Measures

An extended real-valued function on a class of subsets of a nonempty set Ω is a **set function**. A set function μ on an algebra $\mathcal{F}$ on Ω is a **measure** if (1) $\mu(A) \in [0, \infty], \quad \forall A \in \mathcal{F}$, (2) $\mu(\emptyset) = 0$, (3) $A_1, A_2, \dots \in \mathcal{L}$ **disjoint** with $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$, then $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.

If $\mu(\Omega) = 1$, then μ is a **probability measure**. If $\exists B_1, B_2 \in \mathcal{F}$ such that $\Omega = \bigcup_{n=1}^{\infty} B_n$ and $\mu(B_i) < \infty, \ \forall i \geq 1$, then μ is σ**-finite**.

Extensions of Measures: If μ is σ-finite meas on alg $\mathcal{F}$, then $\exists$ *unique* extsn of μ to $\sigma\langle\mathcal{F}\rangle$.

Monotonicity: If $A, B \in \mathcal{F}$ such that $A \subset B$, then $P(A) \leq P(B)$.

Inclusion-exclusion formula: If $A_1, A_2, \dots, A_n \in \mathcal{F}$, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \cdots + (-1)^{n-1} P(A_1 \cap \cdots \cap A_n).$$

Countable subadditivity: If $A_1, A_2, \dots \in \mathcal{F}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$, then $P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$.

Monotone continuity from below (mcfb): If $A, A_1, A_2, \dots \in \mathcal{F}$ such that $A_n \uparrow A$ (i.e., $A_n \subset A_{n+1}$ and $A = \bigcup A_n$), then $P(A_n) \uparrow P(A)$.

Monotone continuity from above (mcfa): If $A, A_1, A_2, \dots \in \mathcal{F}$ such that $A_n \downarrow A$ (i.e., $A_n \supset A_{n+1}$ and $A = \bigcap A_n$), then $P(A_n) \downarrow P(A)$.

Measurable Transformations

If $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ are measurable spaces, then $T : \Omega_1 \to \Omega_2$ is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$**-measurable** if $T^{-1}(A) \equiv \{\omega \in \Omega_1 : T(\omega) \in A\} \in \mathcal{F}_1$ (equivalently, $T^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1$).

Let $(\Omega, \mathcal{F}, P)$ be a psp. Then, $X : \Omega \to \mathbb{R}$ is a **r.v.** if it is $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-measurable.

Checking for measurability: $T : \mathbb{R} \to \mathbb{R}$ is $\langle \mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}) \rangle$-measurable $\iff T^{-1}(-\infty, r) = \{\omega \in \mathbb{R} : T(\omega) < r\} \in \mathcal{B}(\mathbb{R}) \ \forall r \in \mathbb{R}$.

If T_i is $\langle \mathcal{F}_i, \mathcal{F}_{i+1} \rangle$-measurable for $i = 1, 2$, then the composition $T = T_1 \circ T_2 = T_1(T_2(\cdot))$ is $\langle \mathcal{F}_1, \mathcal{F}_3 \rangle$-measurable.

If $f : \mathbb{R}^k \to \mathbb{R}^p$ is continuous, then f is $\langle \mathcal{B}(\mathbb{R}), \mathcal{B}(\mathbb{R}) \rangle$-measurable.

If $f_1, f_2, \dots, f_n$ are each $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-measurable transformations from Ω to $\mathbb{R}$, then $f = (f_1, f_2, \dots, f_n)'$ is $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-measurable. Also, $\sum_{i=1}^{n} f_i, \prod_{i=1}^{n} f_i, \sup_n f_n, \inf_n f_n, \limsup f_n, \liminf f_n$, and $\mathbb{I}_{\{\omega \in \Omega : \lim_{n\to\infty} f_n(\omega) \text{ finitely exists}\}} \lim_{n\to\infty} f_n$ are all $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-measurable.

Induced Measures & Distribution Functions

If $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ are measurable spaces and $T : \Omega_1 \to \Omega_2$ is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable, then for any measure μ on $(\Omega_1, \mathcal{F}_1)$, the set function μT^{-1} defined by $\mu T^{-1}(A) = \mu \circ T^{-1}(A) = \mu(T^{-1}(A)), \quad \forall A \in \mathcal{F}_2$ is a measure on $\mathcal{F}_2$ and is called the **measure induced** by T under μ on $\mathcal{F}_2$.

For a r.v. X on a psp $(\Omega, \mathcal{F}, P)$, the probability **distribution** of X (or the law of X), denoted P_X, is the induced measure of X under P on $\mathcal{B}(\mathbb{R})$, i.e.,
$P_X(A) = P(X^{-1}(A)) = P(\{\omega \in \Omega : X(\omega) \in A\}) = P(X \in A), \quad A \in \mathcal{B}(\mathbb{R})$.

The **cumulative distribution function** (cdf) of a r.v. X is
$F(x) = P_X((-\infty, x]) = P(X \leq x), \quad x \in \mathbb{R}$.
(1) F is right continuous: if $x_n \downarrow x_0$ and $x_n \geq x$ then
$F(x_n) = P_X((-\infty, x_n]) \downarrow P_X((-\infty, x_0]) = F(x_0)$ by mcfa.
(2) F is monotone nondecreasing: if $x \leq y \implies (-\infty, x] \subset (-\infty, y]$ then
$F(x) = P_X((-\infty, x]) \leq P_X((-\infty, y]) = F(y)$ by monotonicity
(3) $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$: show using argument similar to (1).

If $F : \mathbb{R} \to [0, 1]$ satisfies (1)-(3) above, then there exists a r.v. X on a psp $(\Omega, \mathcal{F}, P)$ such that F is the cdf of X.

Integrals

A measurable function $f : \Omega \to \overline{\mathbb{R}}$ on $(\Omega, \mathcal{F}, \mu)$ is μ**-integrable** if $\int_\Omega |f| \, d\mu < \infty$.

If $(\Omega, \mathcal{F}, \mu)$ is a msp and $f : \Omega \to \mathbb{R}$ is measurable and either μ-integrable or nonnegative. Then, for any $A \in \mathcal{F}$, $\int_A f \, d\mu = \int_\Omega \mathbb{I}_A f \, d\mu$.

Using the DCT (see next section), it can be shown that for **disjoint** $A_1, A_2, \dots \in \mathcal{F}$ and measurable $f : \Omega \to \mathbb{R}$ either μ-integrable or nonnegative, then $\int_\Omega f \mathbb{I}_{\bigcup_{n=1}^{\infty} A_n} d\mu = \sum_{n=1}^{\infty} \int_{A_i} f \, d\mu$.

Convergence Theorems

Monotone Convergence Theorem (MCT): If $f_n : \Omega \to \overline{\mathbb{R}}$ is an increasing sequence of nonnegative measurable functions, i.e., $f_n(\omega) \leq f_{n+1}(\omega) \ \forall \omega \in \Omega$ and $f_n(\omega) \uparrow f(\omega)$ a.e.(μ), then $\int_\Omega f_n \, d\mu \uparrow \int_\Omega f \, d\mu$. That is, $\int_\Omega f \, d\mu = \int_\Omega \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int_\Omega f_n \, d\mu$..

Fatou's Lemma: If $f_n : \Omega \to \overline{\mathbb{R}}$ is a sequence of nonnegative functions, then $\int_\Omega \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n \, d\mu$.

Dominated Convergence Theorem (DCT): Suppose (1) $g : \Omega \to \overline{\mathbb{R}}$ is a nonnegative, μ-integrable function; (2) $|f_n| \leq g$ a.e.(μ) $\forall n \geq 1$; and (3) $f_n \to f$ a.e.(μ). Then, f is μ-integrable and $\lim_{n\to\infty} \int_\Omega |f_n - f| \, d\mu = 0$ and $\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu$.

Result under weaker conditions, UI and $\mu(\Omega) < \infty$: If $(\Omega, \mathcal{F}, \mu)$ is a msp with $\mu(\Omega) < \infty$ and $f, f_n : \Omega \to \overline{R}$ are measurable such that $f_n \to f$ a.e.(μ) and $\{f_n : n \geq 1\}$ is UI (see next section), then f is μ-integrable and $\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu$.

Note: for a nonnegative measurable real-valued function f on a msp $(\Omega, \mathcal{F}, \mu)$, $\nu(A) = \int_A f \, d\mu \ \forall A \in \mathcal{F}$ is a measure on $(\Omega, \mathcal{F})$ and f is a density of the measure ν.

Scheffe's Theorem: Let $(\Omega, \mathcal{F}, \mu)$ be a msp and for $n \geq 0$, let $\nu_n(A) = \int_A f_n \, d\mu \ \forall A \in \mathcal{F}$ be finite measures on $\mathcal{F}$ with densities $f_n \geq 0$. If $\nu_n(\Omega) = \nu_0(\Omega) < \infty$ for all $n \geq 1$ and $f_n \to f$ a.e.(μ), then $\lim_{n\to\infty} \int_\Omega |f_n - f_0| \, d\mu = 0$.
Also, $\sup_{A \in \mathcal{F}} |\nu_n(A) - \nu_0(A)| = \frac{1}{2} \int_\Omega |f_n - f_0| \, d\mu \to 0$ as $n \to \infty$.

Uniform Integrability

Recall if $f : \Omega \to \mathbb{R}$ is μ-integrable, then by the DCT $\lim_{n\to\infty} \int_{|f|>n} |f| \, d\mu = \lim_{n\to\infty} \int_\Omega \mathbb{I}_{|f|>n} |f| \, d\mu = 0$.

A family of μ-integrable functions $\{f_\lambda : \lambda \in \Lambda\}$ on a msp $(\Omega, \mathcal{F}, \mu)$ is **uniformly integrable** (UI) w.r.t. μ if $\sup_{\lambda \in \Lambda} \int_{|f_\lambda|>t} |f_\lambda| \, d\mu \to 0$ as $t \to \infty$.

Suppose $\mathcal{A} \equiv \{f_\lambda : \lambda \in \Lambda\}$ is a collection of μ-integrable functions on a msp $(\Omega, \mathcal{F}, \mu)$. Then, (1) if Λ is a finite set, then $\mathcal{A}$ is UI; (2) if $\exists \varepsilon > 0$ such that $\sup\left\{\int |f_\lambda|^{1+\varepsilon} \, d\mu : \lambda \in \Lambda\right\} < \infty$, then $\mathcal{A}$ is UI; (3) if $|f_\lambda| \leq f$ a.e.(μ) and $\int f \, d\mu < \infty$, then $\mathcal{A}$ is UI; (4) if $\mathcal{A}$ is UI and $\mu(\Omega) < \infty$, then $\exists M > 0$ such that $\sup\left\{\int |f_\lambda| \, d\mu : \lambda \in \Lambda\right\} \leq M$; (5) if $\{f_\lambda : \lambda \in \Lambda\}$ and $\{g_\lambda : \lambda \in \Lambda\}$ are both UI, then $\{f_\lambda + g_\lambda : \lambda \in \Lambda\}$ is also UI.

Independence

A collection $A_i, \ i \in I$ of sets in $\mathcal{F}$ are **ind** if $\forall i_1, i_2, \dots, i_n \in I$ distinct indices and fixed $1 \leq n < \infty$, $P\left(\bigcap_{j=1}^{n} A_{i_j}\right) = \prod_{j=1}^{n} P\left(A_{i_j}\right)$. ($2^n - n - 1$ ind conditions!)

Suppose $\mathcal{G}_i \subset \mathcal{F}$ is a collection of measurable sets for each $i \in I$. Then the family of sets $\{\mathcal{G}_i : i \in I\}$ is called **ind** if any possible collection $\{A_i : i \in I\}$ of sets are ind, where $\{A_i : i \in I\}$ is formed by choosing an arbitrary set A_i from $\mathcal{G}_i$ for each $i \in I$. That is, $\forall i_1, i_2, \dots, i_n \in I$ distinct indices, fixed $1 \leq n < \infty$ and $\forall A_{i_1}, A_{i_2}, \dots, A_{i_n} \in \mathcal{G}_{i_n}$, $P\left(\bigcap_{j=1}^{n} A_{i_j}\right) = \prod_{j=1}^{n} P\left(A_{i_j}\right)$.
Note the family of sets $\{\mathcal{G}_i : i \in I\}$ are ind iff for each finite $T \subset I$, $\{\mathcal{G}_i : i \in T\}$ are ind.

A collection of r.v.'s $X_i, \ i \in I$ on $(\Omega, \mathcal{F}, P)$ are **ind** if the family $\{\sigma\langle X_i\rangle : i \in I\}$ is ind, where $\sigma\langle X_i\rangle = \left\{X_i^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\right\} = X_i^{-1}(\mathcal{B}(\mathbb{R}))$ is the σ-algebra generated by X_i. That is, $\forall i_1, i_2, \dots, i_n \in I$ distinct indices, fixed $1 \leq n < \infty$ and $\forall B_{i_1}, B_{i_2}, \dots, B_{i_n} \in \mathcal{B}(\mathbb{R})$, $P\left(X_{i_1} \in B_{i_1}, X_{i_2} \in B_{i_2}, \dots, X_{i_n} \in B_{i_n}\right) = \prod_{j=1}^{n} P\left(X_{i_j} \in B_{i_j}\right)$. In terms of distribution functions, $X_i, i \in I$ are ind iff $\forall x_1, x_2, \dots, x_n \in \mathbb{R}$ and $\forall i_1, i_2, \dots, i_n \in I$ distinct indices, $P\left(X_{i_1} \leq x_1, X_{i_2} \leq x_2, \dots, X_{i_n} \leq x_n\right) = \prod_{j=1}^{n} P\left(X_{i_j} \leq x_j\right)$.

Independence of generated σ**-algebras**: If $(\Omega, \mathcal{F}, P)$ is a psp, $\mathcal{G}_i \subset \mathcal{F}$ is a π-class for each index $i \in I$, and the family $\{\mathcal{G}_i : i \in I\}$ is ind. Then, the family $\{\sigma\langle\mathcal{G}_i\rangle : i \in I\}$ is ind.

Borel-Cantelli Lemmas

If $(\Omega, \mathcal{F})$ be a measurable space and $A_1, A_2, \dots \in \mathcal{F}$, then
$\limsup_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n \in \mathcal{F}$ and
$\liminf_{n\to\infty} A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n \in \mathcal{F}$. Also,
$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$,
("A_n i.o")$^c = \left(\overline{\lim} A_n\right)^c = \underline{\lim} A_n^c =$ "A_n^c eventually,", and
("A_n eventually")$^c = (\underline{\lim} A_n)^c = \overline{\lim} A_n^c =$ "A_n^c i.o.".

Borel-Cantelli Lemma: Let $(\Omega, \mathcal{F}, P)$ be a psp and $A_1, A_2, \dots \in \mathcal{F}$.
(1) If $\sum_n P(A_n) < \infty$, then $P\left(\overline{\lim} A_n\right) = 0$.
(2) If $\{A_n\}$ are ind and $\sum_n P(A_n) = \infty$, then $P\left(\overline{\lim} A_n\right) = 1$.

Borel 0-1 Law: If $A_1, A_2, \dots$ are ind events, then

$$P(\overline{\lim} A_n) = P(A_n \text{ i.o.}) = \begin{cases} 0 & \text{iff } \sum_n P(A_n) < \infty, \\ 1 & \text{iff } \sum_n P(A_n) = \infty. \end{cases}$$

Tail Events & K's 0-1 Law

The **tail** σ**-algebra** of a sequence of r.v.'s $X_1, X_2, \dots$ on a psp $(\Omega, \mathcal{F}, P)$ is $\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma\left\langle\left\{X_j : j \geq n\right\}\right\rangle$, where $\sigma\left\langle\left\{X_j : j \geq n\right\}\right\rangle = \sigma\left\langle\left\{X_j^{-1} : B \in \mathcal{B}(\mathbb{R}), j \geq n\right\}\right\rangle$ is the σ-algebra generated by $X_j, j \geq n$. Any set (event) $A \in \mathcal{T}$ is a **tail event**.

An extended real-valued r.v. $T : \Omega \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is a **tail r.v.** if T is $\langle \mathcal{F}, \mathcal{B}(\overline{\mathbb{R}}) \rangle$-measurable. That is, $\forall r \in \mathbb{R}, \ T^{-1}([-\infty, r)) = \{\omega \in \Omega : T(\omega) < r\} \in \mathcal{T}$.

Kolmogorov's 0-1 Law: Tail events of a seq $X_1, X_2, \dots$ of ind r.v.'s have probs 0 or 1, i.e., if $\{X_n\}_{n\geq 1}$ are ind and $A \in \mathcal{T} = \bigcap_n \sigma\langle\left\{X_j : j \geq n\right\}\rangle$, then $P(A) \in \{0, 1\}$.

Corollary: For a psp $(\Omega, \mathcal{F}, P)$ and tail σ-algebra $\mathcal{T}$ defined by a sequence $X_1, X_2, \dots$ of ind r.v.'s, if $T : \Omega \to \overline{\mathbb{R}}$ is a tail r.v. (that is, T is $\langle \mathcal{T}, \mathcal{B}(\overline{\mathbb{R}}) \rangle$-measurable), then T is degenerate. That is, $\exists c \in \overline{\mathbb{R}}$ such that $P(T = c) = 1$.

Convergence of r.v.'s

A sequence of r.v.'s $X_1, X_2, \dots$ on a psp $(\Omega, \mathcal{F}, P)$ **converge almost surely** to a r.v. X_0 on $(\Omega, \mathcal{F}, P)$ if $P\left(\{\lim_{n\to\infty} X_n(\omega) = X_0(\omega)\}\right) = 1$.
TFAE:
(1) $X_n \xrightarrow{\text{a.s.}} 0$,
(2) $P(|X_n| > \varepsilon \text{ i.o}) = 0$ for all $\varepsilon > 0$,
(3) $\sup_{j\geq n} \left|X_j - X_0\right| \xrightarrow{\text{p}} 0$ as $n \to \infty$,
(4) $\lim_{n\to\infty} P\left(\bigcap_{j=n}^{\infty} \left[\left|X_j - X_0\right| \leq \varepsilon\right]\right) = 1$ for all $\varepsilon > 0$.

A sequence of r.v.'s $X_1, X_2, \dots$ on a psp $(\Omega, \mathcal{F}, P)$ **converge in probability** to a r.v. X_0 on $(\Omega, \mathcal{F}, P)$ if $\lim_{n\to\infty} P(|X_n - X_0| > \varepsilon) = 0, \quad \forall \varepsilon > 0$.
TFAE:
(1) $X_n \xrightarrow{\text{p}} 0$,
(2) $\sup_{m\geq n} (|X_m - X_n| > \varepsilon) \to 0$ as $n \to \infty$ for all $\varepsilon > 0$,
(3) $\forall \left\{n_j\right\}$ of $\{X_n\}$, $\exists \{n_{j_k}\}$ such that $X_{n_{j_k}} \xrightarrow{\text{a.s.}} X_0$.

For a sequence of r.v.'s $X_1, X_2, \dots$ on a psp $(\Omega, \mathcal{F}, P)$,
(1) if $X_n \xrightarrow{\text{a.s.}} X_0$ and $g : \mathbb{R} \to \mathbb{R}$ is continuous, then $g(X_n) \xrightarrow{\text{a.s.}} g(X_0)$,
(2) if $X_n \xrightarrow{\text{p}} X_0$ and $g : \mathbb{R} \to \mathbb{R}$ is continuous, then $g(X_n) \xrightarrow{\text{p}} g(X_0)$.

A sequence $X_1, X_2, \dots$ of $\mathcal{L}_r(\Omega, \mathcal{F}, P) \equiv \left\{\text{mble } X \in \mathbb{R} : \int_\Omega |X|^r \, dP < \infty\right\}$ funcs **converges in** $\mathcal{L}_r$ to a mble func X if $\lim_{n\to\infty} \int_\Omega |X_n - X|^r \, dP = 0$.

If $X \in \mathcal{L}_r$, then $t^r P(|X| > t) \to 0$ as $t \to \infty$, that is, $\uparrow r \implies$ faster convergence. If $\exists p \in (0, \infty)$ such that $t^p P(|X| > t) \to 0$, then $X \in \mathcal{L}_r \ \forall r \in (0, p)$.

If the r.v.'s $X_1, X_2, \dots \in \mathcal{L}_r$, then $\exists X \in \mathcal{L}_r$ such that $X_n \xrightarrow{\mathcal{L}_r} X$ iff $\sup_{m\geq n} \mathrm{E}\,|X_m - X_n|^r \to 0$ as $n \to \infty$.

For a r.v. X, the fixed, real-valued $m(X)$ is a **median** if $P(X \geq m(X)) \geq 1/2$ and $P(X \leq m(X)) \geq 1/2$. Can be defined as $\inf\{x \in \mathbb{R} : P(X \leq x) \geq 1/2\}$. If $P(|X| \geq c) < \varepsilon \leq 1/2$ then $|m(X)| \leq c$.

Levy's Inequality: If $X_1, X_2, \dots, X_n$ are ind r.v.'s on a psp $(\Omega, \mathcal{F}, P)$ and $S_j = \sum_{i=1}^{j} X_i, 1 \leq j \leq n$, then $\forall \varepsilon > 0$,
(1) $P\left(\max_{1\leq j\leq n} \left[S_j - m(S_j - S_n)\right] \geq \varepsilon\right) \leq 2P(S_n \geq \varepsilon)$,
(2) $P\left(\max_{1\leq j\leq n} \left[S_j - m(S_j - S_n)\right] \geq \varepsilon\right) \leq 2P(|S_n| \geq \varepsilon)$.

Levy's Theorem: If $X_1, X_2, \dots$ are ind r.v.'s on a psp $(\Omega, \mathcal{F}, P)$ and $S_n = \sum_{i=j}^{n} X_j, n \geq 1$, then S_n converges a.s.(P) iff S_n converges in probability.

Khintchine-Kolmogorov Convergence Theorem: If $X_1, X_2, \dots$ are ind r.v.'s on a psp $(\Omega, \mathcal{F}, P)$ with $\mathrm{E}(X_n) = 0$ and $\mathrm{E}(X_n^2) < \infty$ for all $n \geq 1$ and $\sum_n \mathrm{E}(X_n^2) < \infty$, then $S_n = \sum_{i=j}^{n} X_j, n \geq 1$ converges a.s.(P) and in $\mathcal{L}_2$ to some random variable $S = \sum_n X_n$. Also, $\mathrm{E}(S) = 0$, $\mathrm{E}(S^2) = \sum_n E(X_n^2)$.

Corollary: If $X_1, X_2, \dots$ are ind r.v.'s on a psp $(\Omega, \mathcal{F}, P)$ with $\sum_n \mathrm{E}(X_n) < \infty$ and $\sum_n \sigma^2_{X_n} < \infty$, then $S_n = \sum_{j=1}^{n} X_j, n \geq 1$ converges a.s.(P) to $S = \sum_n X_n$.

Two sequences of r.v.'s $\{X_n\}$ and $\{Y_n\}$ are **tail equivalent** if $\sum_n P(X_n \neq Y_n) < \infty$. If $\{X_n\}$ and $\{Y_n\}$ are tail equivalent, then
(1) By Borel-Cantelli, $P\left(\overline{\lim}(X_n \neq Y_n)\right) = 0 \implies P\left(X_n = Y_n \text{ for large } n\right) = 1$,
(2) $S_n = \sum_{j=1}^{n} X_j \xrightarrow{\text{a.s.}} S \iff S'_n = \sum_{j=1}^{n} Y_j \xrightarrow{\text{a.s.}} S'$,
(3) If $b_n \to \infty$, then $\frac{\sum_{j=1}^{n} X_j}{b_n} \xrightarrow{\text{a.s.}} 0 \iff \frac{\sum_{j=1}^{n} Y_j}{b_n} \xrightarrow{\text{a.s.}} 0$.

Berry-Esseen Lemma: If $X_1, X_2, \dots, X_n$ are ind r.v.'s with $\mathrm{E}(X_i) = 0$ and $\mathrm{E}\left|X_i\right|^3 < \infty, 1 \leq i \leq n$, then $\forall n \geq 4$,
$\sup_{x\in\mathbb{R}} \left|P\left(\frac{S_n}{\sigma_n} \leq x\right) - \Phi(x)\right| \leq \frac{2.75}{\sigma_n^3} \sum_{i=1}^{n} \mathrm{E}\left|X_i\right|^3$, where
$S_n = \sum_{j=1}^{n} X_j$, $\sigma_n^2 = \mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{E}(X_i^2)$, and $\Phi(\cdot)$ is the $\mathcal{N}(0, 1)$ cdf.

Kolmogorov's 3-Series Theorem: If $X_1, X_2, \dots$ are ind r.v.'s on a psp $(\Omega, \mathcal{F}, P)$, for fixed $c > 0$, define $\sum_n P(|X_n| > c), \quad \sum_n \mathrm{E}(X_n^{(c)}), \quad \sum_n \mathrm{Var}(X_n^{(c)})$, where $X_n^{(c)} = X_n \mathbb{I}_{|X_n|\leq c}$. Then,

(1) if the 3 series conv for *some* $c > 0$, then $S_n = \sum_{j=1}^{n} S_j$, $n \geq 1$ conv a.s.(P),
(2) if $S_n = \sum_{j=1}^{n} S_j$, $n \geq 1$ converges a.s.(P), then the 3 series converge for *all* $c > 0$.

Corollary: If $X_1, X_2, \ldots$ are ind r.v.'s on a psp $(\Omega, \mathcal{F}, P)$ with $\mathrm{E}(X_n) = 0$, then
(1) if $\sum_n \left[\mathrm{E}(X_n^{(c)})^2 + \mathrm{E}\,|X_n| \mathbb{I}_{|X_n|>c} \right] < \infty$ for *some* $c > 0$, then $S_n = \sum_{j=1}^{n} X_j$ a.s.(P),
(2) if $\sum_n \mathrm{E}\,|X_n|^{\alpha_n} < \infty$ for *some* $\{\alpha_n\} \subset [1, 2]$, then $S_n = \sum_{j=1}^{n} X_j$ conv a.s.(P).

Useful Inequalities

For positive a, b, p, $(a+b)^p \leq 2^p(a^p + b^p)$.

Markov's: if X is a nonnegative r.v. and $a > 0$, then $P(X \geq a) \leq \mathrm{E}(X)/a$.

Holder's: If $1/p + 1/q = 1$, then for measurable f, g, $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Jensen's: $\forall r \in (0, q)$, $\phi(x) = x^{q/r}$ is convex $\implies$ $\left[\mathrm{E}\,|X|^q\right]^{1/q} \geq \left[\mathrm{E}\,|X|^r\right]^{1/r}$.

Laws of Large Numbers

$\{X_n\}$ obeys the LLN if $\exists \{b_n\} \subset \mathbb{R}$ and $0 < a_n \uparrow$ such that

$$\textbf{SLLN:} \frac{S_n - b_n}{a_n} \xrightarrow{\text{a.s.}} 0 \quad \textbf{WLLN:} \frac{S_n - b_n}{a_n} \xrightarrow{\text{p}} 0.$$

Kronecker's Lemma: If $\{a_n\}, \{b_n\} \subset \mathbb{R}$ such that $0 < b_n \uparrow \infty$ and $\sum_{n=1}^{\infty} a_n/b_n$ converges, then $\frac{1}{b_n} \sum_{j=1}^{n} a_n \to 0$ as $n \to \infty$.

Cesaro's Mean Summability Theorem: If $\{x_n\} \subset \mathbb{R}$ such that $\lim_{n\to\infty} x_n = x < \infty$, then $\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} x_j = x$.

Theorem 4.14: If $\{X_n\}$ ind such that $\sum_{n=1}^{\infty} \mathrm{E}\,|X_n|^{\alpha_n} / n^{\alpha_n} < \infty$ for $\alpha_n \in [1, 2]$, then $\frac{S_n - \mathrm{E}\,S_n}{n} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mathrm{E}\,X_i) \xrightarrow{\text{a.s.}} 0$.

Marcinkiewicz-Zygmund SLLN: Let $\{X_n\}$ be iid, $S_n = \sum_{j=1}^{n} X_n$, and $p \in (0, 2)$.
(1) If $\frac{S_n - nc}{n^{1/p}} \xrightarrow{\text{a.s.}} 0$ for some $c \in \mathbb{R}$, then $\mathrm{E}\,|X_1|^p < \infty$.
(2) If $\mathrm{E}\,|X_1|^p < \infty$, then (2) holds with $c = \mathrm{E}\,X_1$ if $p \in [1, 2)$ and (2) holds for any $c \in \mathbb{R}$ if $p \in (0, 1)$.

Kolmogorov's SLLN: If $\{X_n\}$ are iid, then
$\bar{X}_n = \frac{S_n}{n} \xrightarrow{\text{a.s.}} \mathrm{E}\,X_1 \iff \mathrm{E}\,|X_1| < \infty \iff \frac{S_n - n\,\mathrm{E}\,X_1}{n} \xrightarrow{\text{a.s.}} 0$.

Useful Theorem: For any r.v. X and $r > 0$,
$\sum_{n=1}^{\infty} P(|X| > n^{1/r}) \leq \mathrm{E}\,|X|^r \leq \sum_{n=0}^{\infty} P(|X| > n^{1/r})$.

Etemaldi's SLLN: If $\{X_n\}$ are *pairwise* ind and identically distributed, then
$\bar{X}_n = \frac{S_n}{n} \xrightarrow{\text{a.s.}} \mathrm{E}\,X_1 \iff \mathrm{E}\,|X_1| < \infty$.

Theorem 4.18 (general WLLN): $\{X_n\}$ ind and put $S_n = \sum_{j=1}^{n} X_j$. If
$\sum_{j=1}^{n} P\left(\left|X_j\right| > n\right) \to 0$ and $\frac{1}{n^2} \sum_{j=1}^{n} EX_j^{(n)2} \to 0$, then $\frac{S_n - a_n}{n} \xrightarrow{\text{p}} 0$, where $a_n = \sum_{j=1}^{n} \mathrm{E}\,X_j^{(n)}$ and $X_j^{(n)} \equiv X_j I\left(\left|X_j\right| \leq n\right)$.

[Corollary] Feller's WLLN (without a 1st moment hypothesis): if $\{X_n\}$ iid with $\lim_{n\to\infty} xP(|X_1| > x) = 0$, then $\frac{S_n}{n} - \mathrm{E}\,X_1^{(n)} \xrightarrow{\text{p}} 0$.

Empirical Distributions

The **empirical cdf** of $X_1, \ldots, X_n$ is the random cdf is the proportion of observations no larger than a fixed x: $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$, $\quad x \in \mathbb{R}$.
(1) With X_i's on $(\Omega, \mathcal{F}, P)$, for each $\omega \in \Omega$, $F_n(x, \omega) = \frac{1}{n} \sum_{i=1}^{n} I(X_i(\omega) \leq x)$.
(2) $F_n(x)$ is a right-continuous, nondecreasing function of $x \in \mathbb{R}$,
(3) For any $x \in \mathbb{R}$, $F_n(x)$ is a r.v., i.e., is $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-mble:
$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i^{-1}(-\infty, x])(\omega)$.

Glivenko-Cantelli Theorem: $\{X_n\}$ iid with cdf $F(\cdot)$. Let $F_n(\cdot)$ be the empirical cdf based on $X_1, \ldots, X_n$ and define $D_n = \sup_{x\in\mathbb{R}} |F_n(x) - F(x)|$. Then, (i) D_n is a r.v. for any $n \geq 1$ and (ii) $D_n \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$.

Quantile Function: $\phi(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\} \equiv F^{-1}(u)$, $u \in (0, 1)$ which implies $F(x) \geq u \iff x \geq \phi(u)$ and $F(\phi(u)-) \leq u \leq F(\phi(u))$.

Convergence in Distribution

Let μ_n, $n \geq 0$ be probability measures on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ for some $1 \leq k < \infty$.
(1) For $\mathbf{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$, the **cdf** of μ_n is
$F_n(\mathbf{x}) = \mu_n\left((-\infty, x_1] \times \cdots \times (-\infty, x_n]\right)$.
If a random vector X_n has probability distribution μ_n [i.e., $P(X_n \in A) = \mu_n(A)$, $A \in \mathcal{B}(\mathbb{R}^k)$], then F_n is also called the cdf of X_n.
(2) A sequence of prob measures μ_n (or cdfs F_n) **converges weakly** to μ_o (to F_0), denoted as $\mu_n \Rightarrow \mu_0$ (or as $F_n \Rightarrow F_0$), if $\lim_{n\to\infty} F_n(\mathbf{x}) = F_0(\mathbf{x})$ $\forall \mathbf{x} \in C(F_0)$, where $C(F_0) = \left\{\mathbf{x} \in \mathbb{R}^k : F_0 \text{ is continuous at } \mathbf{x}\right\}$.
(3) A sequence of random vectors X_n in R^k (with distributions μ_n) **converges in distribution (law)** to a random variable X_0 (with distribution μ_0) if $\mu_n \Rightarrow \mu_0$, denoted by $X_n \xrightarrow{\text{d}} X_0$. That is, if $X_n = (X_{n,1}, \ldots, X_{n,i})$ has cdf F_n, $n \geq 0$, then

$$\lim_{n\to\infty} F_n(\mathbf{x}) = \lim_{n\to\infty} P(X_{n,1} \leq x_1, \ldots, X_{n,k} \leq x_k)$$
$$= P(X_{0,1} \leq x_1, \ldots, X_{0,k} \leq x_k) = F_0(\mathbf{x}) \quad \forall \mathbf{x} \in C(F_0).$$

Note that (1) $\mathbf{x} = (x_1, \ldots, x_k) \in C(F_0) \iff F_0(\mathbf{x}) = P(X_{0,1} < x_1, \ldots, X_{0,k} < x_k) = F_0(\mathbf{x}-)$ i.e., if also left continuous; (2) $C(F_0)^c$ is at most countable; (3) $X_n \xrightarrow{\text{p}} X_0 \implies X_n \xrightarrow{\text{d}} X_0$ but not the other direction, unless X_0 is degenerate.

Skorohod's Embedding Theorem: If μ_n, $n \geq 0$ are probability measures on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ for some $1 \leq k < \infty$ such that $\mu_n \Rightarrow \mu_0$, then $\exists$ random vectors $\{Y_n\}_{n\geq 0}$ on a *common* probability space such that Y_n has probability distribution μ_n for all $n \geq 0$ and $Y_n \xrightarrow{\text{a.s.}} Y_0$. That is, $P(Y_n \in A) = \mu_n(A)$, $A \in \mathcal{B}(\mathbb{R}^k)$, $n \geq 0$.

Continuous Mapping Theorem:
Version (a): Let μ_n, $n \geq 0$ be probability measures on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ for $1 \leq k < \infty$ and let $h : \mathbb{R}^k \to \mathbb{R}^m$ for $1 \leq m < \infty$ be a $\langle \mathcal{B}(\mathbb{R}^k), \mathcal{B}(\mathbb{R}^m) \rangle$-mble function such that $\mu_0(D_h) = 0$, where $D_n \in \mathcal{B}(\mathbb{R}^k)$ denotes the set of all points of discontinuities of the function h. If $\mu_n \Rightarrow \mu_0$, then the induced measures converge weakly: $\mu_n h^{-1} \Rightarrow \mu_0 h^{-1}$.
Version (b): Let X_n, $n \geq 0$ be R^k-valued random vectors and let measurable $h : \mathbb{R}^k \to \mathbb{R}^m$ be such that $P(X_0 \in D_h) = 0$, where D_h is as above. If $X_n \xrightarrow{\text{d}} X_0$, then $h(X_n) \xrightarrow{\text{d}} h(X_0)$.

Corollary: If X_n, Y_n, $n \geq 0$ be r.v.'s such that $(X_n, Y_n) \xrightarrow{\text{d}} (X_0, Y_0)$, then (1) $X_n + Y_n \xrightarrow{\text{d}} X_0 + Y_0$, (2)$X_n Y_n \xrightarrow{\text{d}} X_0 Y_0$, (3) $X_n / Y_n \xrightarrow{\text{d}} X_0 / Y_0$ if $P(Y_0 = 0) = 0$.

Corollary (Slutsky's Theorem): If X_n, Y_n, $n \geq 1$ be r.v.'s such that $X_n \xrightarrow{\text{d}} X$ and $Y_n \xrightarrow{\text{p}} a$ for some $a \in \mathbb{R}$, then (1) $X_n + Y_n \xrightarrow{\text{d}} X + a$, (2) $X_n Y_n \xrightarrow{\text{d}} aX$, (3) $X_n / Y_n \xrightarrow{\text{d}} X/a$ if $a \neq 0$.

Characterizations of Convergence in Distribution

For a probability measure μ on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, a set $A \in \mathcal{B}(\mathbb{R}^k)$ is called a μ-**continuity** set if $\mu(\partial A) = 0$, where $\partial A = \overline{A} \setminus \text{int}A$. E.g., $\partial(-\infty, x] = (-\infty, x] \setminus (-\infty, x) = \{x\}$.

Helly-Bray Theorem: If μ_n, $n \geq 0$ are probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then
(1) $\mu_n \Rightarrow \mu_0 \iff \mu_n(A) \to \mu_0(A)$ $\forall A \in \mathcal{B}(\mathbb{R}) \ni \mu_0(\partial A) = 0$.
(2) $\mu_n \Rightarrow \mu_0 \iff \int f d\mu_n \to \int f d\mu_0$ for all bounded cont. func. $f : \mathbb{R} \to \mathbb{R}$.
Remarks: (1) $\mu_n \Rightarrow \mu_0 \not\Longleftarrow \mu_n(A) \to \mu_0(A) \forall A \in \mathcal{B}(\mathbb{R}^k)$, (2) Holds for $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, (3) Can generalize to a metric space (S, d).

Lemma 5.8: If $\mu_n \Rightarrow \mu_0$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $f : \mathbb{R} \to \mathbb{R}$ is a bounded, Borel-mble function with $\mu_0(D_f)$ (where $D_f \in \mathcal{B}(\mathbb{R})$ is the set of discontinuity points of f), then $\int f d\mu_n \to \int f d\mu_0$ as $n \to \infty$.

A seq of prob measures $\{\mu_n\}$ on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ is **tight** if $\forall \varepsilon > 0$, $\exists M_\varepsilon > 0$ such that

$$\sup_{n\geq 1} \mu_n \left(\left\{ x \in \mathbb{R}^k : \|x\| > M_\varepsilon \right\} \right) < \varepsilon.$$

For a single prob measure on $\mathbb{R}$, given ε, we can find M_ε such that $\mu([-M_\varepsilon, M_\varepsilon]) < \varepsilon$ and note $\mu([-M_\varepsilon, M_\varepsilon]) \uparrow \mu(\mathbb{R}) = 1$.

A seq of R^k-valued random vectors $\{X_n\}$ is **tight** or **stochastically bounded** if their corresponding $\{\mu_n\}$ is tight. That is, $\forall \varepsilon > 0$, $\exists M_\varepsilon > 0$ such that

$$\sup_{n\geq 1} P(\|X_n\| > M_\varepsilon) = \sup_{n\geq 1} \mu_n \left(\left\{ x \in \mathbb{R}^k : \|x\| > M_\varepsilon \right\} \right) < \varepsilon,$$

where $\mu_n(A) = P(X_n \in A)$, $A \in \mathcal{B}(\mathbb{R}^k)$.

A sequence of random vectors $\{X_n\}$ is **uniformly integrable** if $\forall \varepsilon > 0$, $\exists t_\varepsilon > 0$ such that

$$\sup_{n\geq 1} \mathrm{E}\,\|X_n\| \, I(\|X_n\| > t_\varepsilon) = \sup_{n\geq 1} \int_{\|x\|>t_\varepsilon} \|x\| \, d\mu_n < \varepsilon,$$

where $\mu_n(A) = P(X_n \in A)$, $A \in \mathcal{B}(\mathbb{R}^k)$.

Proposition 5.9: Let $\{X_n\}$ be r.v.'s.
(1) If $X_n \xrightarrow{\text{d}} X_0$, then $\{X_n\}$ is tight.
(2) If $\{X_n\}$ is tight and $Y_n \xrightarrow{\text{p}} 0$ for X_n, Y_n defined on $(\Omega_n, \mathcal{F}_n, P_n)$, then $X_n Y_n \xrightarrow{\text{p}} 0$.
(3) But, weak convergence "almost" implies tightness - see Prokhorov's theorem.

Theorem 5.10: A sequence of r.v.'s $\{X_n\}$ (or probability measures $\{\mu_n\}$) is tight iff for any subsequence X_{n_k} of X_n there exists a further subsequence $X_{n_{k_j}}$ of X_{n_k} and a r.v. (or probability measure μ_0) such that $X_{n_{k_j}} \xrightarrow{\text{d}} X_0$ (or $\mu_{n_{k_j}} \Rightarrow \mu_0$). Note: X_0 (or μ_0) depends on the particular subsequence X_{n_k}.

Corollary: If $\{X_n\}$ is tight and all its convergent subsequences converge in distribution to the *same* r.v. X_0, then $X_n \xrightarrow{\text{d}} X_0$.

Theorem 5.12 (conv in dist + UI $\implies$ conv in mean): If $\{X_n\}$, $n \geq 1$ is UI and $X_n \xrightarrow{\text{d}} X_0$, then $\mathrm{E}\,|X_0| < \infty$ and $\mathrm{E}\,X_n \to \mathrm{E}\,X_0$.

Corollary 5.13 If $X_n \xrightarrow{\text{d}} X_0$ and $\sup_{n\geq 1} \mathrm{E}\,|X_n|^{r+\delta} < \infty$ for some integer $r \geq 1$ and real $\delta > 0$, then $\mathrm{E}\,|X_0|^r < \infty$ and $\mathrm{E}\,X_n^r \to \mathrm{E}\,X_0^r$ (recall that $\sup_{n\geq 1} \mathrm{E}\,|Z_n|^{1+\delta} \implies \{Z_n\}$ is UI).

Fréchet-Shohat Theorem: If $\lim_{n\to\infty} \mathrm{E}\,X_n^r = \beta_r \in \mathbb{R}$ for all integers $r \geq 1$ and if $\{\beta_r : r \geq 1\}$ are the moments of a *unique* r.v. X_0, then $X_n \xrightarrow{\text{d}} X_0$.

Moments uniquely determine distribution when Cardeman's condition is met, $\sum_{r=1}^{\infty} \beta_{2r}^{-1/(2r)} = \infty$, or if the MGF $M_X(t) = \mathrm{E}\,e^{tX} < \infty$ $\forall\,|t| < \varepsilon$ for some $\varepsilon > 0$. Recall: $\mathrm{E}\,X^r = \frac{d^r}{dt^r} M_X(t)\Big|_{t=0}$.

Characteristic Functions

A **complex** number is $a + ib$, where $a, b \in \mathbb{R}$ and $i = \sqrt{-1}$. If $a + bi$ and $c + di$ are complex, then their sum is $(a+b) + (c+d)i$, their product is $(ac - bd) + (ad + bc)i$, and the modulus is $|a + bi| = \sqrt{a^2 + b^2} = \sqrt{(a+bi)(a-bi)}$. IMPORTANT! For any $b \in \mathbb{R}$, $e^{bi} = \cos(b) + i\sin(b)$ and $\left|e^{bi}\right| = \sqrt{\cos^2(b) + \sin^2(b)} = 1$ and $e^{ai}e^{bi} = e^{(a+b)i}$ for $a, b \in \mathbb{R}$. ALSO, for fixed $b \in \mathbb{R}$, the function $g(t) = e^{tbi} : \mathbb{R} \to \mathbb{C}$ is infinitely differentiable in t with nth derivative $(bi)^n e^{tbi}$.

For a random vector X in $\mathbb{R}^k$, the **characteristic function** (CF) is defined as

$$\phi_X(t) = \mathrm{E}\,e^{it'X} = \mathrm{E}\cos(t'X) + i\,\mathrm{E}\sin(t'X), \quad t \in \mathbb{R}^k, i = \sqrt{-1}.$$

Note that $\phi_X(0) = 1$ and $\phi_X(t)$ is uniformly continuous on $\mathbb{R}^k$: by the BCT,

$$\sup_{t\in\mathbb{R}^k} |\phi_X(t+h) - \phi_X(t)| = \sup_{t\in\mathbb{R}^k} \left| \mathrm{E}\,e^{i(t+h)'X} - \mathrm{E}\,e^{it'X} \right|$$
$$\leq \mathrm{E}\left|e^{ih'X - 1}\right| = \int_{R^k} \left|e^{ih'x} - 1\right| d\mu_X(x) \to 0 \text{ as } |h| \to 0.$$

Theorem 5.15: If X is a r.v. with $\mathrm{E}\,|X|^r < \infty$ for some $r \geq 1$, then $\phi_X(t)$ is r-times differentiable on $\mathbb{R}$ and $\phi_X^{(r)}(t) = \mathrm{E}(iX)^r e^{itX}$, $\quad t \in \mathbb{R}$.

Riemann-Lebesgue Lemma: If the distribution of a r.v. X has a density f w.r.t. the Lebesgue measure on $\mathbb{R}$, then $\phi_X(t) \to 0$ as $|t| \to \infty$.

Levy Inversion Formula (use CF to recover dist): If X is a r.v. with CF ϕ_X, then for any $a, b \in \mathbb{R}$ with $P(X = a) = 0 = P(X = b)$,

$$P(a < X \leq b) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \phi_X(t) dt.$$

Corollary: If we also assume $\int \left|\phi_X(t)\right| dt < \infty$ (which implies $C(F) = \mathbb{R}$), then X has a pdf f w.r.t. the Lebesgue measure on $\mathbb{R}$ given by

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_X(t) dt.$$

Levy Continuity Theorem: suppose $\{X_n\}$ is a sequence of r.v.'s each with CF ϕ_{X_n}.
(1) If $X_n \xrightarrow{\text{d}} X_o$, then for any $T > 0$, $\sup_{|t|<T} \left|\phi_{X_n}(t) - \phi_{X_0}(t)\right| \to 0$ as $n \to \infty$.
(2) If $\phi_{X_n}(t) \to g(t)$ as $n \to \infty$ for all $t \in \mathbb{R}$ and $g(t)$ is continuous at zero, then $g(\cdot)$ is a CF and $X_n \xrightarrow{\text{d}} X_0$, where X_0 is the r.v. with CF $g(\cdot)$.

Corollary 5.20: $X_n \xrightarrow{\text{d}} X_0 \iff \phi_{X_n}(t) \to \phi_{X_0}(t)$ as $n \to \infty$ for all $t \in \mathbb{R}$.

Inversion Formula in $\mathbb{R}^k$: Let X be a $\mathbb{R}^k$-valued r.v. with CF $\phi_X(t)$ for $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$. Then, $\forall$ rectangle $A = (a_1, b_1] \times \cdots \times (a_k, b_k]$ with $P(X \in \partial A) = 0$,

$$P(X \in A) = \lim_{T\to\infty} \frac{1}{(2\pi)^k} \int_{-T}^{T} \cdots \int_{-T}^{T} \prod_{j=1}^{k} \frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} \phi_X(t_1, \ldots, t_k) dt_1 \cdots$$

Also, if $\int_{\mathbb{R}^k} \left|\phi_X(t_1, \ldots, t_k)\right| dt_1 \ldots dt_k < \infty$, then X has a bounded, continuous density $f_X(x)$ w.r.t. the Lebesgue measure in $\mathbb{R}^k$ given by

$$f_X(x) = \frac{1}{(2\pi)^k} \int_{\mathbb{R}^k} e^{-i\sum_{j=1}^{k} x_j t_j} \phi_X(t_1, \ldots, t_k) dt_1 \cdots dt_k, \quad x = (x_1, \ldots, x_k) \in \mathbb{R}^k.$$

Theorem 5.22: On a psp $(\Omega, \mathcal{F}, P)$, r.v.'s $X_1, \ldots, X_k$ are ind iff for all $t_1, \ldots, t_k \in \mathbb{R}$,
$\phi_{X_1,\ldots,X_k}(t_1, \ldots, t_k) \equiv \mathrm{E}\,e^{i\sum_{j=1}^{k} X_j t_j} = \prod_{j=1}^{k} \mathrm{E}\,e^{iX_j t_j} = \prod_{j=1}^{k} \phi_{X_j}(t_j)$.

Theorem 5.23: For a sequence $\{X_n\}$, $n \geq 0$ of $\mathbb{R}^k$-valued random vectors,
(1) $X_n \xrightarrow{\text{d}} X_0 \iff \phi_{X_n}(t) \to \phi_{X_0}(t)$ $\forall t \in \mathbb{R}^k$,
(2) (Cramer-Wold device) $X_n \xrightarrow{\text{d}} X_0 \iff t'X_n \xrightarrow{\text{d}} t'X_0$ $\forall t \in \mathbb{R}^k$.