Principles of Statistics

Arthur Conmy*

Part II, Michaelmas Term 2021

These notes are based on lectures given (virtually) by Dr P-L Loh in Michaelmas term 2021. Credit is also due to Evan Chen for the style file for these notes[1].

Contents

§1 Introduction

(skip if not interested in meta notes)

This course follows naturally on from IB Statistics. It also borrows some technology from II Probability and Measure. I haven't currently decided what emphasis to add to these notes: a focus on intuitions or rigour or something else entirely.

§2 Convergence

Since we're in a non-deterministic setting, we need new defintions of convergence. Let $X, X_1, X_2, \ldots : \mathcal{X} \to \mathbb{R}$ be random variables, where $\mathcal{X} \subseteq \mathbb{R}^p$. The first and weakest notion of convergence is *convergence in distribution*. This is 1.

Definition 2.1 (Convergence in distribution)**.** $X_n \xrightarrow{d} X$ if

$$\mathbb{P}[X_n \leq t] \to \mathbb{P}[X \leq t] \tag{1}$$

for all $t \in \mathbb{R}^p$ for which $t \to \mathbb{P}[X \leq t]$ is continuous.

Here, $\leq$ means $x_i \leq y_i$ in all components.

A stronger notion is convergence is convergence in probability.

Definition 2.2 (Convergence in probability)**.** $X_n \xrightarrow{P} X$ if

$$\lim_{n\to\infty} \mathbb{P}[||X_n - X|| > \varepsilon] = 0. \tag{2}$$

*Please send any corrections and/or feedback to asc70@cam.ac.uk

[1]Available here: https://github.com/vEnhance/dotfiles/blob/master/texmf/tex/latex/evan/evan.sty.

The proof that convergence in probability implies convergence in distribution is omitted. Hint: only needing to check around points of continuity allows for a standard analysis argument to be applied. See [1] for the gory details.

Example 2.3 (Convergence in distribution $\not\Rightarrow$ convergence in probability)
Consider $X, X_1, X_2, ...$ independent and identically distributed $\mathcal{N}(0,1)$ random variables. Then all these have identical distribution function, but certainly $X_n \overset{P}{\not\to} X$.

The strongest notion of convergence is convergence *almost surely*. This is closely related to null sets from measure theory, i.e not necessarily empty $A \in \mathfrak{B}$ with $\mu(A) = 0$.

Definition 2.4 (Convergence almost surely)**.** $X_n \overset{a.s}{\to} X$ if

$$\mathbb{P}\left[||X_n - X|| \to 0\right] = 1. \tag{3}$$

Example 2.5 (Convergence in probability $\not\Rightarrow$ convergence in distribution)
Consider independent $X_1, X_2, ...$ with $X_n \sim \text{Ber}(\frac{n-1}{n})$[a], and X the constant random variable 1 (equivalently $X_n \sim \text{Ber}(1)$). Then (check!) $X_n \overset{P}{\to} X$ but $X_n \overset{a.s}{\not\to} X$.

[a]These random variables can be interpreted as the boolean outcomes of whether an ever-improving skeleton archer hits a bullseye on the nth shot at a target.

References

[1] Norris, James, *Probability and Measure*, https://web.archive.org/web/20210507005332/http://www.statslab.cam.ac.uk/~james/Lectures/pm.pdf.