n)$ of θ depends on its *sampling distribution*

This problem often simplifies by considering $T_n \equiv T_n(X_1, \ldots, X_n)$ to be a sequence of random variables indexed by $n = 1, 2, 3, \ldots$

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

$$T_n = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad E(T_n) = \theta \qquad Var(T_n) = \frac{1}{n} \qquad T_n \sim N\left(\theta, \frac{1}{n}\right)$$

pdf of T_n



Modes of Stochastic Convergence

Stochastic Convergence

Just as sequences of numbers $x_1, x_2, \ldots$ can convergence, sequences of random variables $X_1, X_2, \ldots$ exhibit patterns of convergence

- Convergence almost surely
- Convergence in probability
- Convergence in distribution

Convergence almost surely and convergence in probability

X_n converges almost surely to X if

$(\Omega, \mathcal{F}, \mathbb{P})$
UNDERLYING
PROB. SPACE

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1$$

$$P\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

$$P\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\right\}\right) = 0$$

X_n converges in probability to X if, $\forall \epsilon > 0$,

$|\cdot| \Rightarrow$ THE EUCLIDEAN NORM

$$\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$$

$$\lim_{n\to\infty} P\left(\left\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\right\}\right) = 0$$

Convergence in distribution

X_n converges in distribution to X with cdf F_X if

$$\lim_{n\to\infty} P(X_n \leq x) = F_X(x) = P(X \leq x)$$

at all continuity points x of $F_X(x)$

$$\lim_{n \to \infty} P(\{\omega \in \Omega : X_n(\omega) \leq x\}) = P(\{\omega \in \Omega : X(\omega) \leq x\})$$

Portmanteau lemma: The following statements are equivalent:

- $X_n \xrightarrow{d} X$
- $E[f(X_n)] \to E[f(X)]$ for all bounded and continuous functions $f : \mathbb{R} \to \mathbb{R}$.

Relationship between modes of stochastic convergence

We have

$$X_n \to_{as} X \qquad \Rightarrow \qquad X_n \to_p X \qquad \Rightarrow \qquad X_n \to_d X$$

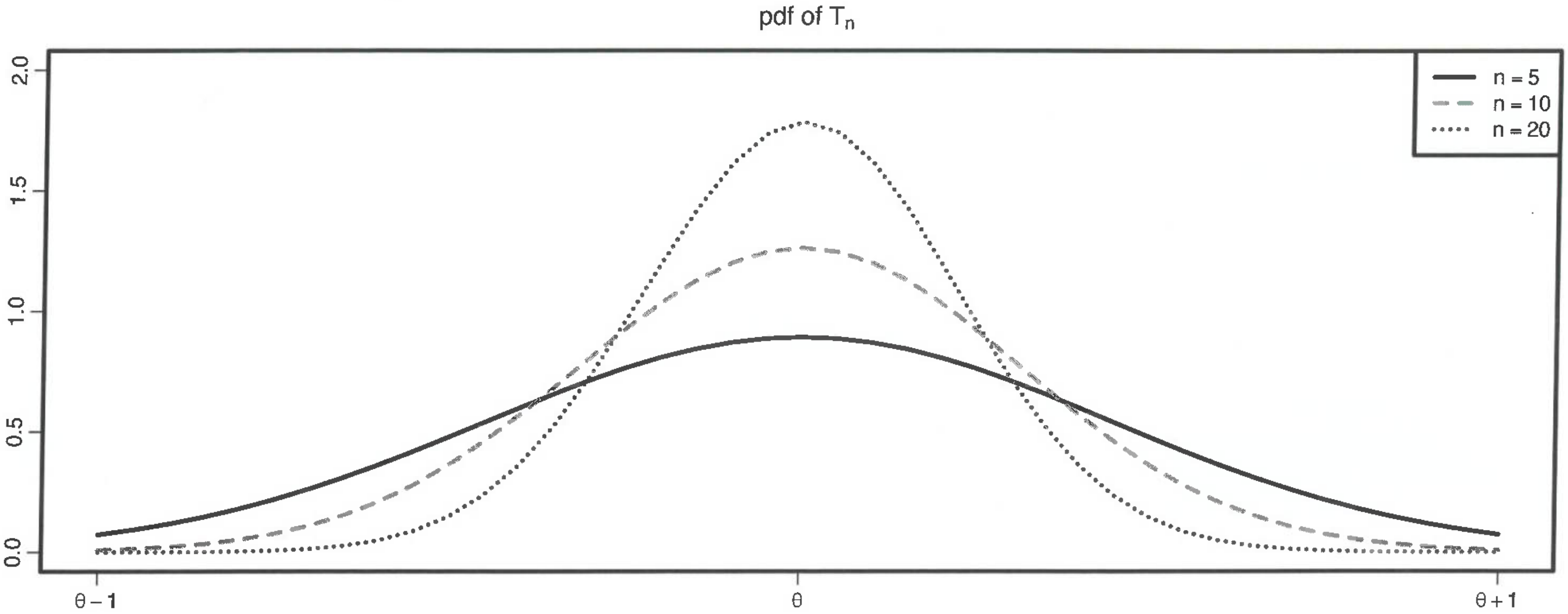
and if $X = c$ is a constant

$$X_n \to_p X \qquad \Leftrightarrow \qquad X_n \to_d X$$

Consistency

Concentration

Estimating θ from $N(\theta, 1)$, we see distribution of T_n concentrates around θ_0



Definition: Consistency

A sequence of estimators $(T_n)_{n\in\mathbb{N}}$ for $g(\theta)$ is called **(weakly) consistent** if for all $\theta \in \Theta$:

→ CONVERGENCE IN PROB.

$$T_n \xrightarrow{P_\theta} g(\theta) \quad (n \to \infty)$$

THANKS TO THE WEAK LAW OF LARGE NUMBERS WE KNOW THAT THE SAMPLE MEAN IS A CONSISTENT ESTIMATOR FOR THE MEAN THAT IS:

$Y_1, \dots, Y_n$ iid

$$T_n = \frac{1}{n} \sum_{i=1}^{n} Y_i \xrightarrow{P_\theta} E[Y_1] = \mu \quad , \forall \mu \in \Theta$$

Definition: Asymptotically Unbiased Estimator

A sequence of estimators $(T_n)_{n\in\mathbb{N}}$ for $g(\theta)$ is called **asymptotically unbiased** if for all $\theta \in \Theta$:

$$E_\theta(T_n) \to g(\theta) \quad (n \to \infty)$$

$Y_1, \dots, Y_n$ iid WITH UNKNOWN MEAN μ

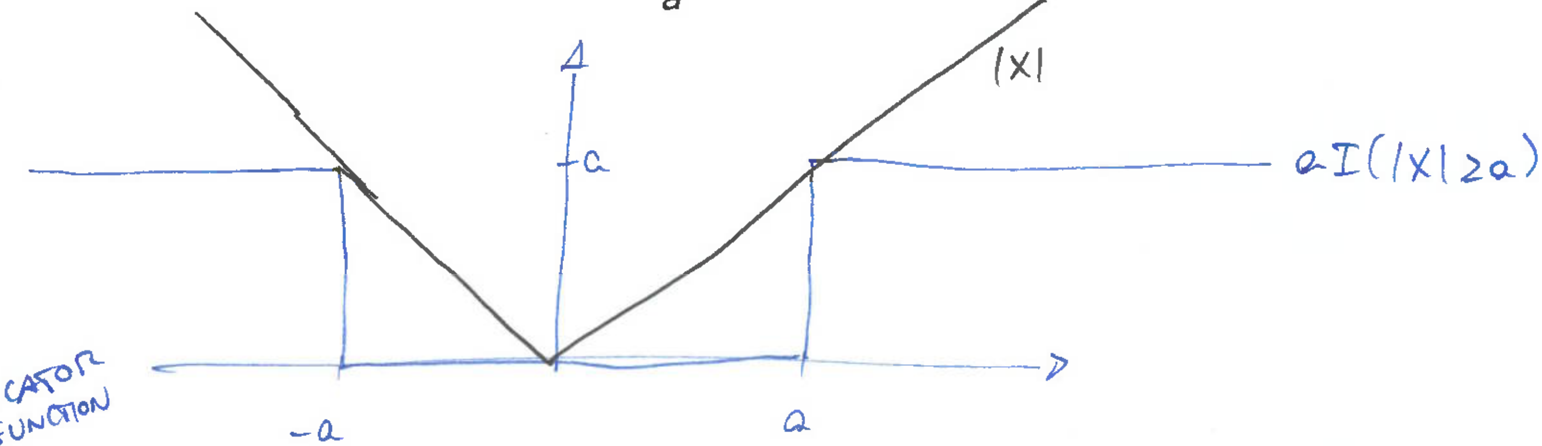
$T_n = \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2$ UNBIASED FOR μ^2? NO! BUT IT IS ASYMPTOTICALLY UNBIASED

$$E[T_n] = E\left[\frac{1}{n^2}\left(\sum Y_i\right)^2\right] = \frac{\sigma^2}{n} + \mu^2 \longrightarrow \mu^2 \quad \text{, AS } n \to \infty$$

(*) ~~SEE~~ SEE THE LAST SLIDE FOR DERIVATION

Markov's Inequality

$$P(|X| \geq a) \leq \frac{E|X|}{a} \qquad \forall a > 0$$



INDICATOR FUNCTION

Proof

$$a\mathcal{I}(|X| \geq a) \leq |X| \qquad \Rightarrow \qquad P(|X| \geq a) = E\{\mathcal{I}(|X| \geq a)\} \leq \frac{1}{a}E|X|$$

$$E[\cancel{a}I(|X| \geq a)] \leq \frac{E[|X|]}{a}$$

$$P(X \in A) = E[I(X \in A)]$$

Lemma: MSE consistency

If T_n is asymptotically unbiased for $g(\theta)$ and $\forall\theta \in \Theta$

$$Var_\theta(T_n) \to 0 \quad (n \to \infty)$$

then T_n is consistent for $g(\theta)$

Proof

Applying Markov's inequality, we have

MARKOV INEQUALITY

$$\forall\epsilon > 0: \quad P_\theta(|T_n - g(\theta)| \geq \epsilon) = P_\theta((T_n - g(\theta))^2 \geq \epsilon^2) \leq \frac{E_\theta(T_n - g(\theta))^2}{\epsilon^2} = \frac{MSE_\theta(T_n)}{\epsilon^2}$$

$$= \frac{1}{\epsilon^2}(\underbrace{Var_\theta T_n}_{\to 0} + \underbrace{(E_\theta T_n - g(\theta))^2}_{\to 0})$$

MSE = VAR + BIAS ^{2}

BY ASSUMPTION OF THE LEMMA

Example: $X_i \sim \text{Bernoulli}(\theta)$, $\theta \in \Theta = [0,1]$, iid

$$T_n(x_1, \dots, x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\left.\begin{aligned} E_\theta T_n(X_1, \dots, X_n) &= \theta \quad \forall \theta \in \Theta \\ Var_\theta(T_n) &= \frac{1}{n}\theta(1-\theta) \to 0 \quad (n \to \infty) \end{aligned}\right\} \implies T_n \text{ consistent}$$

Summary

- Consistency is an important property, but does not tell us all about the sampling distribution of T_n
- To establish approximations to the sampling distribution of T_n, we will appeal to convergence in distribution of sequences of the form $a_n(T_n - g(\theta))$ for deterministic a_n
- We typically find that a_n is proportional to $SE(T_n)$
- Many estimators "look like" sample means, so $a_n = \sqrt{n}$ and $N(0, \sigma^2)$ limits are extremely common!

$$E\left[\frac{1}{n^2}\sum\sum Y_i Y_j\right] = \frac{1}{n^2}\sum_{i\neq j}\mu^2 + \frac{1}{n}E[Y_j^2] =$$

$$= \frac{1}{n^2}\, n(n-1)\,\mu^2 + \frac{1}{n}(\sigma^2+\mu^2)$$

$$= \frac{1}{n}(n-1)\mu^2 + \frac{1}{n}(\sigma^2+\mu^2) = \mu^2 + \frac{\sigma^2}{n}$$

$$n^2 - n = n(n-1)$$