2 Preservation of Stochastic Convergence under Transformations

If $\{a_n\}$ and $\{b_n\}$ are two sequences of real numbers such that $a_n \to a < \infty$ and $b_n \to b < \infty$, then it is easy to show that $a_n + b_n \to a + b$, $a_n b_n \to ab$, $a_n/b_n \to a/b$ (provided $b_n, b \neq 0$), etc. More generally, if $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a continuous function, then $a_n \to a$ and $b_n \to b$ imply $f(a_n, b_n) \to f(a, b)$. In other words, convergence properties of sequences of reals are preserved under various transformations. Some versions of this statement carry over to sequences of random variables.

Theorem 1 (Stochastic convergence of sums of sequences). The following implications hold:

1. If $X_n \xrightarrow{\text{a.s.}} X$ and $Y_n \xrightarrow{\text{a.s.}} Y$, then $X_n + Y_n \xrightarrow{\text{a.s.}} X + Y$.
2. If $X_n \xrightarrow{\text{p.}} X$ and $Y_n \xrightarrow{\text{p.}} Y$, then $X_n + Y_n \xrightarrow{\text{p.}} X + Y$.
3. If $X_n \xrightarrow{\text{m.s.}} X$ and $Y_n \xrightarrow{\text{m.s.}} Y$, then $X_n + Y_n \xrightarrow{\text{m.s.}} X + Y$.
4. If $X_n \xrightarrow{\text{d.}} X$ and $Y_n \xrightarrow{\text{d.}} c$, then $X_n + Y_n \xrightarrow{\text{d.}} X + c$.

Remark 1. In part 4, the convergence of one of the sequences to a deterministic random variable is crucial: it is easy to construct examples when $X_n \xrightarrow{\text{d.}} X$ and $Y_n \xrightarrow{\text{d.}} Y$, but $X_n + Y_n$ does not converge in distribution to $X + Y$.

Proof:

1. Define the events $A = \{\omega : X_n(\omega) \to X(\omega)\}$ and $B = \{\omega : Y_n(\omega) \to Y(\omega)\}$. By hypothesis, $P(A) = P(B) = 1$. Consider the event $C = A \cap B$. For any $\omega \in C$, we have both $X_n(\omega) \to X(\omega)$ and $Y_n(\omega) \to Y(\omega)$. Therefore, if $\omega \in C$, then $X_n(\omega) + Y_n(\omega) \to X(\omega) + Y(\omega)$. To finish the proof, we need to show that $P(C) = 1$. To that end,

$$P(C^c) = P((A \cap B)^c) = P(A^c \cup B^c) \leq P(A^c) + P(B^c) = 0,$$

where we have used De Morgan's identity and the union bound. Thus, $P(C) = 1 - P(C^c) = 1$, so $X_n + Y_n \xrightarrow{\text{a.s.}} X + Y$.

2. By the triangle inequality,

$$|(X_n + Y_n) - (X + Y)| \le |X_n - X| + |Y_n - Y|.$$

Therefore, if $|(X_n + Y_n) - (X + Y)| \ge \varepsilon$, then at least one of $|X_n - X|$ and $|Y_n - Y|$ has to be greater than or equal to $\varepsilon/2$. Consequently,

$$\begin{aligned} &P\left(|(X_n + Y_n) - (X + Y)| \ge \varepsilon\right) \\ &\qquad \le P\left(\{|X_n - X| \ge \varepsilon/2\} \cup \{|Y_n - Y| \ge \varepsilon/2\}\right) \\ &\qquad \le P\left(|X_n - X| \ge \varepsilon/2\right) + P\left(|Y_n - Y| \ge \varepsilon/2\right) \\ &\qquad \xrightarrow{n\to\infty} 0, \end{aligned}$$

where the second step uses the union bound, and the last step is because $X_n \xrightarrow{\text{p.}} X$ and $Y_n \xrightarrow{\text{p.}} Y$.

3. Using the properties of expectation,

$$\begin{aligned} &\mathbb{E}[((X_n + Y_n) - (X + Y))^2] \\ &\qquad = \mathbb{E}[((X_n - X) + (Y_n - Y))^2] \\ &\qquad = \mathbb{E}[(X_n - X)^2] + 2\,\mathbb{E}[(X_n - X)(Y_n - Y)] + \mathbb{E}[(Y_n - Y)^2]. \end{aligned}$$

The first and the last terms converge to zero because $X_n \xrightarrow{\text{m.s.}} X$ and $Y_n \xrightarrow{\text{m.s.}} Y$. For the middle term, Cauchy–Schwarz inequality gives

$$\mathbb{E}[(X_n - X)(Y_n - Y)] \le \sqrt{\mathbb{E}[(X_n - X)^2]\,\mathbb{E}[(Y_n - Y)^2]} \xrightarrow{n\to\infty} 0.$$

Thus, $X_n + Y_n \xrightarrow{\text{m.s.}} X + Y$.

4. The cdf of $Z = X + c$ is given by

$$F_Z(z) = P(Z \le z) = P(X + c \le z) = P(X \le z - c) = F_X(x - c),$$

and the set of its continuity points consists precisely of those z such that $z - c$ is a continuity point of F_X. If z is such a point, then

$$\begin{aligned} P(X_n + Y_n \le z) &= P(X_n + Y_n \le z, Y_n \ge c) + P(X_n + Y_n \le z, Y_n < c) \\ &\le P(X_n \le z - c) + P(Y_n < c). \end{aligned}$$

Since $Y_n \xrightarrow{\text{d.}} c$, $P(Y_n < c) = 1 - P(Y_n = c) \to 0$. Therefore,

$$\limsup_{n\to\infty} P(X_n + Y_n \le z) \le F_X(z - c) = F_Z(z).$$

On the other hand,

$$\begin{aligned}P(X_n + Y_n > z) &= P(X_n + Y_n > z, Y_n > c) + P(X_n + Y_n > z, Y_n \le c) \\ &\le P(Y_n > c) + P(X_n > z - c) \\ &= P(Y_n > c) + 1 - P(X_n \le z - c),\end{aligned}$$

so

$$\liminf_{n\to\infty} P(X_n + Y_n \le z) \ge F_X(z - c) = F_Z(z).$$

Consequently,

$$F_Z(z) \le \liminf_{n\to\infty} P(X_n + Y_n \le z) \le \limsup_{n\to\infty} P(X_n + Y_n \le z) \le F_Z(z),$$

which implies that $X_n + Y_n \xrightarrow{\text{d.}} Z = X + c$. (See course notes for the definitions of liminf and limsup.)

■

More generally:

Theorem 2. Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a continuous function. Then:

1. If $X_n \xrightarrow{\text{a.s.}} X$ and $Y_n \xrightarrow{\text{a.s.}} Y$, then $f(X_n, Y_n) \xrightarrow{\text{a.s.}} f(X, Y)$.
2. If $X_n \xrightarrow{\text{p.}} X$ and $Y_n \xrightarrow{\text{p.}} Y$, then $f(X_n, Y_n) \xrightarrow{\text{p.}} f(X, Y)$.
3. If $X_n \xrightarrow{\text{m.s.}} X$ and $Y_n \xrightarrow{\text{m.s.}} Y$, then $f(X_n, Y_n) \xrightarrow{\text{m.s.}} f(X, Y)$.
4. If $X_n \xrightarrow{\text{d.}} X$ and $Y_n \xrightarrow{\text{d.}} c$, then $f(X_n, Y_n) \xrightarrow{\text{d.}} f(X_n, c)$.

As a corollary, we immediately obtain the following:

Theorem 3. Let $\{X_n\}$ and $\{Y_n\}$ be two sequences of random variables defined on the same probability space. Then:

1. If $X_n \xrightarrow{*} X$ and $Y_n \xrightarrow{*} Y$, where $*$ is either a.s. or p., then $aX_n + bY_n \xrightarrow{*} aX + bY$, $X_nY_n \xrightarrow{*} XY$, and $X_n/Y_n \xrightarrow{*} X/Y$ (provided $Y_n \neq 0$ for all n and $Y \neq 0$ with probability one).
2. If $X_n \xrightarrow{\text{d.}} X$ and $Y_n \xrightarrow{\text{d.}} c$, then $X_nY_n \xrightarrow{\text{d.}} cX$.

Remark 2. The result that says $X_n \xrightarrow{\text{d.}} X$ and $Y_n \xrightarrow{\text{d.}} c$ implies $X_n + Y_n \xrightarrow{\text{d.}} X + c$ and $X_nY_n \xrightarrow{\text{d.}} cX$ is known as *Slutsky's theorem*.