Modes of Convergence

Will Perkins

February 7, 2013

Limit Theorems

We are often interested in statements involving the limits of random variables. We want to say things like:

$$\lim_{n\to\infty} X_n = X$$

where the X_n's and X are random variables.

But what does this actually mean?

We'll see 4 different types of convergence:

1. Convergence in Distribution (or 'Weak' convergence)
2. Convergence in Probability (or convergence in measure)
3. Almost sure convergence
4. Convergence in Mean (or l_p convergence)

Convergence in Distribution

Let's start with a very simple example, not about convergence in distribution, but about equality in distribution.

Let X be the number of heads in 10 flips of a fair coin and Y the number of tails.
Does $X = Y$? No, of course not. $X = 10 - Y$ and with significant probability they are not equal. But as random variables on their own, they are very similar, and in fact

$$F_X(t) = F_Y(t)$$

for all t. I.e., their *distributions* are the same.

We can say

$$X \stackrel{D}{=} Y$$

Convergence in Distribution

Notice that equality in distribution is just determined by the marginal distributions of two random variables - it doesn't say anything about their joint distribution or that they are even defined on the same probability space!
This is important to keep in mind.

Convergence in Distribution

Definition

We say a sequence of random variables X_n *converges in distribution* to a random variable X if

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

for every $t \in \mathbb{R}$ at which $F_X(t)$ is continuous.

We sometimes write a double arrow to indicate convergence in distribution:

$$X_n \Rightarrow X$$

Convergence in Distribution

A basic example:
Let X be any random variable and let $X_n = X + 1/n$. Then

$$F_{X_n}(t) = \Pr[X_n \leq t] = \Pr[X \leq t - 1/n] = F_X(t - 1/n)$$

And $\lim_{n\to\infty} F_X(t - 1/n) = F_X(t)$ but only at continuity points of F_X, since the function is right continuous. This example shows why we only require convergence at continuity points.

Convergence in Distribution

Example: Let $X_n \sim Bin(n, \lambda/n)$. Let $Y \sim Pois(\lambda)$. Show that

$$X_n \Rightarrow Y$$

[hint: enough to show that $\Pr[X_n = k] \to \Pr[Y = k]$ for all k. Why?]

Convergence in Distribution

Why is it also called Weak Convergence?

Theorem

X_n converges to X in distribution if and only if

$$\lim_{n\to\infty} \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)]$$

for every bounded continuous function $g(x)$.

Q: Does this mean that $\mathbb{E}X_n \to \mathbb{E}X$?
No! $f(x) = x$ is not bounded. Give a counterexample to show that this is not necessarily true.

Weak Convergence

We can write the previous statement using the distributions of our random variables:

$$\lim_{n\to\infty} \int_{\mathbb{R}} g(x)\, d\mu_{X_n}(x) = \int_{\mathbb{R}} g(x)\, d\mu_X(x)$$

for all bounded, continuous $g(x)$.

In functional analysis, this is the definition of the weak convergence of measures. (The convergence of linear functionals on the space of probability measures).

Convergence

All other modes of convergence depend how the sequence of random variables and the limiting random variable are defined together on the same probability space.

Convergence in Probability

Definition

X_n converges *in probability* to X if for every $\epsilon > 0$,

$$\lim_{n \to \infty} \Pr[|X_n - X| > \epsilon] = 0$$

We've seen this type of convergence before: in the proof of the weak law of large numbers.

In other areas of math, this type of convergence is called *convergence in measure*.

Convergence in Probability

Lemma

If X_n converges in distribution to a constant c, then X_n converges in probability to c.

Proof: in the HW.

Convergence in Probability

An example: Let $U \sim Unif[0, 1]$. Let $U_n \sim \frac{1}{n} Bin(n, U)$. Then

$$U_n \xrightarrow{p} U$$

But if $V \sim Unif[0, 1]$ is independent of U, then U_n converges to V in distribution but not in probability. Q: How do we prove that U_n does not converge to V in probability?

Convergence in Probability

The first- and second-moment methods are two ways we know of proving convergence in probability.

Almost Sure Convergence

Definition

X_n converges *almost surely* (or a.s. or a.e.) to X if

$$\Pr[\lim_{n \to \infty} X_n = X] = 1$$

A short way to remember the difference is that convergence in probability talks about the limit of a probability, while almost sure convergence talks about the probability of a limit.

Almost Sure Convergence

Almost sure convergence says that with probability 1, the infinite sequence $X_1(\omega), X_2(\omega), \ldots$ has a limit, and that the limit is $X(\omega)$.

In other words, with probability 1, for every $\epsilon > 0$, $|X_n - X| > \epsilon$ only finitely many times.

Almost Sure Convergence

An example: $U \sim Uniform[0,1]$. $X_n = 1/n$ if $U \leq 1/2$, $X_n = 1 - 1/n$ if $U > 1/2$. Show that X_n converges almost surely. To what does X_n converge?
Notice X_n's are very dependent.

Example 2: Consider an infinite sequence of fair coin flips. Let H_n be the indicator rv that you've seen at least one head by flip n. Show that H_n converges a.s.

Borel-Cantelli Lemma

Theorem

Let $A_1, A_2, \ldots$ be an infinite sequence of events. Then

1. *If $\sum_{i=1}^{\infty} \Pr(A_i) < \infty$, with probability 1 only finitely many A_i's occur.*
2. *If $\sum_{i=1}^{\infty} \Pr(A_i) = \infty$ and the A_i's are independent, then with probability 1 infinitely many A_i's occur.*

Proofs:

1. Linearity of Expectation (and Fubini's Theorem)
2. Basic properties of probability and the inequality $1 - p \leq e^{-p}$

l_p Convergence

Definition

We say X_n converges in l_p to X if

$$\lim_{n\to\infty} ||X_n - X||_p = 0$$

where $||f||_p = \left(\int |f|^p\right)^{1/p}$

$||X||_2$ is the usual Euclidean length. We will primarily be interested in $p = 1$ and $p = 2$.

l_p Convergence

A weak law for l_2 convergence:
Let $X_1, X_2, \ldots$ be iid with mean μ and variance σ^2. Prove that

$$\frac{X_1 + \cdots + X_n}{n} \to \mu$$

in l_2.

l_p Convergence

Show that if $X_n \to X$ in l_1, then $\mathbb{E}X_n \to \mathbb{E}X$.
Show that the converse is false.
What if $X_n \to X$ in l_2?

Implications

1. Convergence in distribution is the weakest type of convergence. All other types imply convergence in distribution.
2. Almost sure convergence implies convergence in probability.
3. l_p convergence implies convergence in probability

None of the other directions hold in general.

Counterexamples

We need examples of the following:

1. Convergence in probability but not almost surely.
2. Convergence in probability but not in l_p
3. Convergence in l_p but not almost surely.
4. Convergence almost surely but not in l_p