Three Modes of Convergence

Jin-Lung Lin

1. Converge in distribution
2. Converge in probability
3. Converge with probability 1

1 Definitions

1.1 Converge in distribution

Suppose that $F_1, F_2, \cdots$ is a sequence of cumulative distribution functions corresponding to random variables $X_1, X_2, \cdots$, and that F is a distribution function corresponding to a random variable X. We say that the sequence X_n converges towards X in distribution, if

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for every real number x at which F is continuous. Since $F(x) = Pr(X \le x)$, this means that the probability that the value of X is in a given range is very similar to the probability that the value of X_n is in that range, provided n is large enough. Convergence in distribution is often denoted by adding the letter D over an arrow indicating convergence:

$$X_n \stackrel{D}{\to} X$$

1.2 Convergence in probability

The sequence X_n is said to converge towards X in probability if

$$\lim_{n\to\infty} \Pr\left(|X_n - X| \ge \varepsilon\right) = 0$$

for every $\varepsilon > 0$. Convergence in probability is, indeed, the (pointwise) convergence of probabilities. Pick any $\varepsilon > 0$ and any $\delta \ge 0$, Let P_n be the probability that X_n is outside a tolerance ε of X. Then, if X_n converges in probability to X then there exists a value N such that, for all $n > N$, P_n is itself less than δ.

Convergence in probability is often denoted by adding the letter 'P' over an arrow indicating convergence:

$$X_n \stackrel{P}{\to} X$$

Equivalently, $X_n \stackrel{P}{\to} X$ if for any $\varepsilon > 0, \delta > 0, \exists N$ such that $n > N \Rightarrow P(|X_n - X| < \varepsilon) > 1 - \delta$

1.3 Converge with probability 1

To say that the sequence X_n converges almost surely or almost everywhere or with probability 1 or strongly towards X means

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

This means that the values of X_n approach the value of X, in the sense (see almost surely) that events for which X_n does not converge to X have probability 0. Using the probability space (Ω, F, P) and the concept of the random variable as a function from Ω, to R, this is equivalent to the statement

$$\Pr\left(\left\{\omega \in \Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

and is denoted as

$$X_n \overset{wp1}{\to} X$$

Equivalent definition 1, $X_n \overset{wp1}{\to} X$ if for any $\varepsilon > 0, \lim_{n\to\infty} P(|X_m - X| < \varepsilon,\ \text{all } m \geq n) = 1$
Equivalent definition 2: $X_n \overset{wp1}{\to} X$ if for any $\varepsilon > 0, \delta > 0, \exists\ N$ and all $r > 0$ such that $n > N \Rightarrow P(|X_n - X| < \varepsilon, |X_{n+1} - X| < \varepsilon, \cdots, |X_{n+r} - X| < \varepsilon) > 1 - \delta$

2 Relationship between three modes of convergence

2.1 Implications

1. $X_n \overset{wp1}{\to} X \Rightarrow X_n \overset{P}{\to} X$
2. $X_n \overset{P}{\to} X \Rightarrow X_n \overset{D}{\to} X$

Convergence in probability sufficiently fast implies convergence wp 1.
Theorem: If $\sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < \infty\ \forall \varepsilon > 0$, then $X_n \overset{wp1}{\to} X$
Theorem: *If* $X_n \overset{P}{\to} X$ then there exist a subsequence X_{n_i} of X_n ($n_i, i = 1, 2, \cdots$, is a subset of $n = 1, 2, \cdots$, say $1, 2, 4, 9, \cdots, n^2, \cdots$) such that $X_{n_i} \overset{wp1}{\to} X$
Theorem: Let $X_1, X_2, \cdots$ be independent, then if $S_n = X_1 + X_2 + \cdots + X_n \overset{P}{\to} X \Rightarrow S_n \overset{wp1}{\to} X$
Convergence in probability to a constant implies convergence in distribution
Theorem: If $X_n \overset{D}{\to} c$ where c is a constant, then $X_n \overset{P}{\to} c$.

2.2 counter examples

1. Let X_n be independent with $P(X_n = 1) = 1/n, P(X_n 0) = 1 - 1/n$. Then, obviously, $X_n \overset{P}{\to} 0$ but $X_n \overset{wp1}{\to} 0$ does not hold. Actually, X_n does not converge with probability 1.

2. Let X, Y be independent with $P(X = 0) = P(X = 1) = P(Y = 0) = P(Y = 1) = 1/2$,. Let $X_n = Y$, then $X_n \overset{D}{\to} X$ but $X_n \overset{P}{\to} X$ does not hold since $P(|X - Y| = 1) = 1/2$.

3 Law of large numbers

3.1 Weak law of large number

If $X_1, X_2, X_3, \cdots$ is an infinite sequence of random variables, where all the random variables have the same expected value μ and variance σ^2 and are uncorrelated (i.e., the correlation between any two of them is zero), then the sample average

$$\overline{X}_n = (X_1 + \cdots + X_n)/n$$

converges in probability to μ.

3.2 Strong law of large number

If $X_1, X_2, X_3, \cdots$ is an infinite sequence of random variables that are pairwise independent and identically distributed with

$$E(X_i) = \mu \quad \text{and} \quad E(|X_i|) < \infty,$$

then

$$P\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1,$$