Mathematical Statistics I

Chapter 5: Limit theorems

Jesse Wheeler

Outline

Convergence Concepts

Introduction

- This material comes primarily from Rice (2007, Chapter 5), but will be supplemented with material from Casella and Berger (2024, Chapter 5).
- In this chapter, we are interested in the convergence of sequences of random variables.
- In particular, we are interested in the convergence of the sample mean, $\bar{X}_n = (X_1 + X_2 + \ldots + X_n)/n$, as the number of samples n grows.
- Because $\bar{X}_n$ is itself a random variable, we have to carefully define what it means for the convergence of a random variable.
- In this class, we are mainly concerned with three types of convergence.

Introduction II

- Because convergence of random variables is a tricky topic, we will treat them in varying amounts of detail.

- The first type of convergence is one of the weaker types, and is usually easy(ish) to verify.

Definition: Convergence in Probability

A sequence of random variables $X_1, X_2, \ldots$ converges in probability to a random variable X if, for every $\epsilon > 0$,

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

or, equivalently,

$$\lim_{n\to\infty} P\big(|X_n - X| < \epsilon\big) = 1.$$

Convergence in Probability II

- We often use the shorthand $X_n \xrightarrow{P} X$ to denote "X_n converges in probability to X as n goes to infinity".
- Note that the X_i in the definition above do *not* need to be independent and identically distributed.
- The distribution of X_n changes as the subscript changes, and each of the convergence concepts we will discuss will describe different ways in which the distribution of X_n converges to some limiting distribution as the subscript becomes large.
- A special case is when the limiting random variable X is a constant.

Convergence in Probability III

Example: The (Weak) Law of Large Numbers

Let $X_1, X_2, \ldots$ be iid random variables with $E[X_i] = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$. Define $\bar{X}_n = (1/n)\sum_{i=1}^{n} X_i$. Then $\bar{X}_n \xrightarrow{P} \mu$.

Proof.

Convergence in Probability IV

- The WLLN is very elegant; under general conditions, the sample mean of independent random variables approaches the population mean as $n \to \infty$.
- This is also used for proportions, as proportions are just means of indicator random variables.
- The WLLN can also be extended to show that the results hold even if the variance is infinite, the only condition needed is that the expectation is finite. However, the proof in this case is beyond the scope of this course.
- When a sequence of the "same" sample quantity approaches a constant, we say that the sample quantity is *consistent*.

Convergence in Probability V

- A natural extension of the definition of the convergence of probability, is convergence of functions of random variables: $h(X_1), h(X_2), \ldots$.

Theorem: Convergence in probability for continuous functions

Let $X_1, X_2, \ldots$ be a sequence of random variables that converges in probability to a random variable X, and let h be a continuous function.

Then, $h(X_1), h(X_2), \ldots$ converges in probability to $h(X)$.

- Our next convergence concept is stronger than convergence in probability.

Definition: Almost Sure Convergence

A sequence of random variables $X_1, X_2, \ldots$ converge almost surely to a random variable X if, for every $\epsilon > 0$,

$$P\big(\lim_{n\to\infty} |X_n - X| < \epsilon\big) = 1,$$

or

$$P\big(\lim_{n\to\infty} X_n = X\big) = 1.$$

- Almost sure convergence is often written as $X_n \overset{a.s.}{\to} X$.

Almost sure convergence II

- It appears similar to convergence in probability, but they are in fact very different. In particular, almost sure convergence is a stronger concept.
- One way to think about this difference is that the probability gives a weight to individual sets.
- For convergence in probability, the set where $|X_n - X| > \epsilon$ can have positive probability, but that probability converges to zero for large n.
- For almost sure convergence, the set where $|X_n - X| > \epsilon$ has probability zero. This doesn't imply that the set $|X_n - X| > \epsilon$ is empty, but it has zero probability.

Almost sure convergence III

- Almost sure convergence is very similar to pointwise convergence of a sequence of functions. This is no accident, as random variables *are* functions:

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

- In the equivalent definition above, we see we must have point-wise convergence almost-everywhere, except for the possibility that for some set $N \subset \Omega$ such that $P(N) = 0$, we allow $s \in N$ to not converge: $\lim_{n\to\infty} X_n(s) \neq X(s)$.

Almost sure convergence IV

Example: Convergence in prob, not a.s.

Let the sample space $\Omega = [0,1]$, and assign the uniform probability on this interval. Define the sequence of random variables X_i as: $X_1(s) = s + 1_{[0,1]}(s)$, $X_2(s) = s + 1_{[0,\frac{1}{2}]}(s)$, $X_3(s) = s + 1_{[\frac{1}{2},1]}(s), X_4(s) = s + 1_{[0,\frac{1}{3}]}(s)$, $X_5(s) = s + 1_{[\frac{1}{3},\frac{2}{3}]}(s)$, $X_6(s) = s + 1_{[\frac{2}{3},1]}(s), \ldots$, and then define $X(s) = s$. We can see that $X_n \xrightarrow{P} X$. However, X_n does not converge almost surely, because there is no values $s \in \Omega$ that satisfy $X_n(s) \rightarrow X(s)$. For every ω, the value of $X_n(s)$ alternates between s and $s+1$ infinitely often.

Almost sure convergence V

Theorem: almost sure convergence implies convergence in probability

If $X_1, X_2, \ldots$ are a sequence of random variables such that $X_n \stackrel{a.s.}{\to} X$, for some random variable X, then $X_n \stackrel{P}{\to} X$.

- The converse of the statement above is false. That is, convergence in probability does not imply almost sure convergence.
- A proof of the theorem above, as well as additional treatment of the connection between almost sure convergence and convergence in probability is found in Resnick (2019, Chatper 6).

Almost sure convergence VI

- Note: As stated, the weak-law of large numbers (WLLN) can actually be shown to hold a.s., in which case we call it the strong-law of large numbers (SLLN).

Convergence in Distribution

- The final form of convergence we will consider in this course is convergence in distribution.

Definition: Convergence in Distribution

A sequence of random variables $X_1, X_2, \ldots$ converges in distribution to a random variable X if

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

at all points x where $F_X(x)$ is continuous.

- One way to think about convergence in distribution is that it's really a statement about the long-run behavior of a sequence of random variables, as it's a statement about the CDFs.

Convergence in Distribution II

- This is different from the other types of convergence, which are concerned with the random variable itself.
- A quick recap of how the different types of convergence are related:
  - a.s. convergence $\implies$ convergence in prob $\implies$ convergence in Distribution.
- In a *few* special scenarios, we can talk about more connections between the types of convergence.
- One such example is convergence in probability to a constant. Casella and Berger (Theorem 5.5.13 of 2024) shows that $X_n \xrightarrow{P} a$ for some constant a if and only if $X_n \xrightarrow{d} a$.

The Central Limit Theorem

- TODO: Add continuity theorem (Rice CH 5.3)
- Add some primer about the CLT
- Prove the CLT.

Test Section

Test Frame

References and Acknowledgements

Casella G, Berger R (2024). *Statistical inference*. Chapman and Hall/CRC.

Resnick S (2019). *A probability path*. Springer.

Rice JA (2007). *Mathematical statistics and data analysis*, volume 371. 3 edition. Thomson/Brooks/Cole Belmont, CA.

- Compiled on October 20, 2025 using R version 4.5.1.
-

References and Acknowledgements II

- We acknowledge students and instructors for previous versions of this course / slides.