1. Convergence of random variables

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables, and X be another random variable defined on the same probability space. We will discuss only real-valued random variables.

Convergence in quadratic mean (q.m.): Suppose that $\{X_n\}_{n\geq 1}$ and X are such that $\mathbf{E}[X_n^2] < \infty$ for all $n \geq 1$ and $\mathbf{E}[X^2] < \infty$. Then, the sequence $\{X_n\}_{n\geq 1}$ converges to X in q.m. if

$$\lim_{n\to\infty} \mathbf{E}\left[(X_n - X)^2\right] = 0.$$

This is denoted by $X_n \overset{q.m.}{\to} X$ (or, $X_n \overset{L^2}{\to} X$).
By definition, $X_n \overset{q.m.}{\to} X$ if and only if $X_n - X \overset{q.m.}{\to} 0$ as $n \to \infty$.

Convergence in probability (P): The sequence $\{X_n\}_{n\geq 1}$ converges to X in probability, denoted by $X_n \overset{P}{\to} X$ if

$$\lim_{n\to\infty} \mathbf{P}\left\{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\right\} = 0$$

for every $\varepsilon > 0$.
By definition, $X_n \overset{P}{\to} X$ if and only if $X_n - X \overset{P}{\to} 0$ as $n \to \infty$.

Convergence in q.m. implies convergence in P: By Chebyshev's inequality, for each $\varepsilon > 0$

$$0 \leq \mathbf{P}\left\{|X_n - X| > \varepsilon\right\} \leq \frac{\mathbf{E}\left[|X_n - X|^2\right]}{\varepsilon^2}.$$

Therefore, $\mathbf{E}\left[|\ X_n - X\ |^2\right] \to 0$ implies that $\mathbf{P}\left\{|X_n - X| > \varepsilon\right\} \to 0$ as $n \to \infty$.

Exercise 1. Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with $\mathbf{E}(X_n) = \mu_n$ and $\mathrm{Var}(X_n) = \sigma_n^2$ for $n = 1, 2, \ldots$ Suppose that $\lim_{n\to\infty} \mu_n = \mu$ and $\lim_{n\to\infty} \sigma_n^2 = 0$. Then, $X_n \overset{P}{\to} \mu$ as $n \to \infty$.

Convergence in P does not necessarily imply convergence in q.m.:

Example 2. Consider the $U(0,1)$ probability, and the sequence $X_n = \sqrt{n}\mathbf{1}_{(0,1/n)}$ for $n \geq 1$ and $X = 0$. Clearly, X_n converges in probability to X. (Why? - Note that $|X_n - X| = |X_n|$).
But, not in q.m., because $\mathbf{E}\left[(X_n - X)^2\right] = \mathbf{E}\left[X_n^2\right] = 1$ for every $n \geq 1$.

Transformations: Let $h : \mathbb{R} \to \mathbb{R}$ be a continuous function.

(1) If $X_n \xrightarrow{P} X$, then $h(X_n) \xrightarrow{P} h(X)$ as $n \to \infty$.

(2) If $X_n \xrightarrow{q.m.} X$, then $h(X_n)$ does not necessarily converge to $h(X)$ in q.m.
(Think of a *counter-example*!)

2. Weak law of large numbers

Let $X_1, X_2, \dots$ be i.i.d random variables (independent random variables each having the same marginal distribution). Assume that the second moment of X_1 is finite. Then, $\mu = \mathbf{E}[X_1]$ and $\sigma^2 = \text{Var}(X_1)$ are well-defined. (Why?)

Let $S_n = X_1 + \cdots + X_n$ (partial sums) and $\bar{X}_n = \frac{S_n}{n} = \frac{X_1 + \cdots + X_n}{n}$ (*sample mean*). Then, by the properties of expectation and variance, we have

$$\mathbf{E}[S_n] = n\mu, \ \ \text{Var}(S_n) = n\sigma_1^2, \qquad \mathbf{E}[\bar{X}_n] = \mu, \text{ and } \text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}.$$

In particular, s.d.$(\bar{X}_n) = \sigma/\sqrt{n}$ decreases with n. If we apply Chebyshev's inequality to $\bar{X}_n$, we get for any $\delta > 0$ that

$$0 \le \mathbf{P}\{|\bar{X}_n - \mu| \ge \delta\} \le \frac{\sigma^2}{\delta^2 n}.$$

This goes to zero as $n \to \infty$ (with $\delta > 0$ being fixed). This means that for large n, the sample mean is unlikely to be far from μ (sometimes called the "population mean"). This is consistent with our intuitive idea that if we toss a coin (with probability of head p) many times, we can get a better guess of what the value of p is.

Weak law of large numbers (Jacob Bernoulli): With the above notations, for any $\delta > 0$, we have

$$0 \le \mathbf{P}\{|\bar{X}_n - \mu| \ge \delta\} \le \frac{\sigma^2}{\delta^2 n} \to 0 \quad \text{as } n \to \infty.$$

This is very general, in the sense that we only assume the existence of variance. If X_k are assumed to have more moments, one can get better bounds. For example, when X_k are i.i.d. Ber(p), we have the following theorem.

Hoeffding's inequality: Let $X_1, \dots, X_n$ be i.i.d. Ber(p). Then, for any $\delta > 0$

$$0 \le \mathbf{P}\{|\bar{X}_n - p| \ge \delta\} \le 2e^{-2n\delta^2}.$$

A more general version

Let $X_1, \dots, X_n$ be independent, bounded random variables with $X_i \in [a, b]$ for $i = 1, \dots, n$ and $-\infty < a \le b < \infty$. Then, for all $\delta \ge 0$

$$0 \le \mathbf{P}\left(\left|\bar{X}_n - \mathbf{E}[X]\right| \ge \delta\right) \le 2e^{-2n\delta^2/(b-a)^2}.$$

Check the file 'WLLN.R'.