Classes of Sets

A collection $\mathcal{F}$ of subsets of $\Omega \neq \emptyset$ is an **algebra** if

1. $\Omega \in \mathcal{F}$,
2. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$,
3. $A, B \in \mathcal{F} \implies A \cup B \in \mathcal{F}$ (equivalently, $A \cap B \in \mathcal{F}$).

A collection $\mathcal{F}$ of subsets of Ω is a **σ-algebra** if

1. $\mathcal{F}$ is an algebra,
2. $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

If $\mathcal{F}_\theta,\ \theta \in \Theta$ is a collection of σ-algebras on Ω, then $\mathcal{G} = \bigcap_{\theta\in\Theta} \mathcal{F}_\theta$ is a σ-algebra but $\bigcup_{\theta\in\Theta} \mathcal{F}_\theta$ may not be.

If $\mathcal{A}$ is a collection of subsets of Ω, then the σ-algebra generated by $\mathcal{A}$ is

$$\sigma\langle\mathcal{A}\rangle = \bigcap_{\substack{\mathcal{F} \text{ is } \sigma\text{-algebra} \\ \mathcal{A}\subset\mathcal{F}}} \mathcal{F}.$$

A class $\mathcal{C}$ of subsets of Ω is a **π-class** if

1. $A, B \in \mathcal{C} \implies A \cap B \in \mathcal{C}$.

A class $\mathcal{L}$ of subsets of Ω is a **λ-class** if

1. $\Omega \in \mathcal{L}$,
2. $A \in \mathcal{L} \implies A^c \in \mathcal{L}$,
3. $A_1, A_2, \dots \in \mathcal{L}$ **disjoint** $\implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{L}$.

λ-class $\not\Longrightarrow$ σ-algebra.
π-class and λ-class $\implies$ σ-algebra (but not converse).

Dynkin's $\pi-\lambda$ Theorem: On a set Ω, if $\mathcal{C}$ is a π-class and $\mathcal{L}$ is a λ-class such that $\mathcal{C} \subset \mathcal{L}$, then $\sigma\langle\mathcal{C}\rangle \subset \mathcal{L}$.

Product Spaces

If $\mathcal{F}$ is a σ-algebra on a nonempty set Ω, then $(\Omega, \mathcal{F})$ is a **measurable space**. If $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2, \dots, n$ are measurable spaces, the n-dimensional **product space** is

$$\prod_{i=1}^{n} \Omega_i = \{(\omega_1, \omega_2, \dots, \omega_n) : \omega_i \in \Omega_i,\ 1 \le i \le n\}.$$

If $A_i \subset \Omega_i$, the n-dimensional **rectangle** is

$$\prod_{i=1}^{n} A_i = \{(\omega_1, \omega_2, \dots, \omega_n) : \omega_i \in A_i,\ 1 \le i \le n\}.$$

The n-dimensional **product σ-algebra** is

$$\prod_{i=1}^{n} \mathcal{F}_i = \sigma\left\langle\left\{\prod_{i=1}^{n} A_i : A_i \in \mathcal{F}_i,\ 1 \le i \le n\right\}\right\rangle.$$

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 $\mu(B_i) < \infty,\ \forall i \ge 1$, then μ is **σ-finite**.

Extensions of Measures: If μ is a σ-finite measure on an algebra $\mathcal{F}$, then there is a *unique* extension of μ to $\sigma\langle\mathcal{F}\rangle$.

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

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

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

Countable subadditivity: If $A_1, A_2, \dots \in \mathcal{F}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$, then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} P(A_n).$$

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

$$T^{-1}(A) \equiv \{\omega \in \Omega_1 : T(\omega) \in A\} \in \mathcal{F}_1 (\text{equivalently}, T^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1).$$

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 $\mathcal{B}(\mathbb{R})$. That is,

$$P_X(A) = P(X^{-1}(A)) = P(\{\omega \in \Omega : X(\omega) \in A\}) = P(X \in A), \quad A \in \mathcal{B}(\mathbb{R}).$$

The **cumulative distribution function** (cdf) of a r.v. X is

$$F(x) = P_X((-\infty, x]) = P(X \le x), \quad x \in \mathbb{R}.$$

If F is the cdf of a r.v. X, then

1. F is right continuous: if $x_n \downarrow x_0$ and $x_n \ge x$ 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 $x \le y \implies (-\infty, x] \subset (-\infty, y]$ then $F(x) = P_X((-\infty, x]) \le P_X((-\infty, y]) = F(y)$ 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

Any extended real-valued function f defined on Ω can be decomposed into positive $f^+ = \mathbb{I}_{f\ge0} f$ and and negative $f^- = -\mathbb{I}_{f<0} f$ parts. Then, $f = f^+ - f^-$ and $|f| = f^+ + f^-$.

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 $\mu(|f| = \infty) > 0$, then $\int_\Omega |f|\, d\mu \equiv \infty$. Note $|f|$ is μ-integrable iff both f^+ and f^- are μ-integrable.

If at least one of $\int_\Omega f^+ d\mu$, $\int_\Omega f^- d\mu$ is $< \infty$, then $\int_\Omega f\, d\mu \equiv \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu$ and we say $\int_\Omega f\, d\mu$ exists.

If $f : \Omega \to \overline{\mathbb{R}}$ is nonnegative and μ-integrable, then $f < \infty$ a.e.(μ). If $f : \Omega \to \overline{\mathbb{R}}$ is nonnegative, then $\int_\Omega f\, d\mu = 0$ iff $f = 0$ a.e.(μ).

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., $f_n(\omega) \le f_{n+1}(\omega)\ \forall \omega \in \Omega$ 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

$$\int_\Omega \liminf_{n\to\infty} f_n\, d\mu \le \liminf_{n\to\infty} \int_\Omega f_n\, d\mu.$$

Dominated Convergence Theorem (DCT): Suppose (1) $g : \Omega \to \overline{\mathbb{R}}$ is a nonnegative, μ-integrable function; (2) $|f_n| \le g$ a.e.(μ) $\forall n \ge 1$; and (3) $f_n \to f$ a.e.(μ). Then, f is μ-integrable and

$$\lim_{n\to\infty} \int_\Omega |f_n - f|\, d\mu = 0 \text{ and } \lim_{n\to\infty} \int_\Omega f_n\, d\mu = \int_\Omega f\, d\mu.$$

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 $\{f_n : n \ge 1\}$ 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 $n \ge 0$, let $\nu_n(A) = \int_A f_n d\mu\ \forall A \in \mathcal{F}$ be finite measures on $\mathcal{F}$ with densities $f_n \ge 0$. If $\nu_n(\Omega) = \nu_0(\Omega) < \infty$ for all $n \ge 1$ and $f_n \to f$ a.e.(μ), then

$$\lim_{n\to\infty} \int_\Omega |f_n - f_0|\, d\mu = 0.$$

Also, the "total variation norm" tends to zero:

$$\sup_{A\in\mathcal{F}} |\nu_n(A) - \nu_0(A)| = \frac{1}{2} \int_\Omega |f_n - f_0|\, d\mu \to 0 \text{ as } n \to \infty.$$

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

$$\sup_{\lambda\in\Lambda} \int_{|f_\lambda|>t} |f_\lambda|\, d\mu \to 0 \text{ as } t \to \infty.$$

Suppose $\mathcal{A} \equiv \{f_\lambda : \lambda \in \Lambda\}$ is a collection of μ-integrable functions on a msp $(\Omega, \mathcal{F}, \mu)$. Then,

- if Λ is a finite set, then $\mathcal{A}$ is UI,
- 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,
- if $|f_\lambda| \le f$ a.e.(μ) and $\int f d\mu < \infty$, then $\mathcal{A}$ is UI,
- if $\mathcal{A}$ is UI and $\mu(\Omega) < \infty$, then $\exists M > 0$ such that $\sup\{\int |f_\lambda|\, d\mu : \lambda \in \Lambda\} \le M$,
- 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

Let $(\Omega, \mathcal{F}, P)$ be a psp and I be a set of indices.

- A collection $A_i,\ i \in I$ of sets in $\mathcal{F}$ are **ind** if $\forall i_1, i_2, \ldots, i_n \in I$ distinct indices and fixed $1 \le n < \infty$,

$$P\left(\bigcap_{j=1}^{n} A_{i_j}\right) = \prod_{j=1}^{n} P\left(A_{i_j}\right).$$

  Note the above has $2^n - n - 1$ independence 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, $\forall i_1, i_2, \ldots, i_n \in I$ distinct indices, fixed $1 \le n < \infty$ and $\forall A_{i_1}, A_{i_2}, \ldots, A_{i_n} \in \mathcal{G}_{i_n}$,

$$P\left(\bigcap_{j=1}^{n} A_{i_j}\right) = \prod_{j=1}^{n} P\left(A_{i_j}\right).$$

  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, $\forall i_1, i_2, \ldots, i_n \in I$ distinct indices, fixed $1 \le n < \infty$ and $\forall B_{i_1}, B_{i_2}, \ldots, B_{i_n} \in \mathcal{B}(\mathbb{R})$,

$$P\left(X_{i_1} \in B_{i_1}, X_{i_2} \in B_{i_2}, \ldots, X_{i_n} \in B_{i_n}\right) = \prod_{j=1}^{n} P\left(X_{i_j} \in B_{i_j}\right).$$

  In terms of distribution functions, $X_i, i \in I$ are ind iff $\forall x_1, x_2, \ldots, x_n \in \mathbb{R}$ and $\forall i_1, i_2, \ldots, i_n \in I$ distinct indices,

$$P\left(X_{i_1} \le x_1, X_{i_2} \le x_2, \ldots, X_{i_n} \le x_n\right) = \prod_{j=1}^{n} P\left(X_{i_j} \le x_j\right).$$

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 $A_1, A_2, \cdots \in \mathcal{F}$, then

$$\limsup_{n\to\infty} A_n = \overline{\lim} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n \in \mathcal{F},$$

$$\liminf_{n\to\infty} A_n = \underline{\lim} A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n \in \mathcal{F}.$$

Also,

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n,$$

$$(\text{“}A_n \text{ i.o”})^c = \left(\overline{\lim} A_n\right)^c = \underline{\lim} A_n^c = \text{“}A_n^c \text{ eventually,”}$$

$$(\text{“}A_n \text{ eventually”})^c = (\underline{\lim} A_n)^c = \overline{\lim} A_n^c = \text{“}A_n^c \text{ i.o.”}$$

Borel-Cantelli Lemma: Let $(\Omega, \mathcal{F}, P)$ be a psp and $A_1, A_2, \cdots \in \mathcal{F}$.

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 $A_1, A_2, \ldots$ 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 $X_1, X_2, \ldots$ on a psp $(\Omega, \mathcal{F}, P)$ is

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma\langle\{X_j : j \ge n\}\rangle,$$

where $\sigma\langle\{X_j : j \ge n\}\rangle = \sigma\left\langle\left\{X_j^{-1} : B \in \mathcal{B}(\mathbb{R}), j \ge n\right\}\right\rangle$ is the σ-algebra generated by $X_j, j \ge n$.

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, $\forall r \in \mathbb{R},\ T^{-1}([-\infty, r)) = \{\omega \in \Omega : T(\omega) < r\} \in \mathcal{T}$.

For fixed, arbitrary $m \ge 1$, a tail event or r.v. is only determined by $X_n, n \ge m$ (so changing finitely man r.v.'s does not effect $\mathcal{T}$).

Examples:

- $\overline{\lim} X_n$, $\underline{\lim} X_n$ are extended real-valued tail r.v.'s,
- $\{\lim X_n$ finitely exists$\}$, $\{\overline{\lim} X_n \ge r\}$, $\{\underline{\lim} X_n < r\}$ are tail events.

Kolmogorov's 0-1 Law: Tail events of a sequence $X_1, X_2, \ldots$ of ind r.v.'s have probabilities 0 or 1. That is, if $\{X_n\}_{n\ge1}$ are ind and $A \in \mathcal{T} = \bigcap_n \sigma\langle\{X_j : j \ge n\}\rangle$, then $P(A) \in \{0, 1\}$.

Corollary: For a psp $(\Omega, \mathcal{F}, P)$ and tail σ-algebra $\mathcal{T}$ defined by a sequence $X_1, X_2, \ldots$ 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 $X_1, X_2, \ldots$ on a psp $(\Omega, \mathcal{F}, P)$ **converge almost surely** to a r.v. X_0 on $(\Omega, \mathcal{F}, P)$ if

$$P\left(\left\{\lim_{n\to\infty} X_n(\omega) = X_0(\omega)\right\}\right) = 1.$$

TFAE:

- $X_n \xrightarrow{\text{a.s.}} 0$,
- $P(|X_n| > \varepsilon \text{ i.o}) = 0$ for all $\varepsilon > 0$,
- $\sup_{j\ge n} |X_j - X_0| \xrightarrow{\text{p}} 0$ as $n \to \infty$,
- $\lim_{n\to\infty} P\left(\bigcap_{j=n}^{\infty} [|X_j - X_0| \le \varepsilon]\right) = 1$ for all $\varepsilon > 0$.

A sequence of r.v.'s $X_1, X_2, \ldots$ on a psp $(\Omega, \mathcal{F}, P)$ **converge in probability** to a r.v. X_0 on $(\Omega, \mathcal{F}, P)$ if

$$\lim_{n\to\infty} P\left(|X_n - X_0| > \varepsilon\right) = 0, \quad \forall\varepsilon > 0.$$

TFAE:

- $X_n \xrightarrow{\text{p}} 0$,
- $\sup_{m\ge n} (|X_m - X_n| > \varepsilon) \to 0$ as $n \to \infty$ for all $\varepsilon > 0$,
- $\forall\{n_j\}$ 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 $X_1, X_2, \ldots$ 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 $X_1, X_2, \ldots$ of

$$\mathcal{L}_r(\Omega, \mathcal{F}, P) \equiv \left\{\text{measurable } X \in \mathbb{R} : \int_\Omega |X|^r\, dP < \infty\right\}$$

functions **converges in $\mathcal{L}_r$** to a measurable function 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 $X_1, X_2, \cdots \in \mathcal{L}_r$, then $\exists X \in \mathcal{L}_r$ such that $X_n \xrightarrow{\mathcal{L}_r} X$ iff

$$\sup_{m\ge n} \mathrm{E}\,|X_m - X_n|^r \to 0 \text{ as } n \to \infty.$$

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

Levy's Inequality: If $X_1, X_2, \ldots, X_n$ are ind r.v.'s on a psp $(\Omega, \mathcal{F}, P)$ and $S_j = \sum_{i=1}^{j} X_i, 1 \le j \le n$, then $\forall\varepsilon > 0$,

1. $P\left(\max_{1\le j\le n}\left[S_j - m(S_j - S_n)\right] \ge \varepsilon\right) \le 2P(S_n \ge \varepsilon)$,
2. $P\left(\max_{1\le j\le n}\left[S_j - m(S_j - S_n)\right] \ge \varepsilon\right) \le 2P(|S_n| \ge \varepsilon)$.

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

Khintchine-Kolmogorov Convergence Theorem: If $X_1, X_2, \ldots$ 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 $n \ge 1$ and $\sum_n \mathrm{E}(X_n^2) < \infty$, then $S_n = \sum_{i=j}^{n} X_j, n \ge 1$ 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 $X_1, X_2, \ldots$ 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 $S_n = \sum_{j=1}^{n} X_j, n \ge 1$ 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 $\sum_n P(X_n \ne Y_n) < \infty$. If $\{X_n\}$ and $\{Y_n\}$ are tail equivalent, then

- By Borel-Cantelli, $P(\overline{\lim}(X_n \ne Y_n)) = 0 \implies P(X_n = Y_n \text{ for large } n) = 1$,
- $S_n = \sum_{j=1}^{n} X_j \xrightarrow{\text{a.s.}} S \iff S_n' = \sum_{j=1}^{n} Y_j \xrightarrow{\text{a.s.}} S'$,
- 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 $X_1, X_2, \ldots, X_n$ are ind r.v.'s with $\mathrm{E}(X_i) = 0$ and $\mathrm{E}\,|X_i|^3 < \infty, 1 \le i \le n$, then $\forall n \ge 4$,

$$\sup_{x\in\mathbb{R}} \left| P\left(\frac{S_n}{\sigma_n} \le x\right) - \Phi(x) \right| \le \frac{2.75}{\sigma_n^3} \sum_{i=1}^{n} \mathrm{E}\,|X_i|^3,$$

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 cdf of a $\mathcal{N}(0, 1)$ r.v.

Kolmogorov's 3-Series Theorem: If $X_1, X_2, \ldots$ 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 $X_n^{(c)} = X_n \mathbb{I}_{|X_n|\le c}$. Then,

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

Useful Inequalities

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

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

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

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