Convergence of Sequences of Random Variables

- Definitions
  - ⋆ $X_n \stackrel{a.s.}{\to} X$ means $P\{s : \lim_{n\to\infty} X(s) = X(s)\} = 1$.
  - ⋆ $X_n \stackrel{P}{\to} X$ means $\forall \epsilon > 0$, $\lim_{n\to\infty} P\{|X_n - X| < \epsilon\} = 1$.
  - ⋆ $X_n \stackrel{d}{\to} X$ means for every continuity point x of F_X, $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$.
- $X_n \stackrel{a.s.}{\to} X$ if and only if $\forall \epsilon > 0$, $\lim_{n\to\infty} P(\cap_{k=n}^{\infty}\{|X_k - X| < \epsilon\}) = 1$.
- $X_n \stackrel{a.s.}{\to} X \Rightarrow X_n \stackrel{P}{\to} X \Rightarrow X_n \stackrel{d}{\to} X$.
- If a is a constant, $X_n \stackrel{d}{\to} a \Rightarrow X_n \stackrel{P}{\to} a$.
- If $\lim_{n\to\infty} f_{X_n}(x) = f_X(x)$ for each x, $X_n \stackrel{d}{\to} X$.
- Let $\mathbf{X}$ and $\mathbf{X}_n$ be random vectors in $\mathbb{R}^k$. $\mathbf{X}_n \stackrel{d}{\to} \mathbf{X}$ if and only if $\lim_{n\to\infty} E[g(\mathbf{X}_n)] = E[g(\mathbf{X})]$ for every bounded continuous function $g : \mathbb{R}^k \to \mathbb{R}$.
- Slutsky Theorems for Convergence in Distribution:
  1. If $\mathbf{X}_n \in \mathbb{R}^m$, $\mathbf{X}_n \stackrel{d}{\to} \mathbf{X}$ and if $f : \mathbb{R}^m \to \mathbb{R}^q$ (where $q \le m$) is continuous except possibly on a set C with $P(\mathbf{X} \in C) = 0$, then $f(\mathbf{X}_n) \stackrel{d}{\to} f(\mathbf{X})$.
  2. If $\mathbf{X}_n \stackrel{d}{\to} \mathbf{X}$ and $(\mathbf{X}_n - \mathbf{Y}_n) \stackrel{P}{\to} 0$, then $\mathbf{Y}_n \stackrel{d}{\to} \mathbf{X}$.
  3. If $\mathbf{X}_n \in \mathbb{R}^d$, $\mathbf{Y}_n \in \mathbb{R}^k$, $\mathbf{X}_n \stackrel{d}{\to} \mathbf{X}$ and $\mathbf{Y}_n \stackrel{d}{\to} \mathbf{c}$, then
  $$\begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} \stackrel{d}{\to} \begin{pmatrix} \mathbf{X} \\ \mathbf{c} \end{pmatrix}$$
- Slutsky Theorems for Convergence in Probability:
  1. If $\mathbf{X}_n \in \mathbb{R}^m$, $\mathbf{X}_n \stackrel{P}{\to} \mathbf{X}$ and if $f : \mathbb{R}^m \to \mathbb{R}^q$ (where $q \le m$) is continuous except possibly on a set C with $P(\mathbf{X} \in C) = 0$, then $f(\mathbf{X}_n) \stackrel{P}{\to} f(\mathbf{X})$.
  2. If $\mathbf{X}_n \stackrel{P}{\to} \mathbf{X}$ and $(\mathbf{X}_n - \mathbf{Y}_n) \stackrel{P}{\to} 0$, then $\mathbf{Y}_n \stackrel{P}{\to} \mathbf{X}$.
  3. If $\mathbf{X}_n \in \mathbb{R}^d$, $\mathbf{Y}_n \in \mathbb{R}^k$, $\mathbf{X}_n \stackrel{P}{\to} \mathbf{X}$ and $\mathbf{Y}_n \stackrel{P}{\to} \mathbf{Y}$, then
  $$\begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} \stackrel{P}{\to} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$$

- Let $g(x)$ have a second derivative that is continuous at $x = \theta$, and let $\sqrt{n}(T_n - \theta) \xrightarrow{d} T$. Then $\sqrt{n}(g(T_n) - g(\theta)) \xrightarrow{d} g'(\theta)T$.

- Strong Law of Large Numbers (SLLN): Let $X_1, \ldots X_n$ be i.i.d. random variables with finite first moment. Then $\overline{X}_n \xrightarrow{a.s.} E(X_1)$.

- Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be i.i.d. random vectors in $\mathbb{R}^k$ with expected value vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$. Then $\sqrt{n}(\overline{\mathbf{X}}_n - \boldsymbol{\mu})$ converges in distribution to a multivariate normal with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$.

- Lindeberg Central Limit Theorem: Consider the triangular array of random variables

$$\begin{array}{lll} X_{11} & & \\ X_{21}, & X_{22} & \\ X_{31}, & X_{32}, & X_{33} \\ \ldots, & & \end{array}$$

  where the random variables in each row are assumed independent with $E(X_{ij}) = 0$ and $Var(X_{ij}) = \sigma_{ij}^2$. Let $S_n = \sum_{j=1}^n X_{nj}$, and let $v_n^2 = Var(S_n) = \sum_{j=1}^n \sigma_{nj}^2$. Then $\frac{S_n}{v_n}$ converges in distribution to a standard normal provided, for all $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{v_n^2} \sum_{j=1}^n E[X_{nj}^2 I(|X_{nj}| \geq \epsilon v_n)] = 0$$

- $U_n = O_p(V_n)$ means $\forall \epsilon > 0$, $\exists M = M(\epsilon)$ and $N = N(\epsilon)$ such that if $n > N$, $P\{\frac{U_n}{V_n} \leq M\} - P\{\frac{U_n}{V_n} \leq -M\} > 1 - \epsilon$. Usually, $V_n = n^{-t}$ for some $t > 0$.