M50010 Revision

Teddy Wu

February 2026

1 Convergence of random variables

Definition 1.1. A sequence of random variables X_n with CDF F_n is said to **converge in distribution** to a random variable X with CDF F_X if

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

for all $x \in \mathbb{R}$ where F_X is continuous. Denote as

$$X_n \xrightarrow{d} X.$$

Definition 1.2. A sequence of random variables X_n is said to **converge in probability** to a random variable X if for any $\epsilon > 0$

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

Denote as

$$X_n \xrightarrow{P} X.$$

Definition 1.3. A sequence of random variables X_n is said to **converge almost surely** to a random variable X if

$$\Pr(X_n \to X) = \Pr(\{\omega \in \Omega : X_n(\omega) \to X(\omega)\}) = 1$$

Denote as

$$X_n \xrightarrow{a.s.} X.$$

Remark 1.4. Convergence in distribution only concerns the shape of the distributions, thus with some abuse of notation we can write

$$X_n \xrightarrow{d} N(0,1)$$

meaning $X_n \xrightarrow{d} X$ for some $X \sim N(0,1)$.

Convergence in probability is stronger, in the sense that if $X_n \xrightarrow{p} X$, we know how X_n and X are related. Specifically, a random sample using X_n will become more and more similar to a random sample using X. Note that now it does not make sense to write $X_n \xrightarrow{P} N(0,1)$: Only knowing the shape of the distribution is not enough to conclude convergence in prob.

Almost sure convergence is the strongest among all three. Its definition involves the sample space, enforcing that for almost every sample, the behavior of X_n will tend to that of X. (Here, behavior simply means $X(\omega)$.) In contrast, convergence in probability states that the overall behavior of the samples in X_n tends to that in X, but for each n, we may have different samples whose performance in X_n deviates from that in X.

We can rewrite the def of convergence in probability to compare the difference:

$$\lim_{n\to\infty} \Pr(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}) = 0,$$

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

Theorem 1.5. *Convergence almost surely implies convergence in probability, which implies convergence in distribution.*

Proof. Suppose $X_n \xrightarrow{a.s.} X$. Define the increasing sequence of sets $(A_N(\epsilon))_{N\in\mathbb{N}}$ by

$$A_N(\epsilon) = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \epsilon \textit{ for all } n \geq N\}.$$

For any $\omega \in \{X_n \to X\}$ we have $X_n(\omega) \to X(\omega)$, by the def of limit, $\omega \in A_N(\epsilon)$ for some N, thus

$$\{X_n \to X\} \subseteq \bigcup_{N=1}^{\infty} A_N(\epsilon)$$

and

$$1 = \Pr(X_n \to X) \leq \Pr\left(\bigcup_{N=1}^{\infty} A_N(\epsilon)\right) \leq 1.$$

Therefore $\Pr\left(\bigcup_{N=1}^{\infty} A_N(\epsilon)\right) = 1$. Note that $(A_N(\epsilon))_{N\in\mathbb{N}}$ is nested, so by right continuity we have

$$\lim_{N\to\infty} \Pr(A_N(\epsilon)) = \Pr\left(\bigcup_{N=1}^{\infty} A_N(\epsilon)\right) = 1.$$

By noting that $A_N(\epsilon) \subseteq \{\omega \in \Omega : |X_N(\omega) - X(\omega)| < \epsilon\}$ we have that

$$\Pr(A_N(\epsilon)) \leq \Pr(\{\omega \in \Omega : |X_N(\omega) - X(\omega)| < \epsilon\}) \text{ for all } N \in \mathbb{N}.$$

Finally,

$$\lim_{N\to\infty} \Pr(\{\omega \in \Omega : |X_N(\omega) - X(\omega)| < \epsilon\}) \geq \lim_{N\to\infty} \Pr(A_N(\epsilon)) = 1$$

which shows $X_n \xrightarrow{P} X$.

Now suppose $X_n \xrightarrow{P} X$. Let x be any continuity point of F and let $\epsilon > 0$.

Observe that if $X_n \le x$, then either $X \le x + \epsilon$ or $|X_n - X| > \epsilon$, so

$$\{X_n \le x\} \subseteq \{X \le x + \epsilon\} \cup \{|X_n - X| > \epsilon\}.$$

By a union bound:

$$F_n(x) = \Pr(X_n \le x) \le \Pr(X \le x + \epsilon) + \Pr(|X_n - X| > \epsilon)$$

Similarly, if $X \le x - \epsilon$ then either $X_n \le x$ or $|X_n - X| > \epsilon$:

$$\Pr(X \le x - \epsilon) \le \Pr(X_n \le x) + \Pr(|X_n - X| > \epsilon).$$

Combining the two inequalities we get a two-sided bound for $F_n(x)$:

$$\Pr(X \le x - \epsilon) - \Pr(|X_n - X| > \epsilon) \le \Pr(X_n \le x) \le \Pr(X \le x + \epsilon) + \Pr(|X_n - X| > \epsilon)$$

as $n \to \infty$, $\Pr(|X_n - X| > \epsilon) \to 0$ by convergence in probability, so

$$F(x - \epsilon) \le \lim_{n\to\infty} \Pr(X_n \le x) \le F(x + \epsilon),$$

$$F(x - \epsilon) \le \lim_{n\to\infty} F_n(x) \le F(x + \epsilon).$$

Let $\epsilon \to 0$. By continuity at x we have

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

thus $F_n(x) \to F(x)$.

Lemma 1.6. *If the pdf of X is symmetric around $x = 0$, then $-X$ and X are identically distributed.*

Proof. Let f_X denote the pdf of X, then $f_X(x) = f_X(-x)$.

$$\begin{aligned} F_{-X}(x) &= \Pr(-X \le x) = 1 - \Pr(X < -x) = 1 - \int_{-\infty}^{-x} f_X(t)dt \overset{u=-t}{=} 1 + \int_{\infty}^{x} f_X(-u)du \\ &= 1 - \int_{x}^{\infty} f_X(u)du = \int_{-\infty}^{x} f_X(u)du = F_X(x). \end{aligned}$$

Remark 1.7. Let $U \sim \text{Unif}(-1, 1)$, and for $n \ge 1$ define $U_n = -U$. By 1.6, U_n and U have the same distribution, so trivially $U_n \xrightarrow{d} U$. But

$$\Pr(|U_n - U| \ge 1/2) = \Pr(|2U| \ge 1/2) = 3/4,$$

thus U_n does not converge to U in probability. This happens because convergence in distribution ignores the relation between U_n and U.

On the same probability space define independent variables X_n such that

$$\Pr(X_n = 1) = 1/n, \quad \Pr(X_n = 0) = 1 - 1/n.$$

We have convergence in probability to 0 because

$$\lim_{n\to\infty} \Pr(|X_n - 0| \geq \epsilon) = \lim_{n\to\infty} \Pr(X_n = 1) = \lim_{n\to\infty} 1/n = 0$$

for all small positive ϵ. However, since

$$\sum_{n=1}^{\infty} \Pr(X_n = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

and X_n independent, by second Borel-Cantelli lemma:

$$\Pr(X_n = 1 \ \ i.o.) = 1.$$

For $\omega \in \{X_n = 1 \ \ i.o.\}$, it follows that $X_n(\omega) \not\to 0$. (If otherwise, then only finitely many $X_n(\omega)$ can be 0.) Hence

$$\Pr(\{\omega \in \Omega : X_n(\omega) \to 0\}) \leq \Pr(\{X_n = 1 \ \ i.o.\}^c) = 1 - \Pr(\{X_n = 1 \ \ i.o.\}) = 0,$$

so almost sure convergence does not hold.

Theorem 1.8. *Suppose* $(X_n) \xrightarrow{d} c$ *for some* $c \in \mathbb{R}$*, then* $X_n \xrightarrow{P} c$*.*

Proof. Here we have to view c as a random variable. It has CDF

$$F(x) = \begin{cases} 0 & x < c \\ 1 & x \geq c \end{cases}$$

Since F is continuous at any point except c, we see that for any $\epsilon > 0$,

$$\lim_{n\to\infty} F_n(c - \epsilon) = 0, \quad \lim_{n\to\infty} F_n(c + \epsilon) = 1$$

by convergence in distribution. Now

$$\begin{aligned} \Pr(|X_n - c| \geq \epsilon) &= \Pr(X_n \leq c - \epsilon) + \Pr(X_n \geq c + \epsilon) \\ &\leq \Pr(X_n \leq c - \epsilon) + \Pr(X_n > c + \epsilon/2) \\ &= \Pr(X_n \leq c - \epsilon) + 1 - \Pr(X_n \leq c + \epsilon/2) \\ &= F_n(c - \epsilon) + 1 - F_n(c + \epsilon/2). \end{aligned}$$

Finally when $n \to \infty$ we have

$$0 \leq \lim_{n\to\infty} \Pr(|X_n - c| \geq \epsilon) \leq \lim_{n\to\infty} [F_n(c - \epsilon) + 1 - F_n(c + \epsilon/2)] = 0.$$

so we get convergence in probability.

Theorem 1.9. *Suppose* $X_n \xrightarrow{P} X$ *and* $Y_n \xrightarrow{P} Y$*, then* $X_n + Y_n \xrightarrow{P} X + Y$ *and* $X_n Y_n \xrightarrow{P} XY$*.*

Proof. We know for any $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr(|X_n - X| \geq \epsilon) = 0, \quad \lim_{n\to\infty} \Pr(|Y_n - Y| \geq \epsilon) = 0.$$

By triangle inequality we have

$$\{|X| < \epsilon/2, \ |Y| < \epsilon/2\} \subseteq \{|X + Y| < \epsilon\}.$$

Taking complement yields

$$\{|X + Y| \geq \epsilon\} \subseteq \{|X| \geq \epsilon/2 \ \textit{or} \ |Y| \geq \epsilon/2\}.$$

Thus

$$\begin{aligned} \Pr(|X + Y| \geq \epsilon) &\leq \Pr(|X| \geq \epsilon/2 \ \textit{or} \ |Y| \geq \epsilon/2) \\ &\leq \Pr(|X| \geq \epsilon/2) + \Pr(|Y| \geq \epsilon/2) \end{aligned}$$