I. Definitions

**Definition 1** (Convergence almost surely and in probability)

Let \((X_n)_{n \geq 0}, X\) be random vectors in \(\mathbb{R}^k\), defined on a probability space \((\Omega, \mathcal{A}, \mathbb{P})\).

i) We say that \(X_n\) converges to \(X\) **almost surely**, or \(X_n \xrightarrow{a.s.} X\) as \(n \to \infty\), if
\[
\mathbb{P}(\omega \in \Omega : \|X_n(\omega) - X(\omega)\| \to 0 \text{ as } n \to \infty) = \mathbb{P}(\|X_n - X\| \to 0 \text{ as } n \to \infty) = 1.
\]

ii) We say that \(X_n\) converges to \(X\) **in probability**, or \(X_n \xrightarrow{P} X\) as \(n \to \infty\), if for all \(\varepsilon > 0\)
\[
\mathbb{P}(\|X_n - X\| > \varepsilon) \to 0.
\]

**Remark** Convergence of vectors is equivalent to convergence of each coefficient, for both these definitions. This follows naturally from the definition for almost sure convergence, and addressed in the Example Sheet for convergence in probability.

**Definition 2** (Convergence in distribution)

Let \((X_n)_{n \geq 0}, X\) be random vectors in \(\mathbb{R}^k\), defined on a probability space \((\Omega, \mathcal{A}, \mathbb{P})\). We say that \(X_n\) converges to \(X\) **in distribution**, or \(X_n \xrightarrow{d} X\) as \(n \to \infty\) if
\[
\mathbb{P}(X_n \leq t) \to \mathbb{P}(X \leq t),
\]
for all \(t\) where the map \(t \mapsto \mathbb{P}(X \leq t)\) is continuous.

**Remark** We write \(\{X \leq t\}\) as shorthand for \(\{X(1) \leq t_1, \ldots, X(k) \leq t_k\}\). For \(k = 1\), the definition becomes \(\mathbb{P}(X_n \leq t) \to \mathbb{P}(X \leq t)\).

II. Propositions

**Proposition 1** The following facts about stochastic convergence follow from these definitions and can be proved in measure theory.

i) \(X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X\) as \(n \to \infty\).

ii) (Continuous mapping theorem) If \(X_n, X\) take values in \(X \subset \mathbb{R}^d\) and \(g : X \to \mathbb{R}\) is continuous, then \(X_n \xrightarrow{a.s./P/d} X\) implies \(g(X_n) \xrightarrow{a.s./P/d} g(X)\).
iii) (Slutsky’s lemma) Let $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} c$ where $c$ is deterministic (or non-stochastic). Then, as $n \to \infty$

- a) $Y_n \xrightarrow{P} c$
- b) $X_n + Y_n \xrightarrow{d} X + c$
- c) $(k = 1) X_n Y_n \xrightarrow{d} cX$ and if $c \neq 0$, $X_n/Y_n \xrightarrow{d} X/c$.
- d) If $(A_n)_{n \geq 0}$ are random matrices such that $(A_n)_{ij} \xrightarrow{P} A_{ij}$, where $A$ is deterministic (or non-stochastic), then $A_nX_n \xrightarrow{d} AX$.

iv) If $X_n \xrightarrow{d} X$ as $n \to \infty$, then $(X_n)_{n \geq 0}$ is bounded in probability, or $X_n = O_P(1)$, i.e.

$$\forall \varepsilon > 0 \exists M(\varepsilon) < \infty \text{ such that for all } n \geq 0, \ P(\|X_n\| > M(\varepsilon)) < \varepsilon.$$