3. Mode of Convergence

3.1. L^p-spaces and Inequalities

3.2. Tail Probabilities and Moments

3.3. Limit Sets

Definition

For a sequence $\{A_n : n \geq 1\}$ of subsets of Ω, we define the *limit superior* and the *limit inferior* by

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k \quad \text{and} \quad \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k.$$

If $\liminf_n A_n = \limsup_n A_n = A$, then we say that the limit of A_n exists and we write $\lim_n A_n = A$.

Remark

1. Note that $\omega \in \limsup_n A_n$ if and only if $\omega \in A_n$ **infinitely often(i.o)** in n:

$$\limsup_n A_n = \{\omega \in \Omega : \forall n \geq 1\ \exists k \geq n \text{ s.t. } \omega \in A_k\} = \{\omega \in \Omega : \omega \in A_n \text{ i.o.}(n)\};$$

   and $\omega \in \liminf_n A_n$ if and only if $\omega \in A_n$ for all but finitely many(**almost all**) n:

$$\liminf_n A_n = \{\omega \in \Omega : \exists n \geq 1 \text{ s.t. } \omega \in A_k\ \forall k \geq n\} = \{\omega \in \Omega : \omega \in A_n \text{ a.a.}(n)\}.$$

2. From 1. it is clear that $\liminf_n A_n \subset \limsup_n A_n$ and that

$$I_{\limsup_n A_n} = \limsup_n I_{A_n} \quad \text{and} \quad I_{\liminf_n A_n} = \liminf_n I_{A_n}$$

   - Note: $I_{\limsup_n A_n} = I_{\cap_n\{\cup_{k\geq n} A_n\}} = \wedge_n I_{\cup_{k\geq n} A_n} = \wedge_n \vee_{k\geq n} I_{A_n} = \limsup I_{A_n}$.

3. By de Morgan's law, $(\limsup_n A_n)^c = \liminf_n A_n^c$ or equivalently $(\liminf_n A_n)^c = \limsup_n A_n^c$.
   - Note: $(\limsup_n A_n)^c = (\cap_n \cup_{k\geq n} A_n)^c = \cup_n \cap_{k\geq n} A_n^c = \liminf_n A_n^c$.

4. The *limit superior* and the *limit inferior* can be expressed as monotone limits:

$$\bigcap_{k=n}^{\infty} A_n \uparrow \liminf_n A_n, \qquad \bigcup_{k=n}^{\infty} A_n \downarrow \limsup_n A_n \qquad \text{as } n \to \infty.$$

Theorem

If $(\Omega, \mathcal{F}, P)$ is a probability space, and $A_n \in \mathcal{F}$, $n \geq 1$, then

1. $P(\liminf_n A_n) \leq \liminf_n P(A_n) \leq \limsup_n P(A_n) \leq P(\limsup_n A_n)$;
2. if $A_n \to A$, then $P(A_n) \to P(A)$.

Theorem (Borel-Cantelli lemma : convergence half)

If $\sum_{k=n}^{\infty} \mu(A_k) < \infty$ for some $n \geq 1$, then $\mu(\limsup_n A_n) = 0$.

3.4. Convergence in Measure

Theorem(Almost Everywhere Convergence)

The followings are equivalent;

1. $f_n \to f$ a.e.
2. $\mu(\{\omega : |f_n(\omega) - f(\omega)| > \epsilon \text{ i.o.}(n)\}) = 0$ for all $\epsilon > 0$.
3. $\mu(\{\omega : |f_n(\omega) - f(\omega)| > 1/k \text{ i.o.}(n)\}) = 0$ for all integer $k \geq 1$.

Corollary

If for each $\epsilon > 0$, there exists an $n \geq 1$ such that $\sum_{k=n}^{\infty} \mu[|f_k - f| > \epsilon] < \infty$, then $f_n \to f$ a.e.

Definition(Convergence in Measure)

A sequence of measurable function $\{f_n : n \geq 1\}$ **converges in measure** to a measurable function f (written $f_n \xrightarrow{\mu} f$) if

$$\mu(\{\omega : |f_n(\omega) - f(\omega)| > \epsilon\}) \to 0 \quad \text{as } n \to \infty, \quad \text{for all } \epsilon > 0.$$

In special case of random variables, we say that X_n **converges in probability** to X (written $X_n \xrightarrow{\mu} X$) if

$$P(|X_n - X| > \epsilon) \to 0 \quad \text{as } n \to \infty, \quad \text{for all } \epsilon > 0.$$

Proposition

If $f_n \xrightarrow{\mu} f$ and $f_n \xrightarrow{\mu} g$, then $f = g$ a.e.

Theorem

If $f_n \xrightarrow{\mu} f$, then there exists a subsequence $\{f_{n_k} : k \geq 1\}$ for which $f_{n_k} \to f$ a.e.

Theorem(Dominated Convergence in Measure)

If $f_n \xrightarrow{\mu} f$ and if there exists an integrable function g for which $|f_n| \leq g$ a.e. for all $n \geq 1$, then f_n, and f are integrable and $\int f_n\, d\mu \to \int f\, d\mu$ as $n \to \infty$.

Definition

A sequence of measurable functions $\{f_n : n \geq 1\}$ is **Cauchy in measure** if for all $\epsilon > 0$,

$$\mu(\{\omega : |f_m(\omega) - f_n(\omega)| > \epsilon\}) \to 0 \quad \text{as } m, n \to \infty.$$

In special case of random variables, we say that X_n **Cauchy in probability** to if for all $\epsilon > 0$,

$$P(|X_m - X_n| > \epsilon) \to 0 \quad \text{as } m, n \to \infty.$$

Remark

An equivalent formation is

$$\sup_{m\geq n} \mu(\{\omega : |f_m(\omega) - f_n(\omega)| > \epsilon\}) \to 0 \quad \text{as } n \to \infty.$$

Similarly,

$$\sup_{m\geq n} P(|X_m - X_n| > \epsilon) \to 0 \quad \text{as } n \to \infty.$$

Theorem(Completeness for Convergence in Measure)

A sequence of measurable functions $\{f_n : n \geq 1\}$ is Cauchy in measure

if and only if there exists a measurable function f for which $f_n \xrightarrow{\mu} f$ as $n \to \infty$.

* Cauchy in measure $\Longleftrightarrow$ convergence in measure.

3.5. Convergence in L^p

Definition

A sequence $\{f_n : n \geq 1\}$ of L^p functions, $0 < p < \infty$ is said to **converge in** L^p to a measurable function f (written $f_n \xrightarrow{L^p} f$) if

$$\int |f_n - f|^p d\mu \to 0 \qquad \text{as } n \to \infty.$$

In special case, if X_n and X are random variables $X_n \xrightarrow{L^p} X$,

$$\int |X_n - X|^p d\mu \to 0 \qquad \text{as } n \to \infty,$$

then we sometimes say that X_n converges to X in pth mean.

For $1 \leq p < \infty$, $f_n \xrightarrow{L^p} f$ is equivalent to $||f_n - f||_p \to 0$, and this is also the definition for the case $p = \infty$,

i.e., a sequence $\{f_n, n \geq 1\}$ of L^∞ functions is said to converge in L^∞ to a measurable function f if $||f_n - f||_\infty \to 0$ as $n \to \infty$.

Proposition

If $f_n \overset{L^p}{\to} f$ form some $0 < p \le \infty$, then $f \in L^p$.

Theorem

If $f_n \overset{L^p}{\to} f$ form some $0 < p \le \infty$, then $f_n \overset{\mu}{\to} f$.

Theorem(Riesz-Fisher)

For $0 < p \le \infty$, the space L^p is complete:

a sequence of L^p function converges in L^p if and only if the sequence is Cauchy in L^p.

3.6. Convergence of Random Variables

Theorem(Borel-Cantelli Lemma: convergence half)

If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\limsup_n A_n) = 0$.

Theorem

The followings are equivalent:

1. $X_n \to X$ a.s.
2. $P(|X_n - X| > \epsilon \quad \text{i.o}(n)) = 0$ for all $\epsilon > 0$.
3. $P(|X_n - X| > 1/k \quad \text{i.o}(n)) = 0$ for all(integer) $k \ge 1$.

Corollary

If $\sum_{n=1}^{\infty} P(|X_n - X| > \epsilon) < \infty$ for all $\epsilon > 0$, then $X_n \to X$ a.s.

Theorem

If $X_n \overset{\Pr}{\to} X$, then there exists a subsequence $\{X_{n_k} : k \ge 1\}$ for which $X_{n_k} \to X$ a.s.

Theorem

A sequence of random variables $\{X_n : n \ge 1\}$ is Cauchy in probability if and only if there exists a random variable X for which $X_n \overset{\Pr}{\to} X$ as $n \to \infty$.

- Completeness for convergence in measure.

Theorem

If $X_n \overset{L^p}{\to} X$ for some $0 < p \le \infty$, then $X_n \overset{\Pr}{\to} X$.

- L^p convergence $\implies$ convergence in measure.

Theorem(Riesz-Fisher)

For $0 < p \le \infty$, the space L^p is complete: a sequence $\{X_n, n \ge 1\}$ of random variables converges in pth mean(i.e., there exists an $X \in L^p$ s.t. $||X_n - X||_p \to 0$) if and only if the sequence is Cauchy in L^p(i.e., $||X_m - X_n||_p \to 0$ as $m, n \to \infty$).

Theorem

$X_n \to X$ a.s. if and only if $\sup_{m \ge n} |X_m - X| \overset{\Pr}{\to} 0$ as $n \to \infty$.

Corollary

If $X_n \to X$ a.s., then $X_n \overset{\Pr}{\to} X$.

Theorem

$\{X_n : n \ge 1\}$ converges a.s. if and only if $\sup_{m \ge n} |X_m - X_n| \overset{\Pr}{\to} 0$ as $n \to \infty$.

Theorem

$X_n \overset{\Pr}{\to} X$ **if and only if** for every subsequence $\{X_{n_k}, k \ge 1\}$ there exists a further subsequence $\{X_{n_{k_j}}, j \ge 1\}$ such that $X_{n_{k_j}} \to X$ a.s.

Theorem(Continuous Mapping Theorem for convergence in probability: first version)

If $X_n \overset{\Pr}{\to} X$ and $f : \mathbb{R} \to \mathbb{R}$ is continuous, then $f(X_n) \overset{\Pr}{\to} f(X)$.

Lemma

Let $f : \mathbb{R} \to \mathbb{R}$. Then, the set $D_f = \{x \in \mathbb{R} : \text{ f is discontinuous at } x\}$ is a Borel set.

Theorem(Continuous Mapping Theorem for convergence in probability: first version)

Let $f : \mathbb{R} \to \mathbb{R}$ be Borel measurable, and let D_f be the set of discontinuity points of f.

If $X_n \overset{\Pr}{\to} X$ and $P(X \in D_f) = 0$, then $X_n \overset{\Pr}{\to} X$.

3.7. Uniform Integrability

Definition

A family of random variables $\{X_t, t \in T\}$ is uniformly integrable (u.i) if

$$\sup_{t \in T} \int_{|X_t| > \alpha} |X_t| \; dP \to 0 \quad \text{as} \quad \alpha \to \infty.$$

Remark

If X_t has distribution function $F_t, t \in T$,then the uniform integrability condition can be written as

$$\sup_{t \in T} \int_{(-\infty, \alpha) \cup (\alpha, \infty} |x| dF_t(x) \to 0 \quad \text{as} \quad \alpha \to \infty.$$

Example

If Y is integrable random variable and $|X_t| \le Y$ for all $t \in T$, then $\{X_t, t \in T\}$ is u.i.

Example

Suppose $P(X_n = n) = 1/n = P(X_n = 0), n \ge 1$. Then, $E(X_n) = 1$ for all $n \ge 1$. But, for any $\alpha > 0$

$$|X_n| I(|X_n| > \alpha) = \begin{cases} X_n, & \text{if } n > \alpha, \\ 0, & \text{if } n \le \alpha. \end{cases} \implies E(|X_n| I(|X_n| > \alpha)) = \begin{cases} E(X_n) = 1, & \text{if } n > \alpha, \\ 0, & \text{if } n \le \alpha. \end{cases}$$

Thus, $\sup_n E(|X_n| I(|X_n| > \alpha)) = 1$ for all $\alpha > 0$ (not u.i.).

Theorem(Crystal Ball Condition)

If for some $p > 1$, $\sup_{t \in T} E(|X_t|^p) < \infty$, then $\{X_t, t \in T\}$ is uniformly integrable.

Corollary

If for some $0 < p < \infty$,

$$\sup_{t \in T} E(|X_t|^p) < \infty,$$

then $\{|X_t|^q, t \in T\}$ is uniformly integrable for all $0 < q < p$.

Theorem(An alternative Definition of UI)

The family of random variables $\{X_t, t \in T\}$ is u.i. if and only if

1. $\sup_{t\in T} E(|X_t|) < \infty$ (integrability);
2. for every $\epsilon > 0$, there exists a $\delta > 0$ such that

$$P(A) < \delta \implies \sup_{t\in T} \int_A |X_t| dP < \epsilon.$$

Lemma

If $X_n \overset{L^p}{\to} X$ for some $0 < p < \infty$, then $E(|X_n|^p) \to E(|X|^p)$.

Lemma

$X_n \overset{L^p}{\to} X$ if and only if

1. $X_n \overset{\Pr}{\to} X$;
2. $\{|X_n|^p, n \geq 1\}$ is uniformly integrable.

Theorem(Uniform Integrability Criterion)

Suppose that $\{X_n, n \geq 1\} \subset L^p$ for some $0 < p < \infty$, and $X_n \overset{\Pr}{\to} X$. Then, the followings are equivalent:

1. $\{|X_n|^p\}$ is uniformly integrable;
2. $X_n \overset{L^p}{\to} X$;
3. $E(|X_n|^p) \to E(|X|^p) < \infty$.

Futhermore, if p is an integer, then each of these condition implies

4. $E(X_n^p) \to E(X^p)$.

4. Independence

4.1. Independent Events

Definition

Two events $A, B \in \mathcal{F}$ are **independent** if $P(A \cap B) = P(A)P(B)$.

Definition

Events $A_1, \dots A_n$ are independent if

$$P(A_{i_1} \cap \cdots \cap A_{i_m}) = P(A_{i_1}) \cdots P(A_{i_m})$$

for all $2 \leq m \leq n$ and $1 \leq i_1 < \cdots < i_m \leq n$.

Lemma

Events $A_1, \dots, A_n$ are independent if and only if

$$P(B_1 \cap \cdots B_n) = P(B_1) \cdots P(B_n)$$

for all $B_1, \dots, B_n$ where for each $i = 1, \dots, n$, either $B_i = A_i$ or $B_i = \Omega$.

Definition

Events $A_t, t \in T$ are *independent* if for every finite collection of distinct indicies $\{t_1, \dots, t_n\} \subset T$,

$$P(A_{t_1} \cap \cdots \cap A_{t_n}) = P(A_{t_1}) \cdots P(A_{t_n}).$$

Equivalently, events $A_t, t \in T$ are independent if and only if the events in every finite subcollection are independent.

Definition(Independent Classes of Events)

Classes of events $\mathcal{A}_1, \dots, \mathcal{A}_n$ are independent if for every choice of events $A_i \in \mathcal{A}_i$, $i = 1, \dots, n$, the events $A_1, \dots, A_n$ are independent.

Lemma

Classes of events $\mathcal{A}_1, \dots, \mathcal{A}_n$ are independent if and only if

$$P(B_1 \cap \cdots \cap B_n) = P(B_1) \cdots P(B_n) \tag{*}$$

for all $B_i \in \mathcal{B}_i, i = 1, \dots, n$, where $\mathcal{B}_i = \mathcal{A}_i \cup \{\Omega\}$.

Theorem

If $\mathcal{A}_1, \dots, \mathcal{A}_n$ are independent π-systems, then $\sigma(\mathcal{A}_1), \dots, \sigma(\mathcal{A}_n)$ are independent σ-fields.

Definition

Classes of events $\mathcal{A}_t, t \in T$ are independent if for every finite collection of distinct indicies, $t_1, \dots, t_n \in T$ and events $A_{t_i} \in \mathcal{A}_{t_i}, i = 1, \dots, n$, the events $A_{t_1}, \dots A_{t_n}$ are independent.

Corollary

If $\mathcal{A}_t, t \in T$ are independent π-systems, then $\sigma(\mathcal{A}_t), t \in T$ are independent σ-fields.

4.2. Independent Classes of Events

Definition

The σ-field generated by a random variable X, denoted $\sigma(X)$ is the smallest σ-field w.r.t which X is measurable (as a mapping into $(\mathbb{R}, \mathcal{R})$).

Similarly, the σ-field generated by a random vector $X = (X_1, \dots, X_k)$, again denoted $\sigma(X)$ is the smallest σ-field w.r.t which X is measurable (as a mapping into $(\mathbb{R}^k, \mathcal{R}^k)$).

Finally, the σ-field generated by an arbitrary collection of random variables $\{X_t, t \in T\}$(defined on a common probability space $(\Omega, \mathcal{F}, P)$), is the smallest σ-field w.r.t which all $X_t, t \in T$ are measurable. This σ-filed is denoted $\sigma(X_t, t \in T)$.

Theorem

Let $X = (X_1, \dots, X_k)$ be a random vector. Then

1. $\sigma(X) = \sigma(X_1, \dots, X_k) = \{X^{-1}(H) : H \in \mathcal{R}^k\}$.
2. A random variable Y is $\sigma(X)$-measurable if and only if $Y = f(X)$ for some Borel measurable function $f : \mathbb{R}^k \to \mathbb{R}$.

Proposition

For a random variable X,

$$\sigma(X) = \sigma\left(\{\omega : X(\omega) \leq x, x \in \mathbb{R}\}\right) = \sigma\left(\left\{X^{-1}((-\infty, x]), x \in \mathbb{R}\right\}\right).$$

Definition

Random variables (random vectors) $X_1, \dots X_k$ are *independent* if the σ-fields $\sigma(X_1), \dots, \sigma(X_k)$ are independent, or equivalently, if

$$P(X_1 \in H_1, \dots X_k \in H_k) = P(X_1 \in H_1) \cdots P(X_k \in H_k) \quad \text{for all } H_1, \dots, H_k \in \mathcal{R}^1.$$

Theorem

Random variables $X_1, \dots, X_k$ are independent if and only if

$$\mu = \mu_1 \times \cdots \times \mu_k, \quad \text{(product measure)},$$

or equivalently,

$$F(x) = F_1(x_1) \cdots F_k(x_k) \quad \text{for all } x = (x_1, \dots, x_k) \in \mathbb{R}^k.$$

Theorem

If $X_1, \dots, X_k$ are independent random variables and $g_1, \dots, g_k$ are Borel measurable functions, then $g_1(X_1), \dots, g_k(X_k)$ are independent random variables.

Theorem

If X and Y are independent random variables, either both nonnegative or both integrable, then

$$E(XY) = E(X)E(Y).$$

Theorem

Suppose that X and Y are independent random vectors (k and m dimensional, respectively) with respective distributions P_X and P_Y. Let $g : \mathbb{R}^{k+m} \to \mathbb{R}$ be a Borel measurable function, and let $A \in \mathcal{R}^m$. if either g is nonnegative, or $g(X,Y)$ is integrable, then

$$E[g(X,Y)I_A(Y)] = \int_A E[g(X,y)]dP_Y(y).$$

4.3. Convolution

Definition

If μ and ν are probability measures on $(\mathbb{R}, \mathcal{R})$, then the measure $\mu * \nu$ defined by

$$(\mu * \nu)(H) = \int_{-\infty}^{\infty} \nu(H - x)\ d\mu(x), \quad H \in \mathcal{R},$$

is called the **convolution** of μ and ν.

If F and G are the distribution functions corresponding to μ and ν, respectively, then the distribution function corresponding to $\mu \times \nu$ is given by

$$(F * G)(y) = (\mu * \nu)((-\infty, y]) = \int_{-\infty}^{\infty} \nu((-\infty, y] - x)d\mu(x) = \int_{-\infty}^{\infty} \nu((-\infty, y - x])d\mu(x) = \int_{-\infty}^{\infty} G(y-x)d\mu(x) = \int_{-\infty}^{\infty} G(y-x)dF(x) \implies (F * G)(y) = \int_{-\infty}^{\infty} G(y-x)dF(x)$$

4.4. Borel-Cantelli Lemma

Lemma (Borel-Cantelli: convergence half)

Let $A_n, n \geq 1$ be a sequence of events in a probability space $(\Omega, \mathcal{F}, P)$.

If $\sum_{n=1} P(A_n) < \infty$, then $P(A_n \ \ \text{i.o}(n)) = 0$.

Lemma (Borel-Cantelli: divergence half)

Let $A_n, n \geq 1$ be a sequence of **independent** events.

If $\sum_{n=1} P(A_n) = \infty$, then $P(A_n \ \ \text{i.o}(n)) = 1$.

Corollary(Borel's Zero-One Law)

Let $A_n, n \geq 1$ be independent events. Then, if

$$\sum_{n=1}^{\infty} P(A_n) < \infty \quad \text{or} \quad \sum_{n=1}^{\infty} P(A_n) = \infty,$$

then,

$$P(A_n \ \ \text{i.o}(n)) = 0 \ \text{ or } \ 1, \ \text{ respectively.}$$

4.5. Kolmogorov's Zero-One Law

Definition(Tail Events)

If $\{X_n, n \geq 1\}$ is a sequence of random variables on $(\Omega, \mathcal{F}, P)$, then the **tail σ-field** determined by $\{X_n\}$ is given by

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots).$$

If $A \in \mathcal{T}$, then A is called a **tail event**.

Example

$\{\omega \in \Omega : \sum_{n=1}^{\infty} X_n(\omega) \text{ converges } \} \in \mathcal{T}$, because

$$\left\{\omega \in \Omega : \sum_{n=1}^{\infty} X_n(\omega) \text{ converges }\right\} = \left\{\omega \in \Omega : \sum_{n=m}^{\infty} X_n(\omega) \text{ converges}\right\} \in \sigma(X_m, X_{m+1}, \ldots)\ \forall m \geq 1 \implies \left\{\omega \in \Omega : \sum_{n=1}^{\infty} X_n(\omega) \text{ converges }\right\} \in \bigcap_{m=1}^{\infty} \sigma(X_m, X_{m+1}, \ldots).$$

Theorem(Kolmogorov's Zero-One Law)

If $\{X_n, n \geq 1\}$ is a sequence of independent random variables and A is a tail event, then either $P(A) = 0$ or $P(A) = 1$

5. Random Series, Weak and Strong Laws

5.1. Convergence of Random Series

Theorem(Kolmogorov's Maximal Inequality)

Suppose that $X_1, \ldots, X_n$ are **independent** random variables with **mean 0**, and let $S_n = \sum_{i=1}^{n} X_i, n \geq 1$. Then for any $\alpha > 0$,

$$P\left(\max_{1 \leq j \leq n} \{|S_j| \geq \alpha\}\right) \leq \frac{1}{\alpha^2} \text{Var}(S_n).$$

Theorem(Etemadi's maximal inequality)

If $X_1, \ldots, X_n$ are **independent** random variables, then for any $\alpha > 0$,

$$P(\max_{1 \leq j \leq n} |S_j| \geq 3\alpha) \leq 3 \max_{1 \leq j \leq n} P(|S_j| \geq \alpha)$$

Theorem(Kolmogorov's Convergence Criterion)

If $X_1, X_2, \ldots$ are **independent mean 0** random variables with $\sum_{n=1}^{\infty} \text{Var}(X_n) < \infty$, then $\sum_{n=1}^{\infty} X_n$ converges a.s. and in L^2. Moreover, $E(\sum_{n=1}^{\infty} X_n) = 0$ and $\text{Var}(\sum_{n=1}^{\infty} X_n) = \sum_{n=1}^{\infty} \text{Var}(X_n)$.

Theorem(Levy's Theorem)

If $X_1, X_2, \ldots$ are **independent** random variables, then as $n \to \infty$,

$S_n \to S_\infty$ a.s. ($\exists$ a r.v. S_∞) $\iff S_n \xrightarrow{\text{Pr}} S_\infty$.

Corollary

If $X_1, X_2, \ldots$ are **independent, mean 0, uniformly bounded** random variables, then

$\sum_{n=1}^{\infty} X_n$ converges a.s. $\iff \sum_{n=1}^{\infty} \text{Var}(X_n) < \infty$.

Lemma

If $X_1, X_2, \ldots,$ and $X_1^*, X_2^* \ldots$ are sequensces of random variables with identical finite dimensional distributions, i.e., with

$$(X_1, \ldots, X_n) \sim (X_1^*, \ldots, X_n^*) \quad \forall\ n \geq 1,$$

then X_n converges a.s. $\iff X_n^*$ converges a.s.

Lemma

For **independent, uniformly bounded** random variables $X_n, n \geq 1$,

if $\sum_{n=1}^{\infty} X_n$ converges a.s., then $\sum_{n=1}^{\infty} E(X_n)$ converges.

5.2. Strong Laws of Large Numbers

Definition(Tail Equivalence)

Two sequences of random variables $\{X_n, n \geq 1\}$ and $\{X_n^*, n \geq 1\}$ are **tail equivalent** if

$$\sum_{n=1}^{\infty} P(X_n \neq X_n^*) < \infty.$$

Remark

If $\{X_n, n \geq 1\}$ and $\{X_n^*, n \geq 1\}$ are tail equivalent, then by the convergence-half of the Borel-Cantelli lemma,

$$P(X_n \neq X_n^* \ \text{ i.o}(n)) = 0,$$

i.e.,

$$P(X_n = X_n^* \ \text{ a.a}(n)) = 1,$$

i.e., for almost all ω, there exists N_ω s.t. $X_n(\omega) = X_n^*(\omega)$ for all $n \geq N(\omega)$. From this, it follows that for tail equivalent sequences,

1. $\sum_{n=1}^{\infty}(X_n - X_n^*)$ converges a.s.
2. $\sum_{n=1}^{\infty} X_n$ converges a.s. $\iff \sum_{n=1}^{\infty} X_n^*$ converges a.s.
3. If $a_n \to \infty$, then

$$\frac{1}{a_n}\sum_{i=1}^{n} X_i \text{ converges a.s.} \iff \frac{1}{a_n}\sum_{i=1}^{n} X_i^* \text{ converges a.s.}$$

Theorem(Kolmogorov's Three Series Theorem)

Suppose that $X_1, X_2, \ldots$ are **independent** random variables. Then,

$\sum_{n=1}^{\infty} X_n$ converges a.s. $\iff \exists\, c > 0$ s.t. the following holds:

1. $\sum_{n=1}^{\infty} P(|X_n| > c) < \infty$;
2. $\sum_{n=1}^{\infty} E\left(X_n I_{\{|X_n| \leq c\}}\right)$ converges;
3. $\sum_{n=1}^{\infty} \text{Var}\left(X_n I_{\{|X_n| \leq c\}}\right) < \infty$.

Furthermore $\sum_{n=1}^{\infty} X_n$ converges a.s. $\implies$ 1,2,3 hold for all $c > 0$(not exists, for all).

Theorem(Marcinkiewicz-Zygmund Convergence Theorem)

Suppose that $0 < p < 2$, and let $X_1, X_2, \ldots$ be **i.i.d** L^p random variable. Define

$$Y_n = n^{-\frac{1}{p}} X_n I_{\left\{n^{-1/p}|X_n| \leq 1\right\}}, \quad n \geq 1.$$

Then,

$$\sum_{n=1}^{\infty}\left[n^{-1/p}X_n - E(Y_n)\right] \quad \text{converges a.s.}$$

Moreover, if either

1. $0 < p < 1$, or
2. $1 < p < 2$ and $E(X_1) = 0$,

then

$$\sum_{n=1}^{\infty} n^{-1/p} X_n \quad \text{converges a.s.}$$

Theorem (Cesaro Averages)

If $x_n \in \mathbb{R}, n \geq 0$ and $x_n \to x_0$ as $n \to \infty$, then

$$\frac{1}{n}\sum_{k=1}^{n} x_k \to x_0 \quad \text{as} \quad n \to \infty.$$

Lemma (Kronecker's Lemma)

For real sequences $\{x_n, n \geq 1\}$ and $\{a_n, n \geq 1\}$, with $0 < a_n \uparrow \infty$,

if

$$\sum_{k=1}^{\infty} \frac{x_k}{a_k} \quad \text{converges (to a finite limit),}$$

then

$$\frac{1}{a_n}\sum_{k=1}^{n} x_k \to 0 \quad \text{as} \quad n \to \infty.$$

Theorem (Marcinkiewicz-Zygmund Strong Law)

Suppose that $X_1, X_2, \ldots$ are **i.i.d** random variables, and $0 < p < 2$. Then, for some $c > 0$,

$$\frac{S_n - nc}{n^{1/p}} \to 0 \quad \text{a.s.} \iff E(|X_1|^p) < \infty.$$

Moreover,

1. if $1 \leq p < 2$, then, necessarily $c = E(X_1)$;
2. if $0 < p < 1$, c is arbitrary (could be 0).

Corollary (Classical Strong Law of Large Numbers)

Let $X_1, X_2, \ldots,$ be **i.i.d** random variables. Then,

1. $S_n/n \to E(X_1)$ a.s. $\iff X_1 \in L^1$.
2. If E(X_1) exists(possibly $-\infty, \infty$), then $S_n/n \to E(X_1)$ a.s.

5.3. The Glivenko-Cantelli Theorem

Theorem (Glivenko-Cantelli Theorem)

Let $X_1, X_2, \ldots$ be **i.i.d(finite-valued)** r.vs with common distribution function F, and let F_n be the empirical distribution function based on $X_1, \ldots, X_n$, i.e.,

$$F_{n,\omega}(x) = \frac{1}{n}\sum_{k=1}^{n} I_{\{X_k \leq x\}}(\omega), \quad -\infty < x < \infty, \ \omega \in \Omega.$$

Then,

$$\sup_{-\infty < x < \infty} |F_n(x) - F(x)| \to 0 \quad \text{a.s.,}$$

i.e., for almost all ω, $F_{n,\omega}$ converges uniformly to F as $n \to \infty$.

5.4. Weak Laws of Large Numbers

Theorem(ChebyShev Weak Law of Large Numbers)

Let $S_n = \sum_{i=1}^{n} X_i, n \geq 1$, where $X_1, X_2, \ldots$ are L^2 random variables **(not necessarily i.i.d)**. If $b_n, n \geq 1$, are positive constants satisfying $\text{Var}(S_n) = o(b_n^2)$, then

$$\frac{S_n - E(S_n)}{b_n} \xrightarrow{L^2} 0 \quad \text{and consequently} \quad \frac{S_n - E(S_n)}{b_n} \xrightarrow{\text{Pr}} 0.$$

Corollary (L^2 Weak Law)

Let $X_1, X_2, \ldots$ are **uncorrelated** L^2 random variables, with $E(X_n) = \mu$, and $\text{Var}(X_n) \leq C < \infty$ for all $n \geq 1$. Then

$$\frac{S_n}{n} \xrightarrow{L^2} \mu \quad \text{and consequently} \quad \frac{S_n}{n} \xrightarrow{\text{Pr}} \mu.$$

Theorem (Chevyshev WLLN for Random Arrays)

Suppose $X_{n,i}$, $1 \leq i \leq m_n, n \geq 1$, are L^2 random variables(defined on the samme probability space), and let $S_n = \sum_{i=1}^{m_n} X_{n,i}$, $n \geq 1$. If for some sequence of positive constants (b_n),

$$\frac{\text{Var}(S_n)}{b_n^2} \to 0 \quad \text{as } n \to \infty,$$

then

$$\frac{S_n - E(S_n)}{b_n} \xrightarrow{L^2} 0 \quad \text{and consequently} \quad \frac{S_n - E(S_n)}{b_n} \xrightarrow{\text{Pr}} 0.$$

Theorem (A Weak Law for Triangular Arrays)

For each $n \geq 1$, suppose that $X_{n,i}, 1 \leq i \leq m_n$ are **independent** random variables, and let $S_n = \sum_{i=1}^{m_n} X_{n,i}$. Suppose further that $0 < b_n \to \infty$ and define

$$X_{n,i}^* = X_{n,i} I_{\{|X_{n,i}| \leq b_n\}}, \quad \text{and} \quad a_n = \sum_{i=1}^{m_n} E(X_{n,i}^*), \quad n \geq 1.$$

If both

1. $\sum_{i=1}^{m_n} P(|X_{n,i}| > b_n) \to 0$ as $n \to \infty$, and
2. $\frac{1}{b_n^2} \sum_{i=1}^{m_n} E({X_{n,i}^*}^2) \to 0$ as $n \to \infty$,

then

$$\frac{S_n - a_n}{b_n} \overset{\text{Pr}}{\to} 0 \quad \text{as } n \to \infty.$$

Theorem (Feller's Weak Law of Large Numbers)

Let $X_1, X_2, \ldots$ be **i.i.d** random variables with

$$nP(|X_1| > n) \to 0 \quad \text{as } n \to \infty.$$

Let $\mu_n = E(X_1 I_{\{|X_1| \leq n\}})$. Then,

$$\frac{S_n}{n} - \mu_n \overset{\text{Pr}}{\to} 0.$$

6. Weak Convergence

6.1. Weak Convergence of Probability Distributions

Definition (Weak Convergence)

A sequence of distribution functions $\{F_n, n \geq 1\}$ **converges weakly** to a distribution function F if

$$F_n(x) \to F(x)$$

for all continuity point x of F (written $F_x \rightsquigarrow F$).

If μ_n and μ are probability measures on $(\mathbb{R}, \mathcal{R})$, with distribution functions F_n and F, respectively, then μ_n *converges weakly* to μ if $F_n \rightsquigarrow F$ (written $\mu_n \rightsquigarrow \mu$).

Finally suppose X_n and X are r.vs with distribution functions F_n and F, respectively. If $F_n \rightsquigarrow F$, then we say X_n **converges in distribution** to X(written $X_n \rightsquigarrow X$).

Proposition

If $F_n \rightsquigarrow F$ and $F_n \rightsquigarrow G$, then $F = G$.

Proposition

$X_n \rightsquigarrow X \not\Rightarrow X_n \overset{\text{Pr}}{\to} X$.

Theorem

If $X_n \rightsquigarrow c$ for some real constant c, then $X_n \overset{\text{Pr}}{\rightsquigarrow} c$.

Theorem (Skorohod Representation Theorem)

Suppose that $\mu_n, n \geq 1$ and μ are probability measures on $(\mathbb{R}, \mathcal{R})$ with $\mu_n \rightsquigarrow \mu$.

Then, there exists a probability space $(\Omega, \mathcal{F}, P)$ and random variables $Y_n, n \geq 1$ and Y, defined on $(\Omega, \mathcal{F}, P)$ such that Y_n has distribution μ_n for all $n \geq 1$, Y has distribution μ, and $Y_n \to Y$ a.s.

Theorem (General Definition of Weak Convergence)

Let $C_b(\mathbb{R})$ denote the space of **bounded, continuous** real-valued functions on $\mathbb{R}$.

For real-valued random variables $X_n, n \geq 1$, and X,

$$X_n \rightsquigarrow X \iff E[g(X_n)] \to E[g(X)] \quad \forall\, g \in C_b(\mathbb{R}).$$

Equivalently,

$$\mu_n \rightsquigarrow \mu \iff \int g\, d\mu_n \to \int g\, d\mu \quad \forall\, g \in C_b(\mathbb{R}).$$

For probability distribution functions on $(\mathbb{R}, \mathcal{R})$, and for distribution functions,

$$F_n \rightsquigarrow F \iff \int g\, dF_n \to \int g\, dF \quad \forall\, g \in C_b(\mathbb{R}).$$

Corollary

$X_n \overset{\text{Pr}}{\to} X \implies X_n \rightsquigarrow X$.

Theorem (Continuous Mapping Theorem for Convergence in Distribution)

If $X_n \rightsquigarrow X$ and if $g : \mathbb{R} \to \mathbb{R}$ is Borel measurable with $P(X \in D_g) = 0$, then $g(X_n) \rightsquigarrow g(X)$.

6.2. Tightness

Theorem (Portmanteau Theorem)

For probability measures $\mu_n, n \geq 1$ and μ on $(\mathbb{R}, \mathcal{R})$, the following are equivalent:

1. $\mu_n \rightsquigarrow \mu$;
2. $\liminf_n \mu_n(B) \geq \mu(B)$ for all open $B \in \mathbb{R}$;
3. $\limsup_n \mu_n(C) \leq \mu(C)$ for all closed $C \in \mathbb{R}$;
4. $\lim_n \mu_n(A) = \mu(A)$ for all μ-continuity sets, i.e., for all $A \in \mathcal{R}$ with $\mu(\partial A) = 0$.

Theorem (Helly's selection theorem)

For any sequence of distribution functions $\{F_n, n \geq 1\}$, there exists a subsequence, $\{F_{n_k}, k \geq 1\}$, and a *subdistribution* function F (nondecreasing, right-continuous, $0 \leq F(x) \leq 1$ for all $x \in \mathbb{R}$) such that

$$F_{n_k}(x) \to F(x) \quad \text{for all } x \in C_F.$$

Definition (Tightness)

A sequence of probability measures $\{\mu_n, n \geq 1\}$ on $(\mathbb{R}, \mathcal{R})$ is **tight** if for every $\epsilon > 0$, there exists $\mathcal{M} = \mathcal{M}_\epsilon > 0$ s.t.

$$\mu_n([-\mathcal{M}, \mathcal{M}]) > 1 - \epsilon \quad \text{for all } n \geq 1.$$

Proposition

A finite collection of probability measures on $(\mathbb{R}, \mathcal{R})$ is tight.

Theorem (Prohorov's theorem)

A sequence of probability measures $\{\mu_n, n \geq 1\}$ is tight

$\iff$ for every subsequence $\{\mu_{n_k}, k \geq 1\}$, $\exists$ a further subsequence $\{\mu_{n_{kj}}, j \geq 1\}$ and a probability measure μ s.t. $\mu_{n_{kj}} \rightsquigarrow \mu$.

Corollary

the sequence of measures $\{\mu_n, n \geq 1\}$ converges weakly $\implies$ it is tight.

Corollary

If $\{\mu_n, n \geq 1\}$ is tight, and if each weakly convergent subsequence converges to the same probability measure, then $\mu_n \rightsquigarrow \mu$.

6.3. Slutsky's Theorem

Definition

A sequence of random variables $\{X_n, n \geq 1\}$ is **bounded in probability** if their associated distributions are tignt, i.e., for any $\epsilon > 0$, there exists a constant $\mathcal{M} > 0$ s.t.

$$P(|X_n| \leq \mathcal{M}) > 1 - \epsilon \quad \text{for all } n \geq 1.$$

Lemma

If $\{X_n\}$ is bounded in probability and $Y_n \rightsquigarrow 0$, then $X_nY_n \rightsquigarrow 0$.

Definition

If $X_n \rightsquigarrow 0$(or equivalently, $X_n \xrightarrow{\Pr} 0$), then we write $X_n = o_p(1)$.

If $\{X_n, n \geq 1\}$ is bounded in probability, then we write $X_n = O_p(1)$.

Lemma (Converging Together Lemma)

$X_n \rightsquigarrow X$ and $Y_n \rightsquigarrow Y \implies Y_n \rightsquigarrow X$

Theorem (Slutsky's Theorem)

If for all $n \geq 1$, X_n, A_n, and B_n are random variables defined on the same probability space, with $X_n \rightsquigarrow X$, $A_n \rightsquigarrow a$, and $B_n \rightsquigarrow b$, $a, b \in \mathbb{R}$, then

$$A_nX_n + B_n \rightsquigarrow aX + b.$$

6.4. Convergence of Moments

Theorem

$X_n \rightsquigarrow X \implies E[|X|] \leq \liminf_n E[|X_n|]$.

Theorem

If $X_n \rightsquigarrow X$ and $\{X_n, n \geq 1\}$ is **uniformly integrable**, then X is integrable and $E[X_n] \to E[X]$.

Corollary

If $X_n \rightsquigarrow X$ and if $\sup_{n\geq 1} E\left[|X_n|^{k+\epsilon}\right] < \infty$ for some integer $k \geq 1$, and some $\epsilon > 0$, then X^k is integrable and $E[X_n^k] \to E[X^k]$.

6.5. Total Variation: Scheffe's Theorem

Remark

For probability measures μ_n and μ on $(\mathbb{R}, \mathcal{R})$, below conditions are ($3 \implies 2 \implies 1$) (progressively stronger forms);

1. $\mu_n \rightsquigarrow \mu$.
2. $\mu(A) \to \mu(A)$ for all $A \in \mathcal{R}$.
3. $\sup_{A\in\mathcal{R}} |\mu_n(A) - \mu(A)| \to 0$.

Lemma

If μ and ν are probability measures, with densities f and g w.r.t some dominating measure γ, then

$$\sup_{A\in\mathcal{F}} |\mu(A) - \nu(A)| = \frac{1}{2}\int |f - g| d\gamma.$$

Theorem (Scheffe's Theorem)

$\mu_n, n \geq 1$ and μ are probability measures having densities f_n and f, respectively, w.r.t some dominating measure γ. If

$$f_n \to f \quad \gamma\text{-a.e.} \quad \text{as } n \to \infty,$$

then μ_n converges to μ in total variation norm, i.e.,

$$\sup_{A\in\mathcal{R}} |\mu_n(A) - \mu(A)| \to 0.$$

7. Characteristic Functions

7.1. Integrals of Complex-Valued Functions

Remark

Let $\mathbb{C}$ represent the field of complex numbers. If $z = x + iy \in \mathbb{C}$, where $x, y \in \mathbb{R}$, then $x = \mathrm{Real}(z)$ and $y = \mathrm{Image}(z)$ are called the real and imaginary parts of z, and $|z| = \sqrt{x^2 + y^2}$ and $\bar{z} = x - iy$ are called the modulus and the complex conjucate of z, respectively.

Recall that in its polar representation, z is written in the form

$$z = |z|e^{i\theta} = |z|(\cos\theta + i\sin\theta),$$

and the angle $\theta = \arg z$, measured in radians, is called the argument of z. Of course, θ is determined by the equations $\cos\theta = x/|z|$ and $\sin\theta = y/|z|$.

Remark

For a measure space $(\Omega, \mathcal{F}, \mu)$, and a function $f : \Omega \to \mathbb{C}$, with $f = g + ih$, where $g, h : \Omega \to \mathbb{R}$, the function f is (Borel) measurable if and only if both its real and imagenary parts g and h are measurable, and we define

$$\int f \, d\mu = \int g \, d\mu + i \int h \, d\mu,$$

as long as both integrals on the right hand side exists. Similarly we say that f is **integrable** if

$$\int |f| \, d\mu < \infty,$$

and since $|f| = (g^2 + h^2)^{1/2} \leq |g| + |h|$ while both $|g| \leq |f|$ and $|h| \leq |f|$, we see that f is integrable if and only if both its real and imaginary parts are integrable.

Remark (Modulus Inequality)

If f is integrable w.r.t μ, then for any $\theta \in \mathbb{R}$, we have

$$\left|\int f \, d\mu\right| \leq \int |f| \, d\mu.$$

7.2. Definition and Derivatives of the Characteristic Function

Definition

The **characteristic function** of a probability measure μ on $(\mathbb{R}, \mathcal{R})$ is the function $\phi : \mathbb{R} \to \mathbb{C}$ defined by

$$\phi(t) = \int e^{itx} \, d\mu(x) = \int \cos(tx) \, d\mu(x) + i \int \sin(tx) \, d\mu(x).$$

If F is the distribution function corresponding to μ, or if X is a random variable with distribution μ, then we say that F or X has characteristic function ϕ. Of course we then may write

$$\phi(t) = \int e^{itx} dF(x) = \int \cos(tx) \, dF(x) + i \int \sin(tx) \, dF(x),$$

or

$$\phi(t) = E\left[e^{itX}\right] = E\left[\cos(tX)\right] + iE\left[\sin(tX)\right].$$

Proposition (Elementary Properties of the Characterisric Function)

Let ϕ be the characteristic function of a random variable X. Then,

1. $\phi(0) = 1$;
2. $|\phi(t)| \leq 1$ (Modulus Inequality);
3. ϕ is uniformly continuous;
4. $\mathrm{Re}\{\phi(t)\} = E[\cos(tX)]$ is an even function of t, and $\mathrm{Im}\{\phi(t)\} = E[\sin(tX)]$ is odd;
5. $\phi(-t) = \overline{\phi(t)}$;
6. For $a, b \in \mathbb{R}$, $\phi_{aX+b}(t) = e^{ibt}\phi_X(at)$, in particular, $\phi_{-X}(t) = \phi_X(-t) = \overline{\phi_X(t)}$.
7. ϕ_X is a real valued function $\iff$ X is symmetrically distributed about the origin i.e., $-X \sim X$.
8. If X and Y are independent, then $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$.

Lemma

For any $x \in \mathbb{R}$,

$\left|e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!}\right| \le \min\left\{\frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!}\right\}$.

Thus, we have at $n = 0$, and $n = 1$,

1. $\left|e^{ix} - 1\right| \le \min\{|x|, 2\}$,
2. $\left|e^{ix} - (1 + ix)\right| \le \min\left\{\frac{|x|^2}{2}, 2|x|\right\}$.

Lemma

If $E[|X|^n] < \infty$ for some integer $n \ge 1$, then $\phi(t) = E[e^{itX}]$ satisfies

$$\left|\phi(t) - \sum_{k=0}^{n} \frac{(it)^k}{k!} E(X^k)\right| \le E\left[\min\left\{\frac{|tX|^{n+1}}{(n+1)!}, \frac{2|tX|^n}{n!}\right\}\right].$$

Theorem

If $E[|X|^n] < \infty$ for some integer $n \ge 1$, then

$$\phi(t) = \sum_{k=0}^{n} \frac{(it)^k}{k!} E(X^k) + R_n(t),$$

where $R_n(t) = o(|t|^n)$ as $t \to 0$.

7.3. Fourier Inversion Theorem

Theorem (Fourier Inversion Theorem)

Suppose that μ is a probability measure with probability distribution function F and characteristic function ϕ. If

$$\int_{-\infty}^{\infty} |\phi(t)| dt < \infty$$

then μ has a bounded, uniformly continuous density given by

$$f(x) = F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t) dt.$$

7.4. Levy Continuity Theorem

Lemma

If μ is a probability measure with characteristic function ϕ, then for any $u > 0$,

$$\frac{1}{u} \int_{-u}^{u} [1 - \phi(t)] dt \ge \mu(\{x : |x| \ge 2/u\}).$$

Theorem (Levy Continuity Theorem)

Let $\mu_n, n \ge 1$ be probability measures with characterisric functions ϕ_n. Then,

1. $\mu_n \rightsquigarrow \mu \iff \phi_n(t) \to \phi(t)$ for all $t \in \mathbb{R}$ where $\phi(t)$ is char.func of μ.
2. $\phi_n(t) \to g(t)$ for all $t \in \mathbb{R}$ and g is continuous at $t = 0 \implies g$ is the characterisric function of μ, and $\mu_n \rightsquigarrow \mu$.

Corollary

Suppose that $g(t) = \lim_{n\to\infty} \phi_n(t)$ exists for each $t \in \mathbb{R}$, and that $\{\mu_n, n \ge 1\}$ is tight. Then there exists a prob measure μ s.t. $\mu_n \rightsquigarrow \mu$ and μ has characteristic function g.

8. The Central Limit Theorem

8.1. The Central Limit Theorem

Theorem (Classical CLT for i.i.d Summands)

If $X_n, n \ge 1$ are **i.i.d** random variables with $E(X_1) = c$ and $\mathrm{Var}(X_1) = \sigma^2$, where $0 < \sigma^2 < \infty$, then

$$\frac{S_n - nc}{\sigma\sqrt{n}} \rightsquigarrow Z \sim N(0, 1).$$

Theorem (Linderberg-Feller Central Limit Theorem)

Suppose that for $n \ge 1$, $X_{n,1}, \dots, X_{n,r_n}$ are **independent** random variables with

$$E(X_{n,k}) = 0 \quad \text{and} \quad \sigma^2_{n,k} = E(X^2_{n,k}) < \infty, \quad k = 1, \dots, r_n.$$

Let

$$S_n = X_{n,1} + \cdots + X_{n,r_n} \quad \text{and} \quad s^2_n = \sigma^2_{n,1} + \cdots + \sigma^2_{n,r_n},$$

and assume that $s^2_n > 0$ for n large. Then

1. $\frac{S_n}{s_n} \rightsquigarrow Z \sim (0, 1)$ as $n \to \infty$;
2. $\frac{\max_{1\le k\le r_n} \sigma^2_{n,k}}{s^2_n} \to 0$ as $n \to \infty$,

are necessary and sufficient that the **Linderberg condition** hold:

$$\lim_{n\to\infty} \frac{1}{s^2_n} \sum_{k=1}^{r_n} \int_{\{|X_{n,k}| \ge \epsilon s_n\}} X_{n,k}\, dP = 0 \quad \text{for all} \quad \epsilon > 0.$$

Proposition

If $X_n, n \ge 1$ are independent random variables with

$$E(X_n) = 0 \quad \text{and} \quad \mathrm{Var}(X_n) = \sigma^2_n, \quad n \ge 1,$$

and $s^2_n = \sum_{k=1}^{n} \sigma^2_k$, then the Linderberg condition

$$\lim_{n\to\infty} \frac{1}{s^2_n} \sum_{k=1}^{r_n} \int_{\{|X_{n,k}| \ge \epsilon s_n\}} X_{n,k}\, dP = 0 \quad \text{for all} \quad \epsilon > 0,$$

is equivalent to

$$\lim_{n\to\infty} \frac{1}{s^2_n} \sum_{k=1}^{r_n} \int_{\{|X_{n,k}| \ge \epsilon s_k\}} X_{n,k}\, dP = 0 \quad \text{for all} \quad \epsilon > 0.$$

Examples

Let $1 \le \alpha < \infty$ and let $X_n, n \ge 1$ be **independent** random variables with

$$P(X_n = n^\alpha) = P(X_n = -n^\alpha) = \frac{1}{6} n^{-2(\alpha-1)},$$

and

$$P(X_n = 0) = 1 - \frac{1}{3} n^{-2(\alpha-1)}.$$

9. Conditioning

9.1. Conditional Expectations

9.2. Properties of Conditional Expectation

9.3. Conditional Expectations as Projections