1 Measures

1.1 Classes of Sets

I. **Algebra, σ-algebra.** A class $\mathcal{F}$ is an algebra if (1) $\Omega \in \mathcal{F}$, (2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$, (3) $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$, and a σ-algebra if (4) $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$ is also true.

Note: even though $\mathcal{F}_1$ and $\mathcal{F}_2$ are σ-algebras, $\mathcal{F}_1 \cup \mathcal{F}_2$ may not be.

II. **Generating σ-algebras.** Given $A \subset \Omega$, denoted $\sigma\langle \mathcal{A} \rangle$, is the intersection of all σ-algebras $\mathcal{F}$ on Ω that contain $\mathcal{A}$,

$$\sigma\langle \mathcal{A} \rangle = \bigcap_{\mathcal{F} \supset \mathcal{A}} \mathcal{F}, \quad \mathcal{F} \text{ is a } \sigma\text{-algebra.}$$

III. **π-classes and λ-classes.**

A class $\mathcal{C}$ is a π-class if finite intersections of elements of $\mathcal{C}$ are also in $\mathcal{C}$.

A class $\mathcal{L}$ is a λ-class when only infinite unions of *disjoint* sets are guaranteed to be in $\mathcal{L}$.

IV. **Dynkin's π–λ theorem.** ***Thm 2.1.*** If $\mathcal{C}$ is a *pi*-class and $\mathcal{L}$ is a λ-class with $\mathcal{C} \subset \mathcal{L}$, then $\sigma\langle \mathcal{C} \rangle \subset \mathcal{L}$.

1.2 Measures

V. **Measure.** A *measure* is a nonnegative function, with $\mu(\varnothing) = 0$ and countable additivity. A *finite measure* has $\mu(\Omega) < \infty$.

VI. **Properties of probability measures.** ***Prop. 2.1, 2.2.*** A probability measure has $P(\Omega) = 1$ with the following properties:

(1) if $A \subset B$ then $P(A) \leq P(B)$

(2) inclusion/exclusion rule

(3) monotone continuity from below & above

(4) countable subadditivity (when not disjoint)

VII. **Uniqueness of measures.** ***Corr. 2.2.*** For finite measures μ_1 and μ_2 (on $\mathcal{F}$). Let $\mathcal{C} \subset \mathcal{F}$ be a π-class such $\mathcal{F} = \sigma\langle \mathcal{C} \rangle$. If $\mu_1(\Omega) = \mu_2(\Omega)$ and $\mu_1(C) = \mu_2(C)$ for all $C \in \mathcal{C}$, then $\mu_1(A) = \mu_2(A)$ for all $A \in \mathcal{F}$.

1.3 Measure Construction

VIII. **Semi-algebras.** A class $\mathcal{C}$ of subsets of Ω is a *semi-algebra* if for any $A, B \in \mathcal{C}$, (1) $A \cap B \in \mathcal{C}$ and (2) $A^c = \bigcup_{i=1}^{k} B_i$ for some disjoint $\{B_i\} \subset \mathcal{C}$. ($A \in \mathcal{C}$ but A^c does not have to be in $\mathcal{C}$, only it must be composed of finite disjoint sets in $\mathcal{C}$.)

IX. **Step 1. Semi-algebra selection and its measure.** Define a semi-algebra $\mathcal{C} = \{(a,b],(a,\infty) : -\infty \leq a < b < \infty,\ a,b \in \mathbb{R}\}$. We define this semi-algebra as it will include all sets of interest regarding probability measures.

Define a measure μ on $\mathcal{C}$: (1) $\mu(A) \in [0,\infty]$, (2) $\mu(\varnothing) = 0$, (3) for any disjoint $\{A_i\} \subset \mathcal{C}$ with $\bigcup_{i=1}^{\infty} A_i \in \mathcal{C}$,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i),$$

and (4) $\Omega \in \mathcal{C}$. (This is only a definition.)

X. **Step 2. Outer measure definition.** Define an outer measure

$$\mu^*(A) = \inf_{\{A_i\}\subset\mathcal{C}} \left\{ \sum_i \mu(A_i) : A \subset \bigcup_i A_i \right\}$$

for any set $A \subset \Omega$.

XI. **Step 3. Set restriction for outer measure.** Restrict the sets E on which μ^* operates to those that $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$ holds for all sets $A \subset \Omega$. Define the field $\mathcal{M}_{\mu^*}$ as the collection of all outer measurable sets in Ω.

XII. **Step 4. Caratheodory extension theorem.** The following are true:

(1) $\mathcal{M}_{\mu^*}$ is a σ-algebra.

Reference: **Collections of Lebesgue measurable sets.** ***Thrm. 1.1.1.***

(2) if $\mu^*(A) = 0$ then A is measurable.

Reference: **Zero measure measurability.** ***Prop. 1.1.3.***

(3) μ^* is a measure on $\mathcal{M}_{\mu^*}$.

Reference: **Countable additivity.** ***Prop. 1.1.5.***

(4) $\mathcal{C} \subset \mathcal{M}_{\mu^*}$ implying $\sigma\langle\mathcal{C}\rangle \subset \mathcal{M}_{\mu^*}$ as $\mathcal{M}_{\mu^*}$ is a σ-algebra.

Reference: **Every Borel set is measurable.** ***Thrm. 1.1.2.***

(5) $\mu^*(A) = \mu(A)$ for all $A \in \mathcal{C}$, which follows from the definition of outer measure when $A \in \mathcal{C}$.

XIII. **Step 5. Measure extension.** Thus, we have defined an outer measure μ^* on $\sigma\langle\mathcal{C}\rangle$ which is an extension of μ on $\mathcal{C}$.

XIV. **Uniqueness of measure extensions.** ***Thrm. 2.2.*** Let μ be a σ-finite measure on semi-algebra $\mathcal{C}$ of Ω, i.e., μ is a measure on $\mathcal{C}$ and there exists a partition $\{A_i\}$ of Ω where $\mu(A_i) < \infty$ for all i. Then there exists a measure ν on $\sigma\langle\mathcal{C}\rangle$ which is a unique extension of μ on $\sigma\langle\mathcal{C}\rangle$.

1.4 Measurable Transformations

XV. Measurable transformations. Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be measurable spaces. If a function $T : \Omega_1 \to \Omega_2$ satisfies

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

for all $A \in \mathcal{F}_2$, then T is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable.

XVI. **Random variables.** Let $(\Omega, \mathcal{F}, P)$ be a probability space. If $X : \Omega \to \mathbb{R}$ is $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-measurable then it is a *random variable.*

XVII. **Properties of measurable transformations.** ***Prop. 2.3.*** Let $(\Omega_i, \mathcal{F}_i)$ be measurable spaces and $T_i : \Omega_i \to \Omega_{i+1}$.

Then:

(1) If $\mathcal{F}_2 = \sigma\langle \mathcal{A} \rangle$ and $T_1^{-1}(A) \in \mathcal{F}_1$ for all $A \in \mathcal{A}$, then T_1 is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable

(2) If T_1 is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable and T_2 is $\langle \mathcal{F}_2, \mathcal{F}_3 \rangle$-measurable, then $T = T_2 \circ T_1$ is $\langle \mathcal{F}_1, \mathcal{F}_3 \rangle$-measurable.

XVIII. **Continuous transformations.** ***Prop. 2.4.*** If $f : \mathbb{R}^k \to \mathbb{R}^p$ is continuous, then then f is $\langle \mathcal{B}(\mathbb{R}^k), \mathcal{B}(\mathbb{R}^p) \rangle$-measurable.

XIX. **Properties of random variables.** ***Prop. 2.5.*** Let $X_i : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be random variables ($\langle \mathcal{F}, \mathcal{B}(\mathbb{R}^p) \rangle$-measurable functions).

Then:

(1) $\sum_{i=1}^k a_i X_i$ and $\prod_{i=1}^k a_i X_i$ are also random variables ($\langle \mathcal{F}, \mathcal{B}(\mathbb{R}^p) \rangle$-measurable functions).

(2) $\sup_n X_n$, $\inf_n X_n$, $\underline{\lim}\, X_n$, and $\overline{\lim}\, X_n$ are also random variables.

(3) The set $A = \{\omega \in \Omega : \lim_{n\to\infty} f_n < \infty\}$ lies in $\mathcal{F}$ and $g = \mathbb{I}_A \cdot \lim_{n\to\infty}$ is $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}^p) \rangle$-measurable.

XX. **Product spaces.** $(\Omega_i, \mathcal{F}_i)$ represent measurable spaces. The *product space* is $(\otimes_{i=1}^k \Omega_i, \otimes_{i=1}^k \mathcal{F}_i)$, noting k could be ∞ for an infinite space.

Let $A = A_1 \times A_2 \times \dots$ be a *measurable rectangle* where each $A_i \in \mathcal{F}_i$. As $\otimes_{i=1}^k \mathcal{F}_i \equiv \sigma\langle \{A_1 \times A_2 \times \dots\} \rangle$ is a σ-algebra on $\otimes_{i=1}^k \Omega_i$, it is a *product σ-algebra.*

A *cylinder set* has form $A = A_1 \times A_2 \times \dots$ if there exists k such $A_i = \Omega_i$ for all $i > k$.

XXI. **Coordinate maps.** Let $Z_t : \otimes_{i=1}^\infty \Omega_i \to \Omega_t$ be the coordinate extracting function $Z_t(\omega_1, \omega_2, \dots) = \omega_t$, making Z_t the tth *coordinate map.*

XXII. **Coordinate map measurability.** To show that $Z_t : \mathcal{B}(\mathbb{R}^\infty) \to \mathcal{B}(\mathbb{R})$ is $\langle \mathcal{B}(\mathbb{R}^\infty), \mathcal{B}(\mathbb{R}) \rangle$-measurable, show $Z_t^{-1}(A)$ is in $\mathcal{B}(\mathbb{R}^\infty)$ for all $A \in \mathcal{B}(\mathbb{R})$.

$Z_t^{-1}(A) = \{(\omega_1, \omega_2, \dots) \in \mathbb{R}^\infty : Z_t(\boldsymbol{\omega}) \in A\} = \{(\omega_1 \in \mathbb{R}, \omega_2 \in \mathbb{R}, \dots)\}$ with $\{\omega_t \in A\}$.

As $A \in \mathcal{B}(\mathbb{R})$, $Z_t^{-1}(\boldsymbol{\omega}) \in \mathcal{B}(\mathbb{R}^\infty)$ making Z_t $\langle \mathcal{B}(\mathbb{R}^\infty), \mathcal{B}(\mathbb{R}) \rangle$-measurable.

XXIII. **Mapping to multiple spaces.** ***Prop. 2.6.*** Let $T(\omega_0) = (T_1(\omega_0), T_2(\omega_0), \dots)$ be a map from $\Omega_0 \to \otimes_{i=1}^{\infty}\Omega_i$. Then:

T is $\langle \mathcal{F}_0, \otimes_{i=1}^{\infty}\mathcal{F}_i\rangle$-measurable $\iff$ T_i is $\langle \mathcal{F}_0, \mathcal{F}_i\rangle$-measurable for all i.

XXIV. **Induced measures.** Let $T : \Omega_1 \to \Omega_2$ be $\langle \mathcal{F}_1, \mathcal{F}_2\rangle$-measurable. For any measure μ on $(\Omega_1, \mathcal{F}_1)$, the set function $\mu T^{-1}(A_2) = \mu \circ T^{-1}(A_2) = \mu(T^{-1}(A_2))$ is a measure on $\mathcal{F}_2$ for all $A_2 \in \mathcal{F}_2$.

This is called the *measure induced by* T (under μ on $\mathcal{F}_2$).

XXV. **Random variables as induced measures.** For a random variable mapping $(\Omega, \mathcal{F}, P)$ onto $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, the induced measure μX^{-1} is the probability measure on $\mathcal{B}(\mathbb{R})$ and is denoted P_X.

$$\mu X^{-1}(A) = P_X(A) = P(X^{-1}(A)) = P(\{\omega \in \Omega : X(\omega) \in A\}) = P(X \in A)$$

for all $A \in \mathcal{B}(\mathbb{R})$. Thus $P(X \in A)$ is the measure $\mu X^{-1}(A)$ on $\mathcal{B}(\mathbb{R})$.

XXVI. **Cumulative distribution functions.** The cdf for a random variable X is

$$F(x) = P_X((-\infty, x]) = P(X \in (-\infty, x]) = \mu X^{-1}((-\infty, x]),$$

as μX^{-1} represents the measure induced by X on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

XXVII. **Properties of cdfs.** If $F(x)$ is the cdf of a random variable X then it is equal to $\mu X^{-1}((-\infty, x])$ and it is (1) right continuous, (2) monotone nondecreasing, and (3) has limits of 0 and 1 as x goes to $-\infty$ and ∞, respectively.

XXVIII. **Existence of probability spaces.** Let $F(x)$ have the three properties of a valid cdf. Then there exists a random variable X that maps $(0,1)$ onto $\mathbb{R}$ such that $F(x)$ is the cdf of X.

Proof: Define $\phi(\omega) = \min\{x \in \mathbb{R} : F(x) \geq \omega\}$ for any $\omega \in \Omega = (0,1)$. Thus $\phi(\omega)$ maps ω to the smallest x on $\mathbb{R}$ where $F(x) \geq \omega$. Thus $F(x) \geq \omega \iff \phi(\omega) \leq x$.

Let X be a random variable, whose cdf is defined as $P(X \leq x) = P(X(\omega) \leq x)$. In the field of Ω, this is $P(\{\omega \in (0,1) : X(\omega) \leq x\})$. Let $X(\omega) = \phi(\omega)$, making $P(X \leq x) = P(\{\omega \in (0,1) : \phi(\omega) \leq x\})$ $= P(\{\omega \in (0,1) : F(x) \geq \omega\})$ from the above iff statement.

Let m be the Lebesgue measure on space $(\,(0,1), \mathcal{B}(0,1)\,)$ Thus $P(X \leq x) = P(\,(0, F(x)]\,) = F(x)$, the last equality from the definition of Lebesgue measure on an interval. This has shown that $X(\omega)$ defined as $\phi(\omega)$ is a random variable from $(\,(0,1), \mathcal{B}(0,1), m)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_F)$ with μ_F as the measure induced by X under m on $\mathcal{B}(\mathbb{R})$.

2 The Lebesgue Measure

XXIX. **A set measure.** The ideal set measure assigns a nonnegative value $m(A)$ to a subset $A \subset \mathbb{R}$ such:

(1) for an interval I, $m(I) = l(I)$

(2) if $\{A_n\}_{n=1}^{\infty}$ is a sequence of disjoint sets, then

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

(3) has translation invariance, i.e., if $A + y = \{x + y : x \in A\}$ then $m(A) = m(A + y)$.

(4) can operate on *any* subset $A \subset \mathbb{R}$.

All properties cannot be attained, so the last one will be dropped, i.e., the set measure will be restricted to certain subsets of $\mathbb{R}$.

XXX. **A cover of** $\boldsymbol{A}$**.** A countable collection of disjoint open intervals $\{I_n\}_{n=1}^{\infty}$ is a *cover* of $A \subset \mathbb{R}$ if $A \subset \bigcup_{n=1}^{\infty} I_n$.

XXXI. **Outer measure.** The Lebesgue *outer measure* $m^\star(A)$ of a subset A is

$$m^\star(A) = \inf\left\{\sum l(I_n) : \text{any cover}\{I_n\} \text{ of } A\right\}$$

This means that for any $\epsilon > 0$ there exists a cover $\{I_n\}$ of A such that

$$\sum_{n=1}^{\infty} l(I_n) < m^\star(A) + \epsilon$$

XXXII. $\boldsymbol{m^\star}$ **of intervals.** ***Prop. 1.1.1.*** For an interval $I \subset \mathbb{R}$, $m^\star(I) = l(I)$.

Proof: (1) The smallest cover of (a, b) is (a, b) itself, so the infimum of sum of lengths of all covers is $l(a, b) = b - a$.

(2) Note that $m^\star[a, b] = m^\star(a, b)$.

$[a, b] = \{a, b\} \cup (a, b)$, $m^\star[a, b] \le m^\star(a, b) + m^\star(\{a, b\})$ using $m^\star$ additivity of disjoint sets and $m^\star$ of points is zero. Also $(a, b) \subset [a, b]$, so $m^\star(a, b) \le m^\star[a, b]$. Together $m^\star(a, b) = m^\star[a, b]$.

XXXIII. $\boldsymbol{m^\star}$ **of countable collections of subsets.** ***Prop. 1.1.2.*** If $\{A_n\}_{n=1}^{\infty}$ is a countable collection of subsets of $\mathbb{R}$, then

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

Proof: For each A_n, given any $\epsilon' > 0$ by definition of the infimum in $m^\star$ there exists a cover $\{I_j^n\}_{j=1}^{\infty}$ of A_n such

$$\sum_{j=1}^{\infty} l(I_j^n) < m^\star(A_n) + \epsilon\, 2^{-n}. \tag{1}$$

As $\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{\infty} I_j^n$,

$$m^\star\left(\bigcup_{n=1}^{\infty} A_n\right) \le m^\star\left(\bigcup_{n=1}^{\infty}\bigcup_{j=1}^{\infty} I_j^n\right) \le \sum_{n=1}^{\infty} m^\star\left(\bigcup_{j=1}^{\infty} I_j^n\right) \le \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} l(I_j^n),$$

using $m^\star$ of subsets, countable subadditivity, and the definition of $m^\star$.

Applying (1):

$$m^\star\left(\bigcup_{n=1}^{\infty} A_n\right) < \sum_{n=1}^{\infty} m^\star(A_n) + \epsilon\, 2^{-n} = \sum_{n=1}^{\infty} m^\star(A_n) + \epsilon.$$

Letting $\epsilon \to 0$ gives the result.

Aside: if $a < b + \epsilon$ is true for any $\epsilon > 0$, then $a \leq b$. Suppose $a > b$, then selecting $\epsilon = (a-b)/2$ results in a contradiction, and the supposition is false. Note that $a < b + \epsilon$ holds for all $\epsilon > 0$ even when $a = b$. Thus $a \leq b$.

XXXIV. **$m^\star$ properties.** ***Cor. 1.1.1.***

(i.) If A is countable, $m^\star(A) = 0$.

Proof: $A = \{a_n\}_{n=1}^{\infty}$. $m^\star(A) \leq \sum_{n=1}^{\infty} m^\star(a_n) = 0$.

(ii.) For any set A and $\epsilon > 0$, there is an open set O such that $A \subset O$ and $m^\star(O) \leq m^\star(A) + \epsilon$.

Proof: Since O is open, $m^\star(O)$ is the sum of its interval lengths. Apply the infimum definition for $m^\star(A)$.

XXXV. **Lebesgue measurability.** A set E is *Lebesgue measurable* if for any set A it is true that $m^\star(A) = m^\star(A \cap E) + m^\star(A \cap E^c)$.

As $m^\star(A) \leq m^\star(A \cap E) + m^\star(A \cap E^c)$ always holds by , only $m^\star(A \cap E) + m^\star(A \cap E^c) \leq m^\star(A)$ needs to be shown for any set A.

XXXVI. **Zero measure measurability.** ***Prop. 1.1.3.*** If $m^\star(E) = 0$ then E is measurable.

Proof: $m^\star(A) \leq m^\star(A \cap E) + m^\star(A \cap E^c) \leq m^\star(E) + m^\star(A) = m^\star(A)$ as $A \subset (A \cap E) \bigcup (A \cap E^c)$, $A \cap E \subset E$, $A \cap E^c \subset A$, and $m^\star(E) = 0$ for any set A. Hence $m^\star(A) = m^\star(A \cap E) + m^\star(A \cap E^c)$ and E is measurable.

XXXVII. **Measurability of unions.** ***Prop. 1.1.4-1.*** If E_1 and E_2 are measurable, then $E_1 \cup E_2$ is measurable.

Proof: Need to show $m^\star(A \cap (E_1 \cup E_2)) + m^\star(A \cap (E_1 \cup E_2)^c) \leq m^\star(A)$ for any set A.

(1) $A \cap (E_1 \cup E_2) = (A \cap E_1) \bigcup (A \cap E_1^c \cap E_2)$.

(2) $(E_1 \cup E_2)^c = E_1^c \cap E_2^c$

$$\begin{aligned} m^\star(A \cap (E_1 \cup E_2)) &+ m^\star(A \cap (E_1 \cup E_2)^c) \\ &\leq m^\star(A \cap E_1) + m^\star(A \cap E_1^c \cap E_2) + m^\star(A \cap E_1^c \cap E_2^c) \\ &\leq m^\star(A \cap E_1) + m^\star(A \cap E_1^c) && (E_2 \text{ is measurable}) \\ &\leq m^\star(A) && (E_1 \text{ is measurable}) \end{aligned}$$

XXXVIII. **Measurable sets.**

(1) If E is measurable, then E^c is measurable by definition.

(2) If E and F are measurable, then $E \cap F$ is measurable.

Proof: As E^c and F^c are measurable, $(E^c \cup F^c)^c = E \cap F$ is measurable.

(3) Any finite combination of unions, intersections, and complements of measurable sets is itself measurable.

XXXIX. **Collections of measurable sets.** ***Cor. 1.1.2.*** The collection of all measurable sets, denoted $\mathcal{M}$, is an algebra.

Proof: An algebra contains all finite collections of complements and unions.

XL. **Collections of Lebesgue measurable sets.** ***Thrm. 1.1.1.*** The collection of all Lebesgue measurable sets is a σ-algebra, i.e., the complement and union of any countable collection of Lebesgue measurable sets is Lebesgue measurable.

Proof: Both Ω and $\varnothing$ are Lebesgue measurable by the fact sets of zero measure are measurable and complements of measurable sets are measurable.

To show infinite unions of measurable sets are measurable, let E be a countable collection of sets. There exists a countable collection $\{E_n\}$ of disjoint sets such that $E = \bigcup_{n=1}^{\infty} E_n$ (via Royden Prop. 1.2).

Let $F_n = \bigcup_{i=1}^{n} E_i$. It follows that

$$\begin{aligned}
m^\star(A) &= m^\star(A \cap F_n) + m^\star(A \cap F_n^c) && (F_n \text{ is measurable})\\
&\geq m^\star(A \cap F_n) + m^\star(A \cap E^c) && (E^c \subset F_n^c \ \forall n)\\
&\geq \sum_{i=1}^{n} m^\star(A \cap E_i) + m^\star(A \cap E^c) && (\text{finite additivity of } m^\star \text{ on measurable } \{E_i\})\\
&\geq \sum_{i=1}^{\infty} m^\star(A \cap E_i) + m^\star(A \cap E^c) && (\text{inequality property when one side is independent of } n)\\
&\geq m^\star(A \cap E) + m^\star(A \cap E^c) && (\text{countable subadditivity})
\end{aligned}$$

Thus E is Lebesgue measurable. $\Rightarrow$ the collection of Lebesgue measurable sets is a σ-algebra.

XLI. **Interval measurability.** ***Prop. 1.1.4-2.*** The interval (a, ∞) is measurable.

Proof: For any set A, suppose $m^\star(A) < \infty$. For any $\epsilon > 0$ there is a collection $\{I_n\}_{n=1}^{\infty}$ of disjoint open intervals that is a cover of A for which $\sum_{n=1}^{\infty} l(I_n) \leq m^\star(A) + \epsilon$, by definition of $m^\star$.

Define $\{I_n^1\}_{n=1}^{\infty}$ and $\{I_n^2\}_{n=1}^{\infty}$ where $I_n^1 = I_n \cap (a, \infty)$ and $I_n^2 = I_n \cap (-\infty, a]$.

$$\begin{aligned}
m^\star(A_1) + m^\star(A_2) &\leq \sum_{n=1}^{\infty} \left(m^\star(I_n^1) + m^\star(I_n^2)\right) && (A_j \subset \textstyle\bigcup_{n=1}^{\infty} I_n^j)\\
&\leq \sum_{n=1}^{\infty} l(I_n) && (m^\star(I_n^1) + m^\star(I_n^2) = l(I_n^1) + l(I_n^2) = l(I_n))\\
&< m^\star(A) + \epsilon && (\text{see XXXIII.})
\end{aligned}$$

Thus $m^\star(A_1) + m^\star(A_2) \leq m^\star(A)$ and (a, ∞) is measurable.

XLII. **Every Borel set is measurable.** ***Thrm. 1.1.2.*** From the fact (a, ∞) is measurable, it is easy to show that (a, b) is measurable. The collection of measurable sets $\mathcal{M}$ is a σ-algebra containing all open intervals. As the collection of Borel sets, $\mathcal{B}$, in the smallest σ-algebra containing all open intervals, $\mathcal{B} \subset \mathcal{M}$. Thus, any Borel set $B \in \mathcal{B}$ is measurable.

XLIII. **Lebesgue measure.** If E is measurable, $E \in \mathcal{M}$, then the Lebesgue measure is $m(E) = m^\star(E)$.

XLIV. **Countable additivity.** ***Prop. 1.1.5.***

For a sequence $\{E_n\}$ of measurable sets, $m\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{n=1}^{\infty} m(E_n)$, by countable subadditivity.

For a sequence of $\{E_n\}$ of disjoint measurable sets, $m\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} m(E_n)$.

Proof:

$$\begin{aligned} m\left(\bigcap_{i=1}^{\infty} E_i\right) &\geq m\left(\bigcap_{i=1}^{n} E_i\right) && \text{(LHS} \supset \text{RHS)} \\ &\geq \sum_{i=1}^{n} m(E_i) && \text{(finite additivity of disjoint sets)} \\ &\geq \sum_{i=1}^{\infty} m(E_i) && \text{(inequality property when one side is independent of } n\text{)} \end{aligned}$$

Thus $m\left(\bigcap_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} m(E_i)$.

XLV. **Decreasing sequences of measurable sets.** ***Prop. 1.1.6.*** Let $\{E_n\}$ be an infinite decreasing sequence of measurable sets, i.e., $E_{n+1} \subset E_n$, with $m(E_1) < \infty$. Then

$$m\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} m(E_n).$$

Proof: Define $F_n = E_n \setminus E_{n+1} \Rightarrow m(F_n) = m(E_n) - m(E_{n+1})$. $\{F_n\}$ are disjoint and measurable.

$$m(E_1) = m\left(E \bigcup \left(\bigcup_{n=1}^{\infty} F_n\right)\right) \Rightarrow m(E_1) - m(E) = \sum_{n=1}^{\infty} m(F_n).$$

Thus

$$\begin{aligned} m(E_1) - m(E) &= \sum_{i=1}^{\infty} (m(E_i) - m(E_{i+1})) = \lim_{n\to\infty} \sum_{i=1}^{n} (m(E_i) - m(E_{i+1})) \\ &= \lim_{n\to\infty} (m(E_1) - m(E_n)) = m(E_1) - \lim_{n\to\infty} m(E_n). \end{aligned}$$

As $m(E_1) < \infty$, $m(E) = \lim_{n\to\infty} m(E_n)$.

XLVI. **Symmetric difference of measurable sets.** ***Prop. 1.1.7.*** Let E be a set with $m^\star(E) < \infty$. E is measurable $\iff$ given $\epsilon > 0$, there is a finite union of disjoint open intervals, U, such that $m^\star(U \Delta E) < \epsilon$.

Proof: "⇒"

Given E is measurable and $m^\star(E) < \infty$. For any $\epsilon' > 0$, there exists a cover of E by $O = \{I_i\}_{i=1}^\infty$ of disjoint open intervals such that $m^\star(O) < m^\star(E) + \epsilon'$. Since $\lim_{n\to\infty} \sum_{i=1}^n m^\star(I_i) = m^\star(O) < \infty$, for any $\epsilon' > 0$, there exists N such that $\sum_{i=N}^\infty m^\star(I_i) < \epsilon'$. Let $U = \bigcup_{i=1}^N I_i$.

$m^\star(U\Delta E) \le m^\star(U\cap E^c) + m^\star(U^c\cap E)$. $m^\star(U\cap E^c) \le m^\star(O\cap E^c) < \epsilon'$. $m^\star(U^c\cap E) \le m^\star(U^c\cap O) < \epsilon'$. Select $\epsilon' = \epsilon/2$ and $m^\star(U\Delta E) < \epsilon$ for any $\epsilon > 0$.

"⇐"

Given $\epsilon > 0$ such that $m^\star(U\Delta E) < \epsilon$. $m^\star(A) \le m^\star(A\cap E) + m^\star(A\cap E^c)$. $m^\star(A\cap E) \le m^\star(A\cap E\cap U) + m^\star(A\cap E\cap U^c)$. $m^\star(A\cap E^c) \le m^\star(A\cap E^c\cap U) + m^\star(A\cap E^c\cap U^c)$.

$$\begin{aligned} &m^\star(A\cap E) + m^\star(A\cap E^c) \\ &\qquad \le m^\star(A\cap E\cap U) + m^\star(A\cap E^c\cap U^c) + 2m^\star(U\Delta E) \\ &\qquad \le m^\star(A\cap E\cap U) + m^\star(A\cap E^c\cap U^c) + 2\epsilon \\ &\qquad \le m^\star(A\cap U) + m^\star(A\cap U^c) + 2\epsilon \\ &\qquad \le m^\star(A) + 2\epsilon \qquad \text{(U is measurable)} \end{aligned}$$

Thus $m^\star(A\cap E) + m^\star(A\cap E^c) \le m^\star(A)$ for any set A.

XLVII. **Measurable functions.** ***Prop. 1.1.8.*** Let $f : E \to \mathbb{R} \cup \{\pm\infty\}$ where E is measurable. Then the following are equivalent:

(1) The set $\{x \in E : f(x) > \alpha\}$ is measurable.

(2) The set $\{x \in E : f(x) \ge \alpha\}$ is measurable.

(3) The set $\{x \in E : f(x) < \alpha\}$ is measurable.

(4) The set $\{x \in E : f(x) \le \alpha\}$ is measurable.

Any implies the set $\{x \in E : f(x) = \alpha\}$ is measurable.

XLVIII. **Function measurability.** A function is measurable if (1) its domain E is measurable and (2) one of the four statements in *1.1.8* is true.

XLIX. **Functions of measurable functions.** ***Prop. 1.1.9.*** If f and g are both measurable and have the same domain, then $c + f$, cf, f^2, $f + g$, $f - g$, and fg are also measurable.

Proof: $\{x \in E : c + f(x) > \alpha\}$ is measurable since $\{x \in E : f(x) > \alpha'\}$ is measurable, where $\alpha' = \alpha - c$.

$\{x \in E : c\,f(x) > \alpha\}$ is measurable since $\{x \in E : f(x) > \alpha'\}$ is measurable, where $\alpha' = \alpha/c$.

$A = \{x \in E : (f(x))^2 > \alpha\}$ yields

$$A = \begin{cases} E & \text{if } \alpha < 0 \\ \{x \in E : |f(x)| > \sqrt{\alpha}\} & \text{if } \alpha \ge 0 \end{cases}$$

A is measurable using property (1) and (3) of *1.1.8* and the fact unions of measurable sets are measurable.

$A = \{x \in E : f(x)+g(x) < \alpha\}$. If $f(x)+g(x) < \alpha$ then there exists $r \in \mathbb{Q}$ such that $f(x) < r < \alpha - g(x)$. As the rationals are countable, the following set is measurable

$$A = \bigcup_r \left(\{x \in E : f(x) < r\} \cap \{x \in E : r < \alpha - g(x)\}\right).$$

fg is measurable by application of these results to $[(f+g)^2 - f^2 - g^2]/2$.

L. **Measurability of sequences of functions.** ***Prop. 1.1.10.*** Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of measurable functions on domain E. Then $\sup f_1, f_2, \dots, f_k$, $\sup f_1, f_2, \dots, f_k$, $\sup_n f_n$, $\inf_n f_n$, $\overline{\lim} f_n$, and $\underline{\lim} f_n$ are all measurable functions.

Proof: Let h_n be defined as $h_n(x) = \sup\{f_1(x), f_2(x), \dots, f_n(x)\}$. The set $A = \{x \in E : h_n(x) > \alpha\} = \bigcup_{i=1}^{n}\{x \in E : f_n(x) > \alpha\}$. As $\{x \in E : f_n > \alpha\}$ is measurable for all n, their union is measurable, thus A is measurable, $\Rightarrow \sup\{f_1(x), f_2(x), \dots, f_n(x)\}$ is a measurable function.

As $n \to \infty$, A is still measurable as the union is countable. Thus $\sup_n f_n$ is a measurable function. Similar arguments confirm infima. $\overline{\lim}$ is an infimum of a measurable function $(\sup_k\{f_k : k \geq n\})$, $\Rightarrow \overline{\lim} f_n$, and $\underline{\lim} f_n$ are measurable functions.

LI. **Almost everywhere.** A property holds *almost everywhere* if the set where it does not hold has measure zero.

LII. **Measurability of equal functions.** ***Prop. 1.1.11.*** If f is a measurable function and $f = g$ a.e., then g is measurable.

Proof: Suppose $E = \{x : f(x) \neq g(x)\}$. $m(E) = 0$, thus E^c is a measurable set. If $\{x : g(x) > \alpha\}$ is measurable, then $g(x)$ is a measurable function. It is true that

$$\begin{aligned}\{x : g(x) > \alpha\} &= (\{x : f(x) > \alpha\} \cap E^c) \bigcup (\{x : g(x) > \alpha\} \cap E) \\ &\equiv A_1 \bigcup A_2\end{aligned}$$

where A_1 is measurable as it is the intersection of two measurable sets, and $A_2 \subset E$ that has zero measure, thus $m(A_2) = 0$ and is A_2 is measurable. Thus $\{x : g(x) > \alpha\}$ is a measurable set, so g is a measurable function.

LIII. **Step functions.** A step function $s(x)$ is measurable and assumes a constant value on the intervals $d_0 = a < d_1 < \cdots < d_n = b$ of $[a, b]$

$$s(x) = c_i \in \mathbb{R} \text{ on } (d_{i-1}, d_i) : i = 1, \dots, n.$$

Note that $s(x)$ may equal any value at the points $\{d_i\}$.

LIV. **Simple functions.** A simple function is measurable and assumes only a finite number of values $\{\alpha_n\}$. It is possible to write a simple function in the form

$$\phi(x) = \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i}(x)$$

where $A_i = \{x : \phi(x) = \alpha_i\}$ are disjoint sets.

LV. **Step and continuous function existence.** ***Prop. 1.1.12.*** If f is a measurable function on $[a, b]$ and assumes values $\pm\infty$ on a set of measure zero. Then, given $\epsilon > 0$, a step function g and continuous function exist such that

$$|f(x) - g(x)| < \epsilon \quad \text{and} \quad |f(x) - h(x)| < \epsilon$$

hold for all $x \in [a, b]$ except for sets E_g and E_h with measures $< \epsilon$. If, in addition, $m \le f \le M$, then g and h may exist on $[m, M]$.

Proof: Let $A_n = \{x \in [a, b] : |f(x)| > n\}$. As $\lim_{n\to\infty} A_n = 0$, for any $\epsilon/3 > 0$ there exists $M > 0$ such that $m(A_n) < \epsilon/3$ for all $n \ge M$. This follows from *1.1.6* as A_n is a decreasing set and $m(\{x \in [a, b] : f(x) = \pm\infty\}) = 0$.

Given $\epsilon/3 > 0$ and M, there exists a number of divisions $M^\star \ge 3M/\epsilon$, $M^\star \in \mathbb{Z}$ such that the following simple function exists

$$\phi(x) = \begin{cases} n\epsilon/3 & n\epsilon/3 \le f(x) < (n+1)\epsilon/3, \; n = \{-M^\star, \ldots, M^\star\} \\ 0 & otherwise \end{cases}$$

so that $|f(x) - \phi(x)| < \epsilon/3$. (The set $\{|f(x) - \phi(x)| \ge \epsilon/3$ occurs where $|f(x)| \ge M$, a set of measure $< \epsilon/3$.) Note that the range of $\phi(x)$ includes $[-M, M]$ as $-M^\star\epsilon/3 \le (3M/\epsilon)(-\epsilon/3) = -M$ and $(M^\star + 1)\epsilon/3 > (3M/\epsilon)(\epsilon/3) = M$.

Given a simple function

$$\phi(x) = \sum_{n=-M^\star}^{M^\star} \frac{n\epsilon}{3} \, \mathbb{I}_{A_n}(x)$$

where $A_n = \{x : n\epsilon/3 \le f(x) < (n+1)\epsilon/3\}$. Each A_n can be almost covered by a finite union of disjoint open intervals $\{I_j^n\}_{j=1}^{N_n}$, where $\{I_j^n\}$ is chosen such, denoting $U_n = \bigcup_{j=1}^{N_n} I_j^n$, $m(U_n \Delta A_n) < \epsilon/6M^\star$, by application *1.1.7*. Define a step function

$$g(x) = \sum_{n=-M^\star}^{M^\star} \frac{n\epsilon}{3} \, \mathbb{I}_{U_n}(x)$$

Defining the set

$$E = \bigcup_n (U_n \Delta A_n) : n = \{-M/\epsilon, \ldots, M/\epsilon\}$$

$m(E) \le \sum_n m(U_n \Delta A_n) : n = \{-M^\star, \ldots, M^\star\}$, and as each $m(U_n \Delta A_n) < \epsilon/6M^\star$, $m(E) \le (2M^\star)(\epsilon/6M^\star) = \epsilon/3$. Since U_n is a finite collection of disjoint open intervals, the union $\bigcup_{n=-M^\star}^{M^\star} U_n$ is a finite collection of open intervals and has a finite number of separation points between a and b, say of quantity $Q_{sep} = \sum_{n=-M^\star}^{M^\star} N_n$.

A step function defined on $[a, b]$ with separation points $\{x_q : q = 1, \ldots, Q_{sep}\}$ can be turned into a continuous function by defining a set A_Q containing a region of ϵ' around every separation point, $[x_q - \epsilon', x_q + \epsilon'], \forall q$. Define $h(x)$ as $g(x)$ on A_Q^c and, when in A_Q, as a connection line from the endpoints in A_Q^c. This will be a continuous function. By choosing $\epsilon' < \epsilon/3Q_{sep}$, the measure of A_Q will be less than $\epsilon/3$. Thus $|f(x) - h(x)| < \epsilon$ except on a set of measure $< \epsilon$.

LVI. **Convergence of measurable functions.** ***Prop. 1.1.13.*** Let E be a measurable set of finite measure and $\{f_n\}$ a sequence of measurable functions that converge to f a.e. on E. Then, given $\epsilon > 0$ and $\delta > 0$, there is a set $A \subset E$ with $m(A) < \delta$ and N such that for all $n \geq N$, $|f_n(x) - f(x)| < \epsilon$ when $x \notin A$.

Proof: Let $A_k = \{x \in E : |f_k(x) - f(x)| \geq \epsilon\}$. Let $E_n = \bigcup_{k=n}^{\infty} A_k = \{x \in E : |f_k(x) - f(x)| < \epsilon$ for some $k \geq n\}$.

$\bigcap_{n=1}^{\infty} E_n = \{x \in E :$ for every n it is true for some $k \geq n$ that $|f_k(x) - f(x)| \geq \epsilon\}$. $\Rightarrow \bigcap_{n=1}^{\infty} E_n = \{x \in E : f_n(x) \nrightarrow f(x)\} = B$.

As $\lim_{n\to\infty} m(E_n) = 0$, for any $\delta > 0$ there exists N such $m(E_n) < \delta$ for all $n \geq N$. Let $A = E_N$.

3 Lebesgue Integration

LVII. **Riemann integrals of a step functions.** The Riemann integral of a step function $s(x)$ is given by

$$\int_a^b s(x)\,dx = \sum_{i=1}^{n} c_i\,(d_i - d_{i-1})$$

As the number of points, $\{d_i\}$, is finite, their measure is zero.

LVIII. **Bounded functions.** Taking an arbitrary division of $[a, b]$, $\{d_i : i = 1, \ldots, n\}$, of function f, a step function $s_L(x)$ exists such that

$$s_L(x) = m_i \text{ on } [d_{i-1}, d_i), \text{ for all } i = 1, \ldots, n,$$

where $m_i = \inf\{f(x) : d_{i-1} \leq x \leq d_i\}, i = 1, \ldots, n$. Also a step function $s_U(x)$ exists such that

$$s_U(x) = M_i \text{ on } [d_{i-1}, d_i), \text{ for all } i = 1, \ldots, n,$$

where $M_i = \sup\{f(x) : d_{i-1} \leq x \leq d_i\}, i = 1, \ldots, n$.

LIX. **Riemann bound integrals.** The upper and lower Riemann integrals of f are

$$R_U \int_a^b f = \inf_{s_U(x)} \left\{ \int_a^b s_U : \text{ any } s_U(x) \geq f(x) \text{ on } [a, b] \right\}$$
$$R_L \int_a^b f = \sup_{s_L(x)} \left\{ \int_a^b s_L : \text{ any } s_L(x) \leq f(x) \text{ on } [a, b] \right\}$$

It is always true that $R_L \int_a^b f \leq R_U \int_a^b f$, and if $R_L \int_a^b f = R_U \int_a^b f$ then f is *Riemann integrable* on $[a, b]$, denoted by $R \int_a^b f$.

LX. **Lebesgue integrals of simple functions.** Let the simple function $\phi(x) = \sum_{i=1}^{n} a_i \mathbb{I}_{E_i}(x)$, where $\{E_i\}$ are measurable. The Lebesgue integral of ϕ is

$$\int \phi = \sum_{i=1}^{n} a_i\, m(E_i), \text{ where } \{a_i\} \subset \begin{cases} (-\infty, \infty) & m(E_i) < \infty,\ \forall\, i \\ [0, \infty) & \text{otherwise} \end{cases}$$

LXI. **Integration on sets.** For any measurable set E, the Lebesgue integral is

$$\int_E \phi = \int \phi\, \mathbb{I}_E = \sum_{i=1}^{n} a_i\, m(E_i \cap E).$$

LXII. **Properties of integrals of simple functions.** ***Prop. 1.2.1.*** Let ϕ and ψ be simple functions which are zero outside of a set of finite measure. Define $\phi(x) = \sum_{i=1}^{n} a_i\ \mathbb{I}_{A_i}(x)$ and $\psi(x) = \sum_{i=1}^{m} b_i\ \mathbb{I}_{B_i}(x)$.

(1) $\int(\alpha\phi + \beta\psi) = \alpha \int \phi + \beta \int \psi$.

Proof: Define $\{c_i\}$ and $\{E_i\}$ for $i = 1, \ldots, n+m$, such

$$E_i = \begin{cases} A_i & i = 1, \ldots, n \\ B_{i-n} & i = n+1, \ldots, n+m \end{cases}$$

$$c_i = \begin{cases} \alpha\, a_i & i = 1, \ldots, n \\ \beta\, b_{i-n} & i = n+1, \ldots, n+m \end{cases}$$

It follows that $\int(\alpha\phi + \beta\psi) = \sum_{i=1}^{n+m} c_i\, m(E_i) = \alpha \sum_{i=1}^{n} a_i\, m(A_i)\ + \beta \sum_{i=1}^{m} b_i\, m(B_i) = \alpha \int \phi + \beta \int \psi$.

(2) If $\phi \geq \psi$ a.e. then $\int \phi \geq \int \psi$.

Proof: $\phi - \psi$ is also a simple function with $d \leq n+m$ values, $\phi(x) - \psi(x) = \sum_{i=1}^{d} h_i\ \mathbb{I}_{H_i}(x)$ where $H_i = \{x : \phi(x) - \psi(x) = h_i\}$, noting that $\{H_i\}$ has finite measure for all i. Thus, $\int(\phi - \psi) = \int \phi - \int \psi = \sum_{i=1}^{d} h_i\ m(H_i)$, using the above result. If $h_i < 0$ then $m(H_i) = 0$ as $\phi < \psi$ only on a set of measure zero. Thus $m(H_i) > 0$ only when $h_i \geq 0$, so $\int \phi - \int \psi \geq 0 \Rightarrow \int \phi \geq \int \psi$.

(3) If $\phi = \psi$ a.e. then $\int \phi = \int \psi$.

Proof: Follows from applying the above result to both $\phi \geq \psi$ a.e. and $\psi \geq \phi$ a.e., both of which are true.

Note that if ϕ and ψ are nonnegative simple functions and a and b are nonnegative, then, using the restricted form of the Lebesgue integral defined in LX, the above conditions all hold.

LXIII. **Lebesgue integrals of bounded functions.** As any measurable function f that is infinite only on a set of zero measure can be approximated to precision ϵ by a simple function ϕ, the upper and lower Lebesgue integrals of f are defined as

$$\inf_{\phi \geq f} \int_E \phi \equiv \inf \left\{ \int_E \phi :\ \text{any simple function } \phi(x) \geq f(x) \right\}$$

$$\sup_{\psi \leq f} \int_E \psi \equiv \sup \left\{ \int_E \psi :\ \text{any simple function } \psi(x) \leq f(x) \right\}$$

for any bounded, measurable f on a measurable set E that has finite measure. A function is *Lebesgue measurable* if $\sup_{\psi \le f} \int_E \psi = \inf_{\phi \ge f} \int_E \phi$.

LXIV. **Lebesgue integrability.** ***Prop. 1.2.2.*** Let f be a bounded function on measurable set E that has finite measure.

f is Lebesgue integrable on E, i.e., $\int_E f = \sup_{\psi \le f} \int_E \psi = \inf_{\phi \ge f} \int_E \phi \iff f$ is measurable.

Proof: "$\Leftarrow$" Suppose f is measurable.

Let $|f| < M$. Define sets $E_{k,n} = \{x \in E : (k-1)M/n < f(x) \le kM/n\}$ for $k = -n, \ldots, n$. The $\{E_{k,n}\}$ are disjoint with $E = \bigcup_{k=-n}^{n} E_{k,n}$ (as E is measurable and as $m(E) < \infty$) such that $m(E) = \sum_{k=-n}^{n} m(E_{k,n})$. Define bounding simple functions for f

$$\psi_n(x) = \sum_{k=-n}^{n} \frac{M(k-1)}{n} \mathbb{I}_{E_{k,n}}(x)$$

$$\phi_n(x) = \sum_{k=-n}^{n} \frac{Mk}{n} \mathbb{I}_{E_{k,n}}(x)$$

such that $\psi_n(x) \le f(x) \le \phi_n(x)$.

The following statement is true

$$\int_E \psi_n \ [1] \le \sup_{\psi \le f} \int_E \psi \ [2] \le \inf_{\phi \ge f} \int_E \phi \ [3] \le \int_E \phi_n \ [4]$$

[1] $\le$ [2] and [3] $\le$ [4] as [1] and [4] are examples of simple functions included in the supremum and infimum of [2] and [3]. [2] is always less than [3].

To show that [2]=[3], it is enough to show that [1]=[4]. As [1] is equal to $\sum_{k=-n}^{n} \frac{M(k-1)}{n} m(E_{k,n})$, and [4] is equal to $\sum_{k=-n}^{n} \frac{Mk}{n} m(E_{k,n})$, [4]$-$[1]$= \frac{M}{n} \sum_{k=-n}^{n} m(E_{k,n}) = \frac{M\, m(E)}{n}$. As n can be any integer, [4]$-$[1] must be zero. Thus f is Lebesgue integrable.

"$\Rightarrow$" Suppose f is Lebesgue integrable, i.e., suppose $\sup_{\psi \le f} \int_E \psi = \inf_{\phi \ge f} \int_E \phi$.

This means the same construction for sequences $\psi_n \le f$ and $\phi_n \ge f$ is possible due to the fact f is Lebesgue integrable, and these sequences are chosen such $\int_E \phi_n - \int_E \psi_n < 1/n$. Then $\psi^\star = \sup \psi_n$ and $\phi^\star = \inf \phi_n$ are both measurable as suprema and infima of sequences of measurable functions are measurable, using *1.1.10*. It is also true that $\psi^\star \le f \le \phi^\star$.

Let $\Delta = \{x : \psi^\star(x) < \phi^\star(x)\}$. $\Delta = \bigcup_{n=1}^{\infty} \Delta_\nu$, where $\Delta_\nu = \{x : \psi^\star(x) < \phi^\star(x) - 1/\nu\}$. As $\psi_n(x) \le \psi^\star(x) \le \phi^\star(x) \le \phi_n(x)$ for all n, $\psi_n(x) \le \psi^\star(x) < \phi^\star(x) - 1/\nu \le \phi_n(x) - 1/\nu$. That implies every Δ_ν is contained in $\{x : \psi_n(x) < \phi_n(x) - 1/\nu\}$.

The integral $\int_E (\phi_n - \psi_n) < 1/n$ (using *1.2.1*) and as $\phi_n - \psi_n > 1/\nu$ on a set of measure $\ge m(\Delta_\nu)$, $1/n > \int_E (\phi_n - \psi_n) \ge (1/\nu)\, m(\Delta_\nu)$. Hence, $m(\Delta_\nu) \le \nu/n$. As n is arbitrary, $m(\Delta_\nu) = 0\ \forall \nu$, and $m(\Delta) \le \sum_{\nu=1}^{\infty} m(\Delta_\nu) = 0$. Thus, $\psi^\star = \phi^\star$ except on a set of measure zero, $\Rightarrow \psi^\star = \psi^\star$ a.e. As $\psi^\star \le f \le \phi^\star$, it also holds that $\psi^\star = f$ a.e. Applying *1.1.11*, f is measurable on E as $\psi^\star$ is measurable on E.

LXV. **Riemann integrability implies Lebesgue integrability.** ***Prop. 1.2.3.*** If a bounded function f is Riemann integrable on $[a, b]$, then it is Lebesgue integrable on $[a, b]$ (and thus is measurable using *1.2.2.*)

Proof: The supremum of simple functions $\leq f$ is greater than or equal to the supremum of step functions $\leq f$, as all step functions are simple functions. Similarly, the infimum of simple functions $\geq f$ is less than of equal to the infimum of step functions $\geq f$. As the supremum of simple functions $\leq f$ is less than of equal to the infimum of simple functions $\geq f$, the following is true

$$R_L \inf f = \sup\{\int s_L\} \leq \sup\{\int \psi\} \leq \inf\{\int \phi\} \leq \inf\{\int s_U\} = R_U \inf f$$

As the first term equals the last term, all terms are equal, and thus $\sup_{\psi \leq f} \int_E \psi = \inf_{\phi \geq f} \int_E \phi$ and f is Lebesgue integrable.

LXVI. **Riemann integrability.** ***Prop. 1.2.4.*** A bounded function f on $[a, b]$ is Riemann integrable $\iff$ the set of points at which f is discontinuous on $[a, b]$ have zero measure.

Proof: not complete.

LXVII. **Lebesgue integrals of nonnegative functions.** Define the Lebesgue integral of a nonnegative measurable function f on measurable set E as

$$\int_E f = \sup\left\{\int_E \psi : \text{ any simple nonnegative } \psi, \psi \leq f\right\}.$$

Equivalently, the Lebesgue integral is $\int_E f = \lim_{n\to\infty} \int_E \psi_n$ for any sequence $\{\psi_n\}$ of simple nonnegative functions such that $\psi_n \uparrow f$ on E.

As $\{\int_E \psi_n\}$ is an increasing sequence, its limit exists on $\overline{\mathbb{R}}$ and equal to $\sup_n \int_E \psi_n$.

LXVIII. **Creation of an increasing sequence of simple functions.** This sequence of simple functions, $\{\psi_n\}$, is increasing and always $\leq f$:

$$\psi_n(x) = \begin{cases} j2^{-n} & j2^{-n} \leq f(x) \leq (j+1)2^{-n} \quad j = 0, \ldots, n2^n - 1 \\ n & \text{otherwise} \end{cases}$$

LXIX. **Properties of nonnegative measurable functions.** ***Prop. 1.2.5.***

(1) $\int_E (af + bg) = a\int_E f + b\int_E g$, if $a, b > 0$.

Define increasing sequences of nonnegative simple functions $\{\psi_{f,n}\}$ and $\{\psi_{g,n}\}$ such that $\psi_{f,n} \uparrow f$ and $\psi_{g,n} \uparrow g$. Note that $a\psi_{f,n} + b\psi_{g,n}$ is also a simple function and that $a\psi_{f,n} + b\psi_{g,n} \uparrow af + bg$.

$$\begin{aligned}
\int_E af + bg &= \lim_{n\to\infty} \int_E (a\psi_{f,n} + b\psi_{g,n}) && \text{(definition of Lebesgue integral)} \\
&= \lim_{n\to\infty} \left[a\int_E \psi_{f,n} + b\int_E \psi_{g,n}\right] && \textit{(Prop. 1.2.1)} \\
&= a\int_E f + b\int_E g && \text{(definition of Lebesgue integral)}
\end{aligned}$$

(2) if $f \geq g$ a.e. then $\int_E f \geq \int_E g$.

Define increasing sequences of nonnegative simple functions $\{\psi_{f,n}\}$ and $\{\psi_{g,n}\}$ such that $\psi_{f,n} \uparrow f$ and $\psi_{g,n} \uparrow g$ with the added constraint that $\psi_{f,n} \geq \psi_{g,n}$ for all n. Using *1.2.1*, $\psi_{f,n} \geq \psi_{g,n}$ implies $\int_E \psi_{f,n} \geq \int_E \psi_{g,n}$ for all n. As these sequences are increasing, their limits exist, and $\lim_{n\to\infty} \int_E \psi_{f,n} \geq \lim_{n\to\infty} \int_E \psi_{g,n}$. Using the definition of the Lebesgue integral for nonnegative sequences of functions, $\Rightarrow \int_E f \geq \int_E g$.

(3) if $f = g$ a.e. then $\int_E f = \int_E g$.

Proof: Apply the previous property to both $f \geq g$ a.e. and $g \geq f$ a.e., both of which are true.

LXX. **Nonnegative integral of zero.** Suppose f is a nonnegative measurable function. $\int f = 0 \iff f = 0$ a.e.

Proof:

"$\Leftarrow$" Let $A = \{x : f(x) > 0\}$. Suppose $f = 0$ a.e. $\Rightarrow m(A) = 0$. Note that $f(x) = \mathbb{I}_A(x)$ a.e. $\int f = \int \mathbb{I}_A$ using *1.2.5*. Then $\int f = 1\, m(A) = 0$, using the definition of Lebesgue integration on simple functions.

"$\Rightarrow$" Suppose $\int f = 0$. Again let $A = \{x : f(x) > 0\}$. Define $A_n = \{x : f(x) > 1/n\}$, such $A = \bigcup_{n=1}^{\infty} A_n$. Then $\int_{A_n} 1/n \leq \int_{A_n} f \leq \int f = 0$ using the facts $1/n \leq f$ on A_n and the integral over A_n is less than the integral over $\mathbb{R}$. As $0 \leq \int_{A_n} 1/n \leq 0$, $0 = \int_{A_n} 1/n = m(A_n)/n \Rightarrow m(A_n) = 0$ for all n. As $m(A) \leq \sum_{n=1}^{\infty} m(A_n)$ by countable subadditivity, $m(A) \leq 0 \Rightarrow m(A) = 0$. As f is nonnegative and $f > 0$ on a measure of zero, $\Rightarrow f = 0$ a.e.

LXXI. **Fatou's Lemma.** ***Thrm. 1.2.6.*** If $\{f_n\}$ is a sequence of nonnegative measurable functions and $f_n \to f$ a.e. on a measurable set E, then

$$\int_E \lim_{n\to\infty} f_n = \int_E f \leq \underline{\lim}_n \int_E f_n.$$

Proof: The equality comes from the definition of a limit.

As integrals over sets of zero measure are zero, separate E into A and A^c where $A = \{x : f_n(x) \to f(x)\}$ and $m(A^c) = 0$. Then $\int_E f = \int_A f + \int_{A^c} f = \int_A f$ and $\underline{\lim} \int_E f_n = \underline{\lim} \int_A f_n + \underline{\lim} \int_{A^c} f_n = \underline{\lim} \int_A f_n$.

Define sequence $\{h_n(x) = \inf_{j\geq n} f_j(x)\}$. As $f_n \to f$, the limit of f_n exists and equals $\underline{\lim}\, f_n$, such that $\sup_n h_n(x) = \underline{\lim}\, f_n(x) = f(x)$. $\Rightarrow h_n(x) \uparrow \sup_n h_n(x) = f(x)$ as $\{h_n(x)\}$ is an increasing sequence.

Define simple functions

$$\psi_{n,m} = \begin{cases} j2^{-m} & j2^{-m} \leq h_n(x) < (j+1)2^{-m} \quad j = 0, \dots, m2^m - 1 \\ m & \text{otherwise} \end{cases}$$

Note that $\psi_{n,m}(x) \leq \psi_{n,m+1}(x)$ for all n and $\psi_{n,m}(x) \leq \psi_{n+1,m}(x)$ for all m. Accordingly, $\psi_{n,n}(x) \leq \psi_{n+1,n+1}(x)$ for all n.

Redefine $\psi_n(x) = \psi_{n,n}(x)$ as an increasing sequence of simple functions with $\psi_n \uparrow f$. By definition of Lebesgue integration over nonnegative functions, $\int_A f = \lim_{n\to\infty} \int_E \psi_n$. As $\{\int_A \psi_n\}$ is an increasing sequence, its limit exists, and thus $\lim_{n\to\infty}\int_A \psi_n = \underline{\lim}\int_A \psi_n$. As $\psi_n(x) \le h_n(x) \le f_n(x)$, using *1.2.5*, $\Rightarrow \int_A \psi_n \le \int_A f_n$ for all n. As the limits of $\{\psi_n\}$ and $\{f_n\}$ exist and the limit of a sequence is equal to its limit inferior when its limit exists, it follows that $\underline{\lim}\int_A \psi_n \le \underline{\lim}\int_A f_n$.

$$\Rightarrow \int_A f_n = \underline{\lim}\int_A \psi_n \le \underline{\lim}\int_A f_n.$$

LXXII. **Monotone Convergence Theorem.** ***Thrm. 1.2.7.*** If $\{f_n\}$ is an increasing sequence of nonnegative measurable functions and $f_n \uparrow f$ a.e., then $\int \lim f_n = \int f = \lim \int f_n$.

Proof: As $f_n \le f$ for all n, $\Rightarrow \int f_n \le \int f$ for all n using *1.2.5*. The limit of an increasing sequence exists and is equal to its limit superior, thus $\overline{\lim}\int f_n = \lim \int f_n \le \int f$. As $f_n \to f$, using Fatou's Lemma (*1.2.6*) $\Rightarrow \int f \le \underline{\lim}\int f_n$. As the limit inferior of a sequence is less than its limit superior

$$\int f \le \underline{\lim}\int f_n = \lim_{n\to\infty}\int f_n = \overline{\lim}\int f_n \le \int f,$$

thus $\int f = \lim \int f_n$.

LXXIII. **Infinite sums of nonnegative functions.** ***Prop. 1.2.8.*** Let $f = \sum_{i=1}^{\infty} h_i$. Then $\int f = \sum_{i=1}^{\infty}\int h_i$.

Proof: f_n is an increasing sequence of nonnegative measurable functions with $f_n \uparrow f$. Applying MCT (*1.2.7*), $\Rightarrow \inf f = \lim \int f_n$. As $\int f_n = \int \sum_{i=1}^{n} h_i$, and for finite sums, application of *1.2.5* yields $\Rightarrow \sum_{i=1}^{n}\int h_i$. All together,

$$\int f = \lim \int f_n = \lim_{n\to\infty}\sum_{i=1}^{n}\int h_i = \sum_{i=1}^{\infty}\int h_i.$$

LXXIV. **Integration over disjoint sets.** ***Prop. 1.2.9.*** Given nonnegative measurable function f and a collection of disjoint measurable sets $\{E_i\}$ with $E = \bigcup_{i=1}^{\infty} E_i$. Then,

$$\int_E f = \sum_{i=1}^{\infty}\int_{E_i} f.$$

Proof: Let $h_i = f\,\mathbb{I}_{E_i}$ and define $f_n = \sum_{i=1}^{n} h_i$. As $\{f_n\}$ is a sequence of nonnegative measurable functions and $f = \sum_{i=1}^{\infty} h_i$, applying *1.2.8* and a property of integration domain yields

$$\int_E f = \sum_{i=1}^{\infty}\int_E h_i = \sum_{i=1}^{\infty}\int_E f\,\mathbb{I}_{E_i} = \sum_{i=1}^{\infty}\int_{E_i} f.$$

LXXV. **Integrability of nonnegative measurable functions.** A nonnegative measurable function f is *integrable over* E if E is a measurable set and $\int_E f < \infty$.

LXXVI. **Integral subtraction of integrable functions.** ***Prop. 1.2.10.*** For nonnegative measurable functions f and g, (where g cannot assume values of ∞ such that $f-g$ is well-defined and measurable), if f is integrable on a measurable set E and $f(x) - g(x) \ge 0$, then $f-g$ is integrable on E and

$$\int_E (f-g) = \int_E f - \int_E g.$$

Proof: $\infty > \int_E f = \int_E (g + f - g) = \int_E g + \int_E (f - g)$, where the last equality is using *1.2.5* as g is nonnegative and measurable. As $\int_E g > 0$ and $\int_E (f-g) > 0$ and their sum is finite, $\Rightarrow \int_E (f-g) < \infty$, and thus $f - g$ is integrable on E.

LXXVII. **Delta-epsilon arguments for integrable functions.** ***Prop. 1.2.11.*** Given a nonnegative measurable function f that is integrable over E. Then given $\epsilon > 0$, there exists some $\delta > 0$ such that for every measurable set $A \subset E$ with $m(A) < \delta$, it is true that $\int_A f < \epsilon$.

Proof: Let $f_n(x) = f(x)$ if $f(x) \leq n$ and n otherwise. Thus it is nonnegative and bounded by n.

$f_n \uparrow f$ and $f - f_n \geq 0$. This implies, using *1.2.10*, $\int_E (f - f_n) = \int_E f - \int_E f_n$. As $f_n \uparrow f$, applying MCT (*1.2.7*), $\lim_{n\to\infty} \int_E f_n = \int_E f$. For any $\epsilon > 0$, there exists N such that $|\int_E f_n - \int_E f| < \epsilon/2$ for all $n \geq N$. Thus $\int_E (f - f_N) < \epsilon/2$.

Pick $\delta = \epsilon/2N$. If $A \subset E$ and $m(A) < \delta$. For any A $\int_A f_N < \int_A N$ using *1.2.5*. The latter equals $N\, m(A)$, so $\int_A f_N < \epsilon/2$.

From *1.2.5* for the first equality,

$$\int_A f = \int_A (f - f_N) + \int_A f_N \leq \int_E (f - f_N) + N\, m(A) < \epsilon.$$

4 General Lebesgue Integrals

LXXVIII. **General Lebesgue integrability.** A measurable function f is said to be *integrable* on a measurable set E if $|f|$ is integrable on E. Define f^+ and f^- such $\int_E f = \int_E f^+ + \int_E f^-$.

LXXIX. **Properties of general Lebesgue integrals.** ***1.2.12.*** For integrable functions f and g on some measurable set E.

(1) $\int_E (af + bg) = a \int_E f + b \int_E g < \infty \;\; \forall a, b \in \mathbb{R}$.

Proof: First, establish that $\int cf = c \int f$ for both cases $c \geq 0$ and $c < 0$. Second, establish that $\int_E (f_1 - f_2) = \int_E f_1 - \int_E f_2$ for both nonnegative and integrable functions f_1 and f_2. $\int_E (af + bg) = \int_E (af)^+ + \int_E (bg)^+ - \int_E (af)^- - \int_E (bg)^-$. Using the above properties, $\Rightarrow \int_E (af) + \int_E (bg) \Rightarrow a \int f + b \int g$.

(2) if $f \geq g$ a.e. then $\int_E f \geq \int_E g$.

Proof: As $f^+ \geq g^+$ a.e. and $f^- \leq g^-$ a.e., $\int_E f = \int_E f^+ - \int_E f^- \geq \int_E g^+ - \int_E g^- = \int_E g$.

(3) if $f = g$ a.e. then $\int_E f = \int_E g$.

Proof: Apply the previous property to both $f \geq g$ a.e. and $g \geq f$ a.e., both of which are true.

(4) $|\int_E f| \leq \int_E |f|$.

(5) for disjoint measurable sets A and B with $A \cup B \subset E$, then $\int_{A\cup B} f = \int_A f + \int_B f$.

LXXX. **Extended Dominated Convergence Theorem.** ***1.2.13.*** Let sequence of measurable functions $\{f_n\}$ converge to f a.e. while $|f_n| \leq g_n$ a.e. on measurable set E. Let sequence of nonnegative measurable functions $\{g_n\}$ converge to integrable function g a.e. If $\lim \int_E g_n = \int_E g$, then (1) f is integrable on E, (2) $\lim \int_E f_n = \int_E f$, and (3) $\lim \int_E |f_n - f| = 0$.

Proof: Show $m(A) = 0$ where A includes all x such $f_n \nrightarrow f$, $g_n \nrightarrow g$, or $f_n > g_n$ for any n. WLOG have $E = A^c$.

As $f_n \to f \Rightarrow |f_n| \to |f|$. Define sequences $f_n^+ \to f^+$ and $f_n^- \to f^-$. Each converges less than ϵ' as $f_n \to f$ and differences in these sequences are less than differences in the entire functions. Choose $\epsilon' = \epsilon/2$, thus proving if $f_n \to f$ then for any $\epsilon > 0$ there exists N such that $\big||f_n| - |f|\big| < \epsilon$ for all $n \geq N$.

As $|f_n| \leq g_n \ \forall n$ and $g_n \to g$, $|f| \leq g$ such $\int |f| \leq \int g < \infty$ using *1.2.5*, thus by definition f is integrable.

As $\lim \int g_n = \int g$, for any $\epsilon > 0$ there exists N such that $\int g_n < \int g + \epsilon < \infty$ for all $n \geq N$. Thus $\{g_n\}$ are integrable for all $n \geq N$. As $|f_n| \leq g_n$, $\Rightarrow \int |f_n| < \infty$, thus $\{f_n\}$ are all integrable for all $n \geq N$.

Note that, for all $n \geq N$, $\{g_n - f_n\}$ and $\{g_n + f_n\}$ are all measurable, integrable, nonnegative, and converge to $g - f$ and $g + f$, respectively.

By Fatou's Lemma $\int(g + f) \leq \underline{\lim} \int(g_n + f_n)$ and by property of the limit inferior, $\underline{\lim} \int(g_n + f_n) \leq \overline{\lim} \int g_n + \underline{\lim} \int f_n$. Thus

$$\int g + \int f = \int (g + f) \leq \underline{\lim} \int (g_n + f_n) \leq \overline{\lim} \int g_n + \underline{\lim} \int f_n,$$

and as $\int g = \overline{\lim} \int g_n$ as the limit exists, $\Rightarrow \int f \leq \underline{\lim} \int f_n$.

Similarly, using Fatou's Lemma $\int(g - f) \leq \underline{\lim} \int(g_n - f_n)$ and by property of the limit inferior, $\underline{\lim} \int(g_n - f_n) \leq \overline{\lim} \int g_n + \underline{\lim} \int(-f_n) =\leq \overline{\lim} \int g_n - \overline{\lim} \int f_n$, the last equality from property that adding the limit inferior of an inverted sequence is equivalent to subtracting the limit superior of that sequence. Thus

$$\begin{aligned} \int g - \int f = \int (g - f) \leq \underline{\lim} \int (g_n - f_n) \\ \leq \overline{\lim} \int g_n + \underline{\lim} \int (-f_n) = \overline{\lim} \int g_n - \overline{\lim} \int f_n, \end{aligned}$$

and as $\int g = \overline{\lim} \int g_n$ as the limit exists, $\Rightarrow -\int f \leq -\overline{\lim} \int f_n$, $\Rightarrow \overline{\lim} \int f_n \leq \int f$.

As $\underline{\lim} \int f_n \leq \overline{\lim} \int f_n$, so

$$\overline{\lim} \int f_n \leq \int f \leq \underline{\lim} \int f_n \leq \overline{\lim} \int f_n,$$

yielding $\underline{\lim} \int f_n = \overline{\lim} \int f_n = \int f$, thus $\lim \int f_n = \int f$.

LXXXI. **Dominated Convergence Theorem.** ***1.2.14.*** Same as EDCT, but for $f_n \leq g$, $\forall n$ (set $g_n = g$, $\forall n$).

LXXXII. **Bounded Convergence Theorem.** ***1.2.15.*** Same as EDCT, but for finite E ($m(E) < \infty$). Suppose $|f_n| < M$ a.e.

μ Integrable

Step 1: Integrals of Simple Functions

LXXXIII. If each $\mu(E_i) < \infty$ for $i = 1, \ldots, n$, define the $\mu-$integral of ϕ as

$$\int \phi d\mu \equiv \int_\Omega \phi d\mu = \sum_{i=1}^{n} a_i \mu(E_i)$$

Here allowing for the possibility that a_i can be positive or negative since each $\mu(E_i) < \infty$.

Alternatively, if each $a_1, \ldots, a_n$ is nonnegative (so that $\phi(\omega) \geq 0, \omega \in \Omega$), define the $\mu-$ integral of ϕ as

$$\int \phi d\mu \equiv \int_\Omega \phi d\mu = \sum_{i=1}^{n} a_i \mu(E_i)$$

Here allowing for the possiblity that some $\mu(E_i) = \infty$ since all the $a_i \geq 0$.

Step 2: Integrals of Non-negative Measurable Functions

LXXXIV. $\mu-$**integral of** f: Let $(\Omega, \mathcal{F}, \mu)$ be a msp. If $f \geq 0$ is a nonnegative measurable function where $f : \Omega \to \overline{\mathbb{R}}$, define the $\mu-$integral of f as

$$\int f d\mu \equiv \int_\Omega f d\mu = \sup \Big\{ \int \psi d\mu : \text{ any simple, nonnegative } \psi, \psi \leq f \text{ on } \Omega \Big\}$$

Alternatively and equivalently, we can define the $\mu-$**integral of** f as

$$\int_\Omega f = \lim_{n\to\infty} \int_\Omega \psi_n$$

for any arbitrary sequence of simple, nonnegative functions $\psi_n(\omega) \geq 0$ such that $\psi_n(\omega) \uparrow f(\omega)$ as $n \to \infty$ for $\omega \in \Omega$.

LXXXV. **Example:** Suppose $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, $\mu(A) = |A \cap \mathbb{N}|$, $A \in \mathcal{F}$. Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = 3^{-x}$ for $x = 1, 2, 3, \ldots$ and $f(x) = 0$ otherwise. Find $\int f d\mu$.

$$\psi_n(x) = \sum_{k=1}^{n} 3^{-k} \mathbb{I}_k(x) \uparrow \sum_{k=1}^{\infty} 3^{-k} \mathbb{I}_k(x) = f(x)$$

Where $\psi_n(x)$ is a simple function and clearly $\psi_n(x) \leq \psi_{n+1}(x) \leq f(x)$. Thus

$$\int f d\mu = \lim_{n\to\infty} \int \psi_n d\mu = \lim_{n\to\infty} \Big(\sum_{k=1}^{n} 3^{-k} \Big) = \sum_{k=1}^{\infty} 3^{-k} = 1/2$$

Note: If $f : \Omega \to \overline{\mathbb{R}}$ is nonnegative measurable, then $\int f d\mu = 0$ iff $f = 0$ a.e. (μ).

Note: Suppose $(\Omega, \mathcal{F}, \mu)$ is a msp and $f : \Omega \to \overline{\mathbb{R}}$ is a nonnegative $\mu-$integrable function. then, $f < \infty$ a.e. (μ) or $\mu(\{\omega \in \Omega : f(\omega) = \infty\}) = 0$.

LXXXVI. **Fatou and MCT:** Let $(\Omega, \mathcal{F}, \mu)$ be a msp. Then Fatou's lemma and the Monotone convergence theorem are defined as before except that we have $f_n : \Omega \to \overline{\mathbb{R}}$.

LXXXVII. **Example:** Let $(\Omega, \mathcal{F}, \mu)$ be a msp and $f : \Omega \to \overline{\mathbb{R}}$ be a nonnegative measurable function. Show that the set function $\nu(A) \equiv \int_A f d\mu, \;\; A \in \mathcal{F}$, defines a measure on $(\Omega, \mathcal{F})$. Proof:

(1) $\nu(A) \geq 0$

(2) $\nu(\emptyset) = 0$

(3) for $B_i, B_2, \ldots \in \mathcal{F}$ such that $\{B_i\}_{i\geq 1}$ are disjoint,

$$\nu(\bigcup_{i=1}^{\infty} B_i) = \int_{\left(\bigcup_{i=1}^{\infty} B_i\right)} f d\mu = \sum_{i=1}^{\infty} \int_{B_i} f d\mu = \sum_{i=1}^{\infty} \nu(B_i) \;\; \blacksquare$$

Step 3: General Integrals

LXXXVIII. **Nonnegative function f:** For a function $f : \Omega \to \overline{\mathbb{R}}$, let $f^+(\omega) = f(\omega)\mathbb{I}_{\{f\geq 0\}}(\omega)$ and $f^-(\omega) = -f(\omega)\mathbb{I}_{\{f<0\}}(\omega)$. Then,

$$f(\omega) = f^+(\omega) - f^-(\omega) \qquad |f(\omega)| = f^+(\omega) + f^-(\omega), \qquad \omega \in \Omega$$

LXXXIX. $\mu-$**integrable**: Suppose $(\Omega, \mathcal{F}, \mu)$ is a msp. A measurable function $f : \Omega \to \overline{\mathbb{R}}$ is said to be $\mu-$**integrable** if $|f|$ is $\mu-$integrable. Then we define

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu$$

Note: $|f|$ is $\mu-$ integrable iff both f^+ and f^- and $\mu-$integrable.

XC. L^p **spaces**. Suppose $(\Omega, \mathcal{F}, \mu)$ is a msp and $0 < p \leq \infty$. Then, the L^p **space** of $(\Omega, \mathcal{F}, \mu)$, denoted as $L^p(\Omega, \mathcal{F}, \mu)$ is a collection of measurable functions $f : \Omega \to \overline{\mathbb{R}}$ defined as:

$$L^p(\Omega, \mathcal{F}, \mu) \equiv \big\{f : |f|^p \text{ is } \mu\text{−integrable}\big\} = \big\{f : \int |f|^p d\mu < \infty\big\}$$

for $0 < p < \infty$, and

$$\begin{aligned} L^\infty(\Omega, \mathcal{F}, \mu) &\equiv \big\{f : f \text{ is bounded a.e. } (\mu)\big\} \\ &= \big\{f : \mu(\{\omega \in \Omega : |f(\omega)| > M\}) = 0 \text{ for some } M \equiv M(f) > 0\big\} \end{aligned}$$

If f is $\mu-$integrable then $f \in L^1(\Omega, \mathcal{F}, \mu)$

XCI. **Scheffe's Theorem.** Let $(\Omega, \mathcal{F}, \mu)$ be a msp and, for n ≥ 1, let $\nu_n(A) = \int_A f_n d\mu, A \in \mathcal{F}$, 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 $f_n \to f_0$ a.e. (μ), then

$$\lim_{n\to\infty} \int |f_n - f_0| d\mu = 0$$

Additionally, $\sup_{A\in\mathcal{F}} |\nu_n(A) - \nu_0(A)| \to 0$ as $n \to \infty$.

XCII. **Example:** Suppose $X_n \sim \text{Poisson}(\lambda_n), n \geq 0$. If $\lambda_n \to \lambda_0$, then $\sum_{x=0}^{\infty} |P(X_n = x) - P(X_0 = x)| \to 0$.

Proof: $\Omega = \mathbb{R}, \mathcal{F} = \mathcal{B}(\mathbb{R}), \mu(A) = |A \cap \mathbb{Z}^+|, A \in \mathcal{F}$.

note: $P(x \in A) = \nu_n(A) = \int_A f_n d\mu = \sum_{x \in A \cap \mathbb{Z}^+} f_n(x)$, Where

$$f_n(x) = \begin{cases} \frac{e^{-\lambda_n}\lambda_n^x}{x!} & \text{if } x \in \mathbb{Z}^+ \\ 0 & \text{else} \end{cases}$$

Remember since $f_n, n \geq 1$ are densities, $\nu_n(\Omega) = \nu_0(\Omega) = 1$. Then for $x \in \mathbb{Z}^+, \lambda_n \to \lambda_0 \Rightarrow \frac{e^{-\lambda_n}\lambda_n^x}{x!} \to \frac{e^{-\lambda_0}\lambda_0^x}{x!}$. that is $f_n \to f_0$ a.e. (μ). Therefore, by Scheffe's theorem,

$$0 = \lim_{n\to\infty} \int |f_n - f_0| d\mu = \lim_{n\to\infty} \sum_{x=0}^{\infty} |f_n(x) - f_0(x)|$$

XCIII. **Uniform Integrability.** Let $(\Omega, \mathcal{F}, \mu)$ be a msp. Then a family of $\mu-$integrable functions $\{f_\lambda : \lambda \in \Lambda\}$ is called **Uniformly Integrable** or (UI) with respect to μ if

$$\lim_{t\to\infty} \sup_{\lambda \in \Lambda} \int_{|f_\lambda|>t} |f_\lambda| d\mu = 0$$

XCIV. **UI Conditions.** Let $\mathcal{A} \equiv \{f_\lambda : \lambda \in \Lambda\}$ be a collection of $\mu-$integrable functions on a msp $(\Omega, \mathcal{F}, \mu)$. then,

- (i) if $\Lambda = \{\lambda_1, \ldots, \lambda_k\}$ for some $1 \leq k < \infty$, then $\mathcal{A}$ is UI.
- (ii) if $\sup\{\int |f_\lambda|^{1+\epsilon} d\mu : \lambda \in \Lambda\} < \infty$ for some $\epsilon > 0$, then $\mathcal{A}$ is UI.
- (iii) if $|f_n| \leq |f| a.e.(\mu)$ and $\int f d\mu < \infty$, the $\mathcal{A}$ is UI.
- (iv) if $\mathcal{A}$ is UI and $\mu(\Omega) < \infty$, then $\sup\{\int |f_\lambda| d\mu : \lambda \in \Lambda\} < M$ for some $M > 0$.
- (v) if $\{f_\lambda : \lambda \in \Lambda\}$ and $\{g_\lambda : \lambda \in \Lambda\}$ are both UI, then $\{f_\lambda + g_\lambda : \lambda \in \Lambda\}$ is UI.

XCV. **UI theorem for finite measure spaces.** Let $(\Omega, \mathcal{F}, \mu)$ be a msp. with $\mu(\Omega) < \infty$. Suppose $f_n : \Omega \to \overline{\mathbb{R}}$ and $f : \Omega \to \overline{\mathbb{R}}$ are measurable and $f_n \to f$ a.e. (μ). If $\{f_n : n \geq 1\}$ is UI, then f is $\mu-$integrable and

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu \qquad \textit{and} \qquad \lim_{n\to\infty} \int |f_n - f_0| d\mu = 0$$

Remark: By prop 3.12 (iii) (above), This theorem yields the convergence of $\int f_n d\mu$ to $\int f d\mu$ under weaker conditions than the DCT, provided that $\mu(\Omega) < \infty$.

5 Radon Nikodym

XCVI. μ **dominated by** ν. Let $(\Omega, \mathcal{F})$ be a measurable space and let μ and ν be measures on $(\Omega, \mathcal{F})$. The measure μ is said to be **Dominated by** ν **or absolutely continuous** with respect to ν denoted $\mu \ll \nu$ if

$$\text{if } A \in \mathcal{F} \text{ and } \nu(A) = 0 \quad \Rightarrow \quad \mu(A) = 0$$

XCVII. **Example.** $(\Omega = \mathbb{R}, \mathcal{F} = \mathcal{B}(\mathbb{R}))$ and let $\mu =$ standard normal probability for $A \in \mathcal{B}(\mathbb{R})$ and let $\nu =$ Lebesgue measure. For $A \in \mathcal{B}(\mathbb{R}), \nu(A) = 0 \Rightarrow \mu(A) = 0$ so $\mu \ll \nu$. Similarly, $\mu(A) = 0 \Rightarrow \nu(A) = 0$ so $\nu \ll \mu$.

XCVIII. **Radon Nikodym Derivative.** Let μ and ν be measures on a measurable space $(\Omega, \mathcal{F})$ and let h be a nonnegative measurable (extended real-valued) function such that

$$\mu(A) = \int_A h d\nu \qquad \text{for all } A \in \mathcal{F}$$

Then h is call the **density** or the **Radon-Nikodym derivative** of μ with respect to ν and $h = \frac{d\mu}{d\nu}$.

XCIX. **Radon Nikodym Theorem.** Let μ and ν be measures on a measurable space $(\Omega, \mathcal{F})$ such that ν is a $\sigma-$finite measure and μ is absolutely continuous with respect to ν. (i.e. $\mu \ll \nu$). Then

(i) there is a nonnegative measurable (extended real valued) function h such that

$$\mu(A) = \int_A h d\nu \qquad \text{for all } A \in \mathcal{F}$$

(ii) The function h is unique in the sense that if g is any other measurable function with the above property then $g = h$ a.e. (ν).

C. **Radon Nikodym Proposition.** Let μ, μ_1, μ_2, and ν be $\sigma-$finite measures on a measurable space $(\Omega, \mathcal{F})$.

(i) If $\mu \ll \nu$ and f is a nonnegative measurable (extended real valued) function, then

$$\int f d\mu = \int f \Big[\frac{d\mu}{d\nu}\Big] d\nu$$

Proof: Let $h = \frac{d\mu}{d\nu}$. for a simple function $\psi = \sum_{i=1}^n a_i \mathbb{I}_{A_i}, \;\; a_i \geq 0$. Check that

$$\int \psi d\mu = \sum_{i=1}^n a_i \mu(A_i) = \sum_{i=1}^n a_i \int \mathbb{I}_{A_i} h d\nu = \int \big(\sum_{i=1}^n a_i \mathbb{I}_{A_i}\big) h d\nu$$

Now $0 \leq \psi_n \uparrow f$,

$$\int f d\mu = \lim_{n \to \infty} \int \psi_n d\mu = \lim_{n \to \infty} \int h \psi_n d\nu$$

where $0 \leq h\psi_n \uparrow hf \Rightarrow$ by MCT

$$\lim_{n \to \infty} \int h \psi_n d\nu = \int hf \; d\nu$$

CI. **Singular measures.** Let $(\Omega, \mathcal{F})$ be a measurable space and let μ and ν be measures on $(\Omega, \mathcal{F})$. The measures μ and ν are called **singular** with respect to each other (denoted $\mu \perp \nu$) if there exist (there only need exist one set) $B \in \mathcal{F}$ such that

$$\mu(B) = 0 \qquad \text{and} \qquad \nu(B^c) = 0$$

That is μ can be positive only on B^c and ν can be positive only on B.

CII. **Lebesgue Decomposition Theorem.** Let μ and ν be $\sigma-$finite measures on a measurable space $(\Omega, \mathcal{F})$. then the measure μ can be uniquely decomposed as

$$\mu = \mu_a + \mu_s$$

Where μ_a, μ_s are $\sigma-$finite measures on $(\Omega, \mathcal{F})$ such that $\mu_a \ll \nu$ and $\mu_s \perp \nu$. By the Radon Nikodym Theorem, There also exists a nonnegative measurable function h (i.e. $h = \frac{d\mu_a}{d\nu}$) such that

$$\mu_a(A) = \int_A h d\nu \qquad \text{for all } A \in \mathcal{F}$$

CIII. **Example.** Find the Lebesgue Decomposition of μ with respect to ν and find $\frac{d\mu_a}{d\nu}$. Let $\mu = N(0,1)$ and let $\nu = \exp(1)$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proof: Define $\mu_s(A) \equiv \mu(A \cap (-\infty, 0))$, then μ_s and ν will be singular $\mu_s \perp \nu$ because for $B = [0, \infty), \mathbb{R} = B \cap B^c$ and $\mu_s(B) = 0$ and $\nu(B^c) = 0$. Thus there exists a set, $B = [0, \infty)$, such that the condition was satisfied. Now

$$\begin{aligned} \mu_a = \mu(A \cap [0,\infty)) &= \int_{A \cap [0,\infty)} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2} dx \\ &= \int_A \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2} e^x f(x) dx \\ \text{(By the R-D theorem)} &= \int_A \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2} e^x d\nu(x) \end{aligned}$$

Where $f(x) = e^{-x} \mathbb{I}_{[0,\infty)}(x)$. So $\frac{d\mu_a}{d\nu} = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2} e^x$.

6 Miscellaneous Measure Theory

CIV. **Product Measures.** Suppose $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are two measurable spaces.

- $\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$ is the product of Ω_1 and Ω_2
- The set $A_1 \times A_2$, where $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$, is called a measurable rectangle.
- The product $\sigma-$algebra is $\mathcal{F}_1 \times \mathcal{F}_2 \equiv \sigma\langle\{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}\rangle$.
- $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ is the product measurable space.

CV. ω**-sections**: Let $A \in \mathcal{F}_1 \times \mathcal{F}_2$, then

(*i*) for any fixed $\omega_1 \in \Omega_1$, the set

$$A_{1\omega_1} \equiv \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\}$$

is called the ω_1-section of A

(*ii*) for any fixed $\omega_2 \in \Omega_2$, the set

$$A_{2\omega_2} \equiv \{\omega_1 \in \Omega_1 : (\omega_1, \omega_2) \in A\}$$

is called the ω_2-section of A

CVI. **Tonelli's Theorem.** Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two $\sigma-$finite measure spaces and let $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ be a non-negative $\langle \mathcal{F}_1 \times \mathcal{F}_2, \mathcal{B}(\mathbb{R})\rangle$-measurable (extended real valued) function. Then,

$$g_1(\omega_1) \equiv \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) : \Omega_1 \to \overline{\mathbb{R}} \text{ is } \langle \mathcal{F}_1, \mathcal{B}(\mathbb{R})\rangle - \text{measurable}$$

$$g_2(\omega_2) \equiv \int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) : \Omega_2 \to \overline{\mathbb{R}} \text{ is } \langle \mathcal{F}_2, \mathcal{B}(\mathbb{R})\rangle - \text{measurable}$$

$$0 \leq \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2) = \int_{\Omega_1} g_1 d\mu_1 = \int_{\Omega_2} g_2 d\mu_2 = \quad \text{nonnegative}$$

Again, if $A \in \mathcal{F}_1 \times \mathcal{F}_2, f(\omega) = \mathbb{I}_A(\omega), \omega = (\omega_1, \omega_2) \Rightarrow g_1(\omega_1) = \mu_2(A_{1\omega_1}),$ and $g_2(\omega_2) = \mu_1(A_{2\omega_2})$. This result is prop 4.2 in notes

CVII. **Fubini's Theorem.** Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be two $\sigma-$finite measure spaces and let $f \in \mathcal{L}^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$, then

(i) $g_1(\omega_1) \equiv \int_{\Omega_2} |f(\omega_1, \omega_2)| d\mu_2(\omega_2) : \Omega_1 \to \overline{\mathbb{R}}$ is $\langle \mathcal{F}_1, \mathcal{B}(\mathbb{R})\rangle -$measurable, μ_1-integrable so that $B_1 \equiv \{\omega_1 \in \Omega_1 : g_1(\omega_1) = \infty\} \in \mathcal{F}_1$ and $\mu_1(B_1) = 0$.

(ii) $g_2(\omega_2) \equiv \int_{\Omega_1} |f(\omega_1, \omega_2)| d\mu_1(\omega_1) : \Omega_2 \to \overline{\mathbb{R}}$ is $\langle \mathcal{F}_2, \mathcal{B}(\mathbb{R})\rangle -$measurable, μ_2-integrable so that $B_2 \equiv \{\omega_2 \in \Omega_2 : g_2(\omega_2) = \infty\} \in \mathcal{F}_2$ and $\mu_2(B_2) = 0$.

(iii) Additionally,

$$h_1(\omega_1) \equiv \mathbb{I}_{B_1^c}(\omega_1) \cdot \int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) : \Omega_1 \to \overline{\mathbb{R}} \text{ is } \langle \mathcal{F}_1, \mathcal{B}(\mathbb{R})\rangle - \text{measurable }, \mu_1 - \text{integrable}$$

$$h_2(\omega_2) \equiv \mathbb{I}_{B_2^c}(\omega_2) \cdot \int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) : \Omega_2 \to \overline{\mathbb{R}} \text{ is } \langle \mathcal{F}_2, \mathcal{B}(\mathbb{R})\rangle - \text{measurable }, \mu_2 - \text{integrable}$$

and

$$\int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2) = \int_{\Omega_1} h_1 d\mu_1 = \int_{\Omega_2} h_2 d\mu_2$$

CVIII. **Absolute Continuity and Bounded Variation.** Two properties of functions, called "bounded variation" and "Absolute continuity", are useful to know. These, in turn, can be connected to integrals with respect to the the Lebesgue measure. This will take us to the "Fundamental Theorem of Lebesgue Integral Calculus" with respect to Lebesgue measure m.

CIX. **monotonely increasing**. A function $f : [a, b] \to \mathbb{R}$ is **monotonely increasing** if $f(x) \leq f(y)$ for all $x \leq y$.

CX. **Proposition.** Suppose $f : [a, b] \to \mathbb{R}$ is an increasing real-valued function on the interval $[a, b]$. Then f is differentiable a.e.(m) on $[a, b]$ and f' is nonnegative measurable with

$$\int_a^b f'(x) dx \leq f(b) - f(a)$$

CXI. **Functions of Bounded Variation.** Take a real valued function $f : [a, b] \to \mathbb{R}$ on an interval $[a, b]$ and let $a = x_0 < x_1 < \cdots < x_n = b$ be any division of $[a, b]$.

Define

$$p_a^b = \sum_{i=1}^{n} \big[f(x_i) - f(x_{i-1})\big]^+ \qquad n_a^b = \sum_{i=1}^{n} \big[f(x_i) - f(x_{i-1})\big]^+$$

$$t_a^b = \sum_{i=1}^{n} \big|f(x_i) - f(x_{i-1})\big| = p_a^b + n_a^b$$

Set

$$\begin{aligned} P_a^b(f) &= \sup\{p_a^b : p_a^b \text{ from some possible division of } [a, b]\} \\ N_a^b(f) &= \sup\{n_a^b : n_a^b \text{ from some possible division of } [a, b]\} \\ T_a^b(f) &= \sup\{t_a^b : t_a^b \text{ from some possible division of } [a, b]\} \end{aligned}$$

CXII. **Bounded Variation.** We say that a real-valued function $f : [a, b] \to \mathbb{R}$ on $[a, b]$ is of **bounded variation** over $[a, b]$ if $T_a^b(f) < \infty$.

CXIII. **Note.** if f is an increasing function, then f is of bounded variation.

CXIV. **Differentiability of functions of bounded variation.** If a function $f : [a, b] \to \mathbb{R}$ is of bounded variation over $[a, b]$, then f is differentiable a.e.(m) on $[a, b]$.

Proof: By definition, $N_a^x(f) - P_a^x(f) = f(x) - f(a)$. Therefore, $f(x) = f(a) + P_a^x(f) - N_a^x(f)$. Since both $P_a^x(f)$ and $N_a^x(f)$ are nondecreasing functions, the result follows from the previous proposition.

CXV. **Absolute Continuity.** A real-valued function $f : [a, b] \to \mathbb{R}$ on $[a, b]$ is said to be **absolutely continuous** on $[a, b]$ if, given $\epsilon > 0$, there is some $\delta > 0$ where

$$\sum_{i=1}^{n} |f(d_i) - f(c_i)| < \epsilon \quad \text{holds}$$

for any collection of intervals $[c_1, d_1], \dots, [c_n, d_n] \subset [a, b]$ (n is arbitrary) where $(c_1, d_1), \dots, (c_n, d_n)$ are disjoint and

$$\sum_{i=1}^{n} |d_i - c_i| < \delta$$

CXVI. **Proposition.** If a function $f : [a, b] \to \mathbb{R}$ is absolutely continuous on $[a, b]$, then f is of bounded variation on $[a, b]$ (and hence f is differentiable a.e.(m) on $[a, b]$ by a previous proposition).

Consider a closed interval for f between $[a, b]$. Let $\epsilon = 1$. Then there exists k large enough such that

$$\sum_{i=1}^{n} |d_i - c_i| < \frac{b-a}{k} \quad \Rightarrow \quad \sum_{i=1}^{n} |f(d_i) - f(c_i)| < 1$$

Let $a = x_0 < x_1 < \cdots < x_n = b$ be a partition of $t_a^b = \sum_{i=1}^m |f(x_i) - f(x_{i-1})|$. Suppose $x_0, x_1, \ldots, x_k \in [a, a+\delta]$,

$$\sum_{i=1}^{k} |f(x_i) - f(x_{i-1})| < 1 \quad \Rightarrow \quad \sum_{i=1}^{m} |f(x_i) - f(x_{i-1})| < k$$

Therefore $T_a^b(f) \leq k$.

CXVII. **Example.** Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation. Show that f is continuous except on an at most countable set.

Proof: If $g : [a,b] \to \mathbb{R}$ is increasing, then let $A = \{x \in (a,b) : g_-(x) < g_+(x)\}$. Let $B_n = \{x \in A : g_+(x) - g_-(x) > \frac{1}{n}\}, n \geq 1$. Then $A = \cup_{n=1}^{\infty} B_n$. Fix n. Let $N \in \mathcal{N}$ be such that $N/n > g(b) - g(a) \geq 0$. Then B_n must have less than N elements. To see this, suppose if possible that B_n has N or more elements. Then there exists distinct $x_1, \ldots, x_N \in B_N$. Let $y_0 = \frac{a+x_1}{2}, y_i = \frac{x_i + x_{i+1}}{2}, y_N = \frac{x_N + b}{2}$, $i = 1, \ldots, N-1$. So

$$\begin{aligned} g(y_N) - g(y_0) &= \sum_{i=1}^{N} g(y_i) - g(y_{i-1}) \\ &\geq \sum_{i=1}^{N} g_+(x_i) - g_-(x_i) \\ &> \frac{N}{n} \geq g(b) - g(a) \end{aligned}$$

Where $x_0 = a$ since $y_{i-1} < x_i < y_i \Rightarrow g(y_{i-1}) \leq g_-(x_i) \leq g_+(x_i) \leq g(Y_i)$ thus completing the contradiction. Therefore, B_n is at most finite for all $n \geq 1$ and $A = \cup_{n=1}^{\infty} B_n$ is at most countable. If f is of bounded variation on $[a,b]$, then $f(x) = P_a^x - N_a^x + f(a), x \in [a,b]$ where P_a^x and N_a^x are increasing real valued functions of x. Therefore, P_a^x and N_a^x have at most a countable number of discontinuities. Hence f has at most countably many discontinuities.

CXVIII. **Lebesgue Fundamental Theorem of Calculus.** *i*): For a Lebesgue integrable function f on $[a,b]$, suppose another function $F(x), x \in [a,b]$, can be written as

$$F(x) = \int_a^x f(t)dt + F(a)$$

Then, F is absolutely continuous (and of bounded variation) on $[a,b]$ and $F'(x) = f(x)$ a.e.(m) on $[a,b]$.

ii.): On the other hand, if F is an absolutely continuous function on $[a,b]$, then there exists some (Lebesgue) integrable function f on $[a,b]$ where

$$F(x) = \int_a^x f(t)dt + F(a), \quad x \in [a,b]$$

where $F'(x) = f(x)$ a.e.(m) on $[a,b]$.

7 Inequalities

CXIX. **Markov's Inequality**

(1) Suppose f is a nonnegative (extended real valued) measurable function on a msp $(\Omega, \mathcal{F}, \mu)$. Then

$$\mu\big(\{\omega \in \Omega : f(\omega) \geq t\}\big) \leq \frac{\int f d\mu}{t}$$

(2) Let X be a random variable on a psp $(\Omega, \mathcal{F}, P)$. Then for any $r \in \mathbb{R}$ and $t > 0$

$$P\big(|X| > t\big) \leq \frac{E\big(|X|^r\big)}{t^r}$$

(3) If $\phi : (0, \infty) \to (0, \infty)$ is non-decreasing, then

$$P\big(|X| > t\big) \leq \frac{E\big(\phi(|X|)\big)}{\phi(t)}$$

CXX. **Chebychev's Inequality.** Let X be a random variable with $E(X^2) < \infty, E(X) = \mu$ and $\text{var}(X) = \sigma^2$. Then for any real $k > 0$,

$$P\big(|X - \mu| > k\sigma\big) \leq \frac{1}{k^2}$$

CXXI. **Convex function.** For $-\infty \leq a < b \leq \infty$, a real valued function $\phi : (a, b) \to \mathbb{R}$ is said to be **convex** if, for all $0 \leq \lambda \leq 1$ and $a < x \leq y < b$,

$$\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y)$$

CXXII. **Jensen's Inequality.** Let X be a random variable on a psp $(\Omega, \mathcal{F}, P)$ and let $\phi : (-\infty, \infty) \to \mathbb{R}$ be convex. Then,

$$\phi\big(EX\big) \leq E\big(\phi(X)\big)$$

provided that $E(X) < \infty$ and $E\big(\phi(X)\big) < \infty$.

CXXIII. **Holder's Inequality.** Let $(\Omega, \mathcal{F}, \mu)$ be a msp. For $1 < p < \infty$ and $q = p/(p - 1)$, suppose $f \in L^p(\Omega, \mathcal{F}, \mu)$, Then

$$||fg||_1 = \int |fg| d\mu \leq ||f||_p ||g||_q = \Big(\int |f|^p d\mu\Big)^{1/p} \Big(\int |g|^q d\mu\Big)^{1/q}$$

Additionally, if $||fg||_1 \neq 0$, equality holds iff $|f| = c|g|$ a.e.(μ) for some $0 < c < \infty$.

CXXIV. **Cauchy-Schwarz Inequality.** If $f, g \in L^2(\Omega, \mathcal{F}, \mu)$, then

$$||fg||_1 = \int |fg| d\mu \leq ||f||_2 ||g||_2 = \Big(\int |f|^2 d\mu\Big)^{1/2} \Big(\int |g|^2 d\mu\Big)^{1/2}$$

CXXV. **Minkowski's Inequality.** If $f, g \in L^p(\Omega, \mathcal{F}, \mu)$ for $1 \leq p < \infty$, then $f + g \in L^p(\Omega, \mathcal{F}, \mu)$ and

$$\Big(\int |f + g|^p d\mu\Big)^{1/p} = ||f + g||_p \leq ||f||_p + ||g||_p = \Big(\int |f|^p d\mu\Big)^{1/p} + \Big(\int |g|^p d\mu\Big)^{1/p}$$

8 Independence

CXXVI. **The many definitions of independence.** Let $(\Omega, \mathcal{F}, P)$ be a psp. Let I be a set of indices.

(*i*) A collection $A_i, i \in I$, of set in $\mathcal{F}$ are called **independent** if, for any finiste subset of distinct indices $i_1, \ldots, i_k \in I, 1 \leq k \leq \infty$,

$$P\Big(\cap_{j=1}^{k} A_{i_j} \Big) = \prod_{j=1}^{k} P(A_{i_j})$$

(*ii*) Suppose $\mathcal{G}_i \subset \mathcal{F}$ is a collection of measurable sets, for each $i \in I$. Then, the family of those sets $\{\mathcal{G}_i : i \in I\}$ is called **independent** if for any possible collection $\{A_i : i \in I\}$ of sets are independent, where $\{A_i : i \in I\}$ is formed by choosing an arbitrary set A_i from $\mathcal{G}_i$ for each $i \in I$.

(*iii*) A collection of random variables $X_i, i \in I$, on $(\Omega, \mathcal{F}, P)$ are called **independent** if the family $\{\sigma\langle X_i\rangle : i \in I\}$ is independent where

$$\sigma\langle X_i\rangle = \{X_i^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$$

is the $\sigma-$algebra generated by $X - i$. In other words, for any finite sunset of distinct indices $i_1, \ldots, i_k \in I$ and any $B_{i_1}, \ldots, B_{i_k} \in \mathcal{B}(\mathbb{R}), 1 \leq k < \infty$,

$$P(X_{i_1} \in B_{i_1}, \ldots, X_{i_k} \in B_{i_k}) = \prod_{j=1}^{k} P(X_{i_j} \in B_{i_j})$$

CXXVII. **π-class Independence Theorem.** Let $(\Omega, \mathcal{F}, P)$ be a psp. Suppose $\mathcal{G}_i \subset \mathcal{F}$ is a $\pi-$class for each index $i \in I$ and that the family $\{\mathcal{G}_i : i \in I\}$ is independent. Then, the family $\{\sigma\langle\mathcal{G}_i\rangle, i \in I\}$ is independent.

CXXVIII. **Example.** Let $X_i, i \in I$ be r.v.'s on a psp $(\Omega, \mathcal{F}, P)$. then $X_i, i \in I$ are independent iff, for any $x_1, \ldots, x_k \in \mathbb{R}$ and any distinct $i_1, \ldots, i_k \in I, 1 \leq k \leq \infty$,

$$P(X_{i_1} \leq x_1, \ldots, X_{i_k} \leq x_k) = \prod_{j=1}^{k} P(X_{i_j} \leq x_j) \tag{2}$$

Proof: Now, suppose the $X_i, i \in I$ are independent, then (1) holds by definition. Now suppose (1) holds. Define $\mathcal{G}_i = \{\{\omega : X_i(\omega) \leq \alpha\} : \alpha \in \mathbb{R}\} \equiv \{X_i^{-1}\big((-\infty, \alpha]\big) : \alpha \in \mathbb{R}\}$. this is a $\pi-$class since $X_i^{-1}\big((-\infty, x] \cap (-\infty, y]\big) \in \mathcal{G}_i$. Thus from (1) we have that $\{\mathcal{G}_i : i \in I\}$ is independent, and by the theorem above, since $\{\mathcal{G}_i : i \in I\}$ is a $\pi-$class, the family $\{\sigma\langle\mathcal{G}_i\rangle, i \in I\}$ is independent. Thus $\sigma\langle\mathcal{G}_i\rangle = \sigma\langle X_i\rangle \equiv X_i^{-1}(\mathcal{B}(\mathbb{R})) = X_i^{-1}(\sigma\langle\{(-\infty, \alpha] : \alpha \in \mathbb{R}\}\rangle)$. and therefore, $X_i, i \in I$ are independent. ■

CXXIX. **Example.** On a psp $(\Omega, \mathcal{F}, P)$, suppose the collection of sets $A_i, i \in I$ in $\mathcal{F}$ are independent. Show that random variables $X_i(\omega) = \mathbb{I}_{A_i}(\omega), i \in I$ are independent.

Proof: for $\omega \in \Omega$, let

$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

$$\sigma\langle X_i \rangle = \sigma\langle X_i^{-1}\big((-\infty, \alpha)\big) : \alpha \in \mathbb{R}\rangle = \{A_i, A_i^c, \emptyset, \Omega\}$$

(1) $\mathcal{G}_i = \{\emptyset, A_i\}, i \in I$, is a $\pi-$ class.

(2) $\{\mathcal{G}_i : i \in I\}$ is independent since $A_i, i \in I$ are independent. That is for any $i_1, \ldots, i_k \in I$, pick a set $B_j \in \mathcal{G}_{i_j}, j = 1, \ldots, k < \infty$. Then check that $P(B_1 \cap \ldots \cap B_k) = \prod_{j=1}^k P(B_j)$ holds trivially for $B_j = \emptyset$ and otherwise holds by A_i independence. $\Rightarrow \{\sigma\langle \mathcal{G}_i \rangle : i \in I\}$ is independent by above prop (5.1 in notes) and $\sigma\langle \mathcal{G}_i \rangle = \sigma\langle X_i \rangle$ and thus $X_i, i \in I$ are independent. ■

CXXX. **Proposition.** Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\{X_1, \ldots X_k\}, 2 \leq k < \infty$ be a collection of random variables on $(\Omega, \mathcal{F}, P)$.

(1) Then $\{X_1, \ldots X_k\}$ is independent iff

$$E \prod_{i=1}^k h_i(X_i) = \prod_{i=1}^k E h_i(X_i)$$

for all bounded Borel measurable functions $h_i : \mathbb{R} \to \mathbb{R}, i = 1, 2, \ldots, k$.

(2) If X_1, X_2 are independent and $E|X_1| < \infty, E|X_2| < \infty$, then

$$E|X_1 X_2| < \infty \qquad and \qquad E X_1 X_2 = E X_1 E X_2$$

Proof(1): Assume the equation holds. Take $h_i = \mathbb{I}_{B_i}$ with $B_i \in \mathcal{B}(\mathbb{R}), i = 1, 2, \ldots, k$ yields the independence of $\{X_1, \ldots X_k\}$. Conversely, if $\{X_1, \ldots X_k\}$ are independent, then the equation holds for $h_i = \mathbb{I}_{B_i}$ for $B_i \in \mathcal{B}(\mathbb{R}), i = 1, 2, \ldots, k$ and hence for simple functions $\{h_1, \ldots, h_k\}$. The equation is established by applying the BCT.

Proof(2): Note that by the change of variable formula,

$$\begin{aligned} E|X_1 X_2| &= \int_{\mathbb{R}^2} |x_1 x_2| dP_{X_1, X_2}(x_1, x_2) \\ E|X_i| &= \int_{\mathbb{R}} |x_i| dP_{X_i}(x_i) \end{aligned}$$

Where P_{X_1, X_2} is the joint distribution of (X_1, X_2) and P_{X_1} is the marginal distribution of $X_i, i = 1, 2$. Also, by the independence of X_1 and X_2, P_{X_1, X_2} is equal to the product measure $P_{X_1} \times P_{X_2}$. Hence, by Tonelli's theorem,

$$\begin{aligned} E|X_1 X_2| &= \int_{\mathbb{R}^2} |x_1 x_2| dP_{X_1, X_2}(x_1, x_2) \\ &= \int_{\mathbb{R}^2} |x_1 x_2| dP_{X_1}(x_1) P_{X_2}(x_2) \\ &= \left(\int_{\mathbb{R}} |x_1| dP_{X_1}(x_1) \right) \left(\int_{\mathbb{R}} |x_2| dP_{X_2}(x_2) \right) \\ &= E|X_1| E|X_2| < \infty \end{aligned}$$

Using Fubini's theorem finishes the proof.

CXXXI. **Borel Cantelli Lemmas.** Let $(\Omega, \mathcal{F}, P)$ be a psp and $A_n, n \geq$ be sets in $\mathcal{F}$:

(*i*) If $\sum_{s=1}^{\infty} P(A_n) < \infty$, then $P\big(\limsup_{n\to\infty} A_n\big) = 0$.

(*ii*) If $A_n, n \geq 1$ are independent and $\sum_{s=1}^{\infty} P(A_n) = \infty$, then $P\big(\limsup_{n\to\infty} A_n\big) = 1$.

note: (ii) is also true if the events are pairwise independent.

CXXXII. **Example.** Suppose $X_n, n \geq 1$ are independent r.v.'s on some psp $(\Omega, \mathcal{F}, P)$. Show that :

$$P\big(\lim_{n\to\infty} X_n = 0\big) = 1 \quad \text{iff } \forall \epsilon > 0, \quad \sum_{s=1}^{\infty} P(|X_n| > \epsilon) < \infty$$

Proof: $(\Rightarrow)$ Assume $\sum_{n=1}^{\infty} P(|X_n| > \epsilon) < \infty$ for all $\epsilon > 0$. Fix $\epsilon > 0$ and let $A_n = \{\omega : |X_n(\omega)| > \epsilon\}$. By (i) in the Borel Cantelli Lemma, $\sum_{s=1}^{\infty} P(A_n) < \infty$, then $P\big(\limsup_{n\to\infty} A_n\big) = 0$ a.s. (P). This is

$$\begin{aligned} P\Big(\{\omega : |X_n| > \epsilon \;\text{ i.o.}\}\Big) &= 0 \\ \Rightarrow P\Big(\{\omega : |X_n| \leq \epsilon \;\text{ i.o.}\}\Big) &= 1 \end{aligned}$$

Let $B(\epsilon) = \{\omega : |X_n| \leq \epsilon\}$. For $\omega \in B(\epsilon)$, there exists N_ω such that $\big|X_n\big| \leq \epsilon$ for $n \geq N_\omega$. Now for $k \geq 1$, $B(\frac{1}{k+1}) \subset B(\frac{1}{k})$. This implies that

$$\begin{aligned} P\Big(\cap_{k=1}^{\infty} B(\frac{1}{k})\Big) &= \lim_{k\to\infty} P\Big(B\Big(\frac{1}{k}\Big)\Big) \quad \text{by m.c.f.a.} \\ &= \lim_{k\to\infty} 1 \\ &= 1 \end{aligned}$$

note: $\bigcap_{k=1}^{\infty} B(\frac{1}{k}) = \{\omega : |X_n(\omega)| \leq \frac{1}{k}$ eventually, any given $k \geq 1\} = \{\omega : \lim_{n\to\infty} X_n(\omega) = 0\}$. If $\omega \in \bigcap_{k=1}^{\infty} B(\frac{1}{k})$, then given $k \geq 1$, then $N_{\omega,k} \geq 1$ where $X_n(\omega) \leq \frac{1}{k}$, for all $n \geq N_{\omega,k}$.

$(\Leftarrow)$ Suppose $\mathrm{P}(\lim_{n\to\infty} X_n = 0) = 1$. For a contradiction, assume that $\sum_{n=1}^{\infty} P(|X_n| > \epsilon) = \infty$ for some $\epsilon > 0$. Now define $A_n = \{\omega : |X_n| > \epsilon\}$, $n \geq 1$. Then by the Borel Cantelli Lemma,

$$P(\limsup_{n\to\infty} A_n) = P(|X_n| > \epsilon \quad \text{i.o}) = 1$$

This means $P(\omega : |X_n(\omega)| \nrightarrow 0) = 1 \Rightarrow P(\omega : |X_n(\omega)| \to 0) = 0$. a contradiction. Thus if $P(\lim_{n\to\infty} |X_n| = 0) = 1$ then for all $\epsilon > 0$, $\sum_{n=1}^{\infty} P(|X_n| > \epsilon) < \infty$. ■

CXXXIII. **Tail σ-algebra.** The **tail $\sigma-$ algebra** of a sequence of random variables $X_n, n \geq 1$ on a psp *psp* is

$$\mathcal{T} \equiv \cap_{n=1}^{\infty} \sigma\langle\{X_j : j \geq n\}\rangle, \subset \mathcal{F}$$

i.e. it is the $\sigma-$algebra generated by $X_j, j \geq n$ (all the r.v.'s beyond some point n)

CXXXIV. **Definition of Tail Event.** Any set (event) $A \in \mathcal{T}$ is called a **tail event**.

CXXXV. **Tail Random Variable.** An *extended real − valued* random variable $T : \Omega \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is called a **tail random variable** if T is $\langle T, \mathcal{B}(\mathbb{R})\rangle-$ measurable (e.g., $T^{-1}((-\infty, \alpha)) = \{\omega \in \Omega : T(\omega) < \alpha\} \in \mathcal{T}$, any $\alpha \in \mathbb{R}$).

CXXXVI. **Example.** $\overline{\lim} X_n$ exists and is finite. pick $m \geq 1$, then

$$\{\omega : \overline{\lim} X_n(\omega) \geq \alpha\} = \{\omega : \inf_{n \geq m} \sup_{j \geq n} X_j(\omega) \geq \alpha\} \in \sigma\langle\{X_j(\omega) : j \geq m\}\rangle \quad m \geq 1$$

Then,

$$\{\omega : \overline{\lim} X_n(\omega) \geq \alpha\} \in \mathcal{T} \equiv \bigcap_{m \geq 1} \sigma\langle\{X_j(\omega) : j \geq m\}\rangle$$

CXXXVII. **Kolmogorov 0-1 law.** Tail events of a sequence $X_n, n \geq 1$ of independent random variables have probabilities of 0 or 1.

$$\text{i.e. } A \in \mathcal{T} \quad \Rightarrow \quad P(A) = 0 \quad \textit{or} \quad P(A) = 1.$$

CXXXVIII. **Degeneracy of tail events.** Suppose $(\Omega, \mathcal{F}, P)$ is a psp and $\mathcal{T}$ is a tail $\sigma-$ algebra defined by a sequence of indpendent random variables $X_n, n \geq 1$. If $T : \Omega \to \overline{\mathbb{R}}$ is a tail random variable (i.e. T is $\langle \mathcal{T}, \mathcal{B}(\mathbb{R})\rangle-$ measurable), then T is degenerate. That is, There exists a $c \in \overline{\mathbb{R}}$ where $P(T = c) = 1$.

Proof : for any $x \in \overline{\mathbb{R}}, P(T \leq x) = 0$ or 1 and $P(T \leq x)$ is right continuous. Let $c = \inf\{x \in \overline{\mathbb{R}} : P(T \leq x) = 1\}$. Therefore $P(T = c) = 1$.

CXXXIX. **Example.** Let $X_n, n \geq 1$ be independent random variables with $E(X_n) = 0$ and $E(X_n^2) = 1, \forall n \geq 1$, sich that $S_n = \sum_{n=1}^{\infty} X_i$,

$$\lim_{n \to \infty} P\Big(\frac{S_n}{\sqrt{n}} \leq x\Big) \to \Phi(x), \quad \forall x \in \mathbb{R}$$

where $\Phi(\cdot)$ is the cdf of a $N(0,1)$. Show that $\limsup_{n \to \infty} \frac{S_n}{\sqrt{n}} = +\infty$ a.s. (P).

Proof: Let: $S \equiv S(\omega) = \limsup_{n \to \infty} \frac{S_n}{\sqrt{n}} = +\infty$. We'll show S is a tail random variable. If so, then there exists a $c \in \overline{\mathbb{R}}$ where $P(T = c) = 1$. If possible suppose $c < \infty$. Let $x = \max\{0, c\} + 1 \in \mathbb{R}$. Then,

$$\begin{aligned}
0 < 1 - \Phi(x) &= \lim_{n \to \infty} P\Big(\frac{S_n}{\sqrt{n}} \leq x\Big) \\
&\leq \lim_{n \to \infty} P\Big(\cup_{m=n}^{\infty} \Big\{\frac{S_m}{\sqrt{m}} > x\Big\}\Big) \\
\text{by m.c.f.a} \quad &= \lim_{n \to \infty} P\Big(\cap_{n=1}^{\infty} \cup_{m=n}^{\infty} \Big\{\frac{S_m}{\sqrt{m}} > x\Big\}\Big) \\
&\leq P\Big(\limsup_{n \to \infty} \frac{S_n}{\sqrt{n}} \leq x\Big) \\
&= P\big(S \geq x\big) = 0
\end{aligned}$$

The last line is because $x > c$ and $P(T = c) = 1$. Thus $P(S \geq x) = 0$ establishing the contradiction. To show S is a tail R.V. Let $S_{m,n} = X_m + \ldots, X_n, S_{m-1} = \sum_{i=1}^{m-1} X_i$, so $S_n = S_{m-1} + S_{m,n}$, for $i \leq m \leq n$. Fix m. Then

$$\limsup_{n \to \infty} \frac{S_n}{\sqrt{n}} = \lim_{n \to \infty} \Big(\frac{S_{m-1}}{\sqrt{n}} + \frac{S_{m,n}}{\sqrt{n}}\Big) = \lim_{n \to \infty} \Big(\frac{S_{m,n}}{\sqrt{n}}\Big) \in \sigma\langle\{X_m, X_{m+1}, \ldots\}\rangle$$

That is, the $\overline{\lim}X_n$ is in the tail $\sigma-$algebra. $\Rightarrow$

$$S = \limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \in \mathcal{T} = \bigcap_{m=1}^{\infty} \sigma\langle\{X_m, X_{m+1}, \ldots\}\rangle \quad \blacksquare$$

9 LLN and Methods of Convergence

CXL. **Convergence in Probability.** A sequence of random variables $X_n, n \geq 1$, on a psp $(\Omega, \mathcal{F}, P)$ is said to **converge in probability** to a random variable X_0 on $(\Omega, \mathcal{F}, P)$ if for any $\epsilon > 0$,

$$P(|X_n - X_0| > \epsilon) \to 0 \quad \text{as} \quad n \to \infty$$

Notation $X_n \xrightarrow{p} X_0$.

CXLI. **Almost sure convergence implies convergence in probability.** On a psp $(\Omega, \mathcal{F}, P)$, if $X_n \to X_0$ a.s.(P), the $X_n \xrightarrow{p} X_0$.

Proof: for any $\epsilon > 0$, let $A_m = \{\omega : |X_m(\omega) - X_0(\omega)| > \epsilon\}$ and $B_n = \bigcup_{m=n}^{\infty} A_m$. Then,

$$\begin{aligned}
\lim_{n\to\infty} P(|X_n - X_0| > \epsilon) &\leq \lim_{n\to\infty} P(\bigcup_{m=n}^{\infty} |X_m - X_0| > \epsilon) \\
&= P(\bigcap_{n=1}^{\infty} B_n) \\
&= P(\overline{\lim} A_n(\epsilon)) \\
&= P(|X_m - X_0| > \epsilon \text{ i.o.}) \\
&= 0
\end{aligned}$$

Thus, Convergence almost surely implies convergence in probability.

CXLII. **Convergence in Probability Equivalencies Theorem.** Let $X_n, n \geq 1$, be random variables on a psp $(\Omega, \mathcal{F}, P)$. Then the following are equivalent:

(i) $X_n \xrightarrow{p} X_0$ as $n \to \infty$.

(ii) $\sup_{m\geq n} P(|X_m - X_n| > \epsilon) \to 0$ as $n \to \infty$, for all $\epsilon > 0$.

(iii) Every subsequence of $\{X_n\}_{n\geq 1}$ has a further subsequence converging almost surely to the same r.v. X_0.

CXLIII. **Example.** If $X_n \xrightarrow{p} X$ and $X_n \xrightarrow{p} Y$, then $P(X = Y) = 1$.

Proof: There exist a subsequence X_{n_k} such that $X_{n_k} \to X$ a.s.(P) (since $X_n \xrightarrow{p} X$). Then there exists a further subsequence $X_{n_{k_i}}$ of X_{n_k} such that $X_{n_{k_i}} \to Y$ a.s. (P) (since $X_n \xrightarrow{p} Y$). But $X_{n_{k_i}}$ (as a subsequence of X_{n_k}) $\to X$ a.s.(P). Therefore X=Y a.s.(P) $\blacksquare$

CXLIV. **Version of the Continuous Mapping theorem.** Let $\{X_n\}_{n\geq 0}$ be random variables on a psp $(\Omega, \mathcal{F}, P)$. Then If $X_n \xrightarrow{p} X_0$ and $g : \mathbb{R} \to \mathbb{R}$ is continuous, then $g(X_n) \xrightarrow{p} g(X_0)$ as $n \to \infty$.

Proof (i): It suffices to show that given any subsequence $\{g(X_{n_k})\}$ of $\{g(X_n)\}$, there exists a further subsequence $\{g(X_{n_{k_i}})\}$ of $\{g(X_{n_k})\}$ where $\{g(X_{n_{k_i}})\} \to g(X_0)$ a.s.(P). note: Since $X_n \xrightarrow{p} X_0$, given $\{X_{n_k}\}$, we can find a further subsequence $\{X_{n_{k_i}}\}$ where $X_{n_{k_i}} \to X_0$ a.s.(P). That is, there exists a set $A \in \mathcal{F}, P(A) = 1$ and if $\omega \in A, X_{n_{k_i}}(\omega) \to X_0(\omega)$ as $n \to \infty$. Thus, using a real analysis result, since $g(\cdot)$ is continuous, $g(X_{n_{k_i}}(\omega)) \to g(X_0(\omega))$ as $n \to \infty$ for a given $\omega \in A$. Therefore, $g(X_{n_{k_i}}) \to g(X_0)$ a.s.(P).

CXLV. **Convergence in Probability Equivalence.** Let $\{X_n\}_{n\geq 0}$ be random variables on a psp $(\Omega, \mathcal{F}, P)$. Then If $X_n \xrightarrow{p} X_0$ if and only if $\sup_{m\geq n} |X_m - X_n| \xrightarrow{p} 0$ as $n \to \infty$.

CXLVI. **Median.** For a random variable X, a real number $m(X)$ is called a **median** if

$$P(X \geq m(X)) \geq 1/2 \qquad \text{and} \qquad P(X \leq m(X)) \geq 1/2$$

(i) Let $m_1 \equiv \inf\{x \in \mathbb{R} : P(X \leq x) \geq 1/2\}$. Then m_1 is a median of X.

(ii) If $P(|X| \geq c) < \epsilon \leq 1/2$, for some $c > 0$ and some $\epsilon > 0$, then $|m(X)| \leq c$.

(iii) $cm(X)$ and $c + m(X)$ are medians of cX and $c + X$ for any $c \in \mathbb{R}$.

Proof(i): By definition, if $\{x_n\} \downarrow m_1$ then $P(X \leq m_1) = \lim_{n\to\infty} P(x \leq x_n) \geq 1/2$. If $\{x_n\} \uparrow m_1$ then $P(X < x_n) \leq P(X \leq x_n) < 1/2$ since $x_n < m_1$. This means $P(X \geq m_1) = \lim_{n\to\infty} P(X \geq x_n) = \lim_{n\to\infty}[1 - p(X < x_n)] \geq 1/2$. Therefore, m_1 is the median of X.

CXLVII. **Levy's Inequality.** Suppose $X_1, \ldots, X_n$ are independent random variables on some psp $(\Omega, \mathcal{F}, P)$. Let $S_j = S_1 + \cdots + X_j, 1 \leq j \leq n$. Then for any $\epsilon > 0$,

$$(i) \quad P\Big(\max_{1\leq j\leq n} \big[S_j - m(S_j - S_n)\big] \geq \epsilon\Big) \leq 2P(S_n \geq \epsilon)$$

$$(ii) \quad P\Big(\max_{1\leq j\leq n} \big|S_j - m(S_j - S_n)\big| \geq \epsilon\Big) \leq 2P(|S_n| \geq \epsilon)$$

CXLVIII. **Levy's Theorem.** Let $\{X_n\}_{n\geq 1}$, be independent random variables on a psp $(\Omega, \mathcal{F}, P)$ and $S_n = \sum_{j=1}^{n} X_j, n \geq 1$. Then S_n converges a.s.(P) if and only if S_n converges in probability.

Proof: By definition almost sure convergence implies convergence in probability. Thus it is left to show if S_n converges in probability, then S_n converges a.s.(P). Suppose S_n converges in probability, then for any $1 > \epsilon > 0$, there exists $n_\epsilon \geq 1$ such that whenever $n \geq n_\epsilon$,

$$\sup_{m\geq n} P\Big(|S_m - S_n| \geq \frac{\epsilon}{2}\Big) \;\leq\; \frac{\epsilon}{2} \;<\; \frac{1}{2} \tag{1}$$

Let $S_{m,n} = S_m - S - n, m \geq n$. By (1) and (ii) (from median section) we know $|m(S_{m,n})| \leq \frac{\epsilon}{2}$ and this holds for all $m \geq n \geq n_\epsilon$. Hence, for all $k \geq n \geq n_\epsilon$,

$$P\Big(\max_{n<m\leq k} |S_m - S_n| \geq \epsilon\Big) = P\Big(\max_{n<m\leq k}$$

to be continued

CXLIX. **Convergence of sums Theorem (CLT-esque).** Let $\{X_n\}_{n\geq 1}$ be independent and identically distributed random variables on a psp $(\Omega, \mathcal{F}, P)$ with $EX_n = 0$ and $EX_n^2 < \infty$, for $n \geq 1$. If $\sum_{n=1}^{\infty} EX_n^2 < \infty$, then $S_n = \sum_{j=1}^{n} X_j, n \geq 1$ converges a.s.(P) (to $S = \sum_{n=1}^{\infty} X_n$).

Proof: By Levy's theorem, it suffices to show S_n converges in probability. For any $\epsilon > 0$,

$$\sup_{m\geq n} (|S_m - S_n| > \epsilon) \leq \sup_{m\geq n} \frac{E(S_m - S_n)^2}{\epsilon^2} \leq \frac{\sum_{m=n+1}^{\infty} EX_m^2}{\epsilon^2} \to 0$$

as $n \to \infty$. Therefore $S_n \xrightarrow{p} S$.

CL. **Berry-Esseen Theorem.** Let $X_1, \ldots, X_n$ be independent random variables with $EX_i = 0$ and $E|X_i|^3 < \infty, 1 \leq i \leq n$. Then for all $n \geq 4$,

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

Where $S_n = \sum_{j=1}^{n}, \sigma_n^2 = \text{var}(S_n) = \sum_{i=1}^{n} EX_i^2$, and $\Phi(\cdot)$ is the standard normal cdf.

CLI. **Kolmogorov 3 Series Theorem.** Let $\{X_n\}_{n\geq 1}$ be independent random variables on a psp $(\Omega, \mathcal{F}, P)$. Given $c > 0$, consider the three series given by

$$\sum_{n=1}^{\infty} P(|X_n| > c), \qquad \sum_{n=1}^{\infty} EX_n^{(c)}, \qquad \sum_{n=1}^{\infty} var(X_n^{(c)}),$$

Where $X_n^{(c)} = X_n \mathbb{I}(|X_n| \leq c)$. Then,

(*i*) If the three series converge for **some** $c > 0$, then $S_n = \sum_{j=1}^{n} X_j, n \geq 1$ converges a.s.(P).

(*ii*) If $S_n = \sum_{j=1}^{n} X_j$ converges a.s.(P), then the three series converge **for all** $c > 0$.

CLII. **3- series corollary.** Let $\{X_n\}_{n\geq 1}$ be independent random variables on a psp $(\Omega, \mathcal{F}, P)$ with $EX_n = 0$, $n \geq 1$,

(*i*) If

$$\sum_{n=1}^{\infty} \left\{E[X_n^{(c)}]^2 + E|X_n|\mathbb{I}(|X_n| > c)\right\} < \infty$$

for some $c > 0$, then $S_n = \sum_{j=1}^{n} X_j$ a.s.(P).

(*ii*) If $\sum_{n=1}^{\infty} E|X_n|^{\alpha_n} < \infty$ for some $\{\alpha_n\} \subset [1, 2]$, then $S_n = \sum_{j=1}^{n} X_j$ converges a.s.(P).

CLIII. **Example** Consider Independent random variables $X_n, n \geq 1$ with $X_n \sim$ Unif$(-a_n, a_n)$ where $a_n \to 0$. Then $S_n = \sum_{j=1}^{n} S_j$ converges a.s.(P) if and only if $\sum_{n=1}^{\infty} a_n^2 < \infty$.

Proof: Assume $S_n = \sum_{j=1}^{n} X_j$ converges a.s.(P), then by the Kolmogorov 3-series theorem, the 3-series converge for all $c > 0$. We know that $E(X_n) = 0$ and $E(X_n^2) = a_n^2/4$. From the three series theorem we have $\sum_{n=1}^{\infty} \text{var}(X_n^{(c)}) < \infty$. Choose $c = \sup_{n\to\infty}\{a_n\}$. Then $\sum_{n=1}^{\infty} \text{var}(X_n^{(c)}) = \frac{1}{4} \sum_{n=1}^{\infty} a_n^2 < \infty$. Thus $\sum_{n=1}^{\infty} a_n^2 < \infty$ by the 3-series theorem.

Assume $\sum_{n=1}^{\infty} a_n^2 < \infty$. We know that $a_n \to 0$. Let $c = 1$. There exists an $N \in \mathbb{N}$ such that whenever $n \geq N, |a_n| < 1$. Thus $\sum_{j=1}^{\infty} P(X_j > 1) \leq n < \infty$. Since $E(X_n^2) = a_n^2/4 < \infty$, then by Jensen's Inequality, $E(X_n) < \infty$. Since $X_n\mathbb{I}(X_n \leq c) \subset X_n$ we have both $\sum_{n=1}^{\infty} E(X_n^{(c)}) < \infty$ and $\sum_{n=1}^{\infty} \text{var}(X_n^{(c)}) < \infty$ for all $c > 0$. Thus all 3 series are finite for $c = 1$ and thus $S_n = \sum_{j=1}^{n} X_j$ converges a.s.(P).

CLIV. **Law of Large Numbers.** A sequence of random variables $X_n, n \geq 1$, on a psp $(\Omega, \mathcal{F}, P)$ is said to satisfy a **strong law of large numbers** (SLLN) or a **weak law of large numbers** (WLLN) if there exists a sequence $b_n \in \mathbb{R}$ (centering value) and $0 < a_n \uparrow$ (scaling value) such that, for $S_n = \sum_{j=1}^{n} X_n, n \geq 1$,

$$\text{SLLN: } \frac{S_n - b_n}{a_n} \to 0 \text{ a.s.(P)} \qquad \text{WLLN: } \frac{S_n - b_n}{a_n} \xrightarrow{p} 0$$

note: if the SLLN holds, this implies the WLLN holds.

CLV. **Marcinkiewicz-Zygmund SLLN.** Let $\{X_n\}_{n\geq 1}$ be independent, identically distributed random variables on a psp $(\Omega, \mathcal{F}, P)$. Let $S_n = \sum_{j=1}^{n} X_n, n \geq 1$ and real number $0 < p < 2$.

(*i*) If for some $c \in \mathbb{R}$,

$$\frac{S_n - nc}{n^{1/p}} \to 0 \quad \text{a.s.(P)} \tag{1}$$

then $E|X_1|^p < \infty$.

(*ii*) If $E|X_1|^p < \infty$, then (1) holds with $c = EX_1$ if $1 \leq p < 2$ and (1) holds for any $c \in \mathbb{R}$ if $0 < p < 1$.

note: if $1 \leq p < 2$ and $E|X|^p < \infty$, then $E|X| < \infty$. If $0 < p < 1$ then $E|X|$ may not exist. If $p = 2$, and $E|X|^2 < \infty, c = E|X|$, then by the CLT

$$\frac{S_n - nE|X|}{n^{1/2}} \xrightarrow{d} N(0, \text{var}(X)) \text{ as } n \to \infty$$

CLVI. **Kolmogorov's SLLN.** Let $\{X_n\}_{n\geq 1}$ be independent identically distributed random variables on a psp $(\Omega, \mathcal{F}, P)$. Then,

$$\frac{S_n}{n} \to EX_1 \text{ a.s.(P) if and only if } E|X_1| < \infty$$

This is a special case for $p = 1$. This can be equivalently written $\frac{S_n - nEX_1}{n} \to 0$ a.s.(P).

CLVII. **Kronecker's Lemma.** Suppose $\{a_n\}$ and $\{b_n\}$ are sequences of real numbers such that $0 < b_n \uparrow \infty$ and $\sum_{n=1}^{\infty} \frac{a_n}{b_n}$ converges, then

$$\frac{1}{b_n} \sum_{j=1}^{n} a_j \to 0 \text{ as } n \to \infty$$

CLVIII. **Theorem.** Let $\{X_n\}_{n\geq 1}$ be independent random variables on a psp $(\Omega, \mathcal{F}, P)$ satisfying $\sum_{n=1}^{\infty} E|X_n|^{\alpha_n}/n^{\alpha_n} < \infty$ for some $1 \leq \alpha_n \leq 2,\ n \geq 1$. Then

$$\frac{1}{n} \sum_{i=1}^{n} (X_i - EX_i) \to 0 \quad \text{a.s.(P)}.$$

Proof: First some useful facts:

(a) for $t > 0$ and $a, b \in \mathbb{R}$, $|a+b|^t \le [2\max\{|a|,|b|\}]^t \le 2^t[|a|^t + |b|^t]$

(b) if $t \ge 1$, $\phi(x) = |x|^t$ is convex for $x \in \mathbb{R}$. For a r.v. Y, by Jensen's inequality we have $\phi(EY) \le E(\phi(Y))$ provided the moments are finite.

Thus we have

$$\sum_{n=1}^{\infty} E|\frac{X_n - EX_n}{n}|^{\alpha_n} \le \sum_{n=1}^{\infty} \frac{2^{\alpha_n}[E|X_n|^{\alpha_n} + |EX_n|^{\alpha_n}]}{n^{\alpha_n}} \qquad \text{by (a)}$$
$$\le \sum_{n=1}^{\infty} \frac{2^2[2E|X_n|^{\alpha_n}]}{n^{\alpha_n}} < \infty \qquad \text{by (b)}$$

Hence, $\sum_{i=1}^{n} \frac{X_i - EX_i}{i}$ converges to something finite a.s.(P) as $n \to \infty$. That is $\sum_{i=1}^{\infty} \frac{X_i - EX_i}{i}$ finitely exists a.s.(P). Now by Kroneckers Lemma ($a_i = X_i - EX_i, b_i = i$), $\frac{1}{n}\sum_{i=1}^{n}(X_i - EX_i) \to 0$ a.s.(P).

CLIX. **Etemaldi's SLLN.** Let $\{X_n\}_{n\ge1}$ be **pairwise** independent and identically distributed random variables on a psp $(\Omega, \mathcal{F}, P)$. Then,

$$\frac{S_n}{n} \to EX_1 \text{ a.s.(P) if and only if } E|X_1| < \infty$$

CLX. **Example application.** Let f be a bounded measurable function on [0,1] that is continuous at $\frac{1}{2}$. Evaluate

$$\lim_{n\to\infty} \int_0^1 \int_0^1 \cdots \int_0^1 f\Big(\frac{x_1 + x_2 + \cdots + x_n}{n}\Big) dx_1 dx_2 \cdots dx_n$$

[Hint: Define $X_1, \ldots, X_n$ as independent uniforms on a common probability space $(\Omega, \mathcal{F}, P)$].

Proof: Define $X_1, \ldots, X_n$ as independent uniforms on a common probability space $(\Omega, \mathcal{F}, P)$, then we have,

$$\lim_{n\to\infty} \int_0^1 \int_0^1 \cdots \int_0^1 f\Big(\frac{X_1 + X_2 + \cdots + X_n}{n}\Big) dP(X_1)dP(X_2)\cdots dP(X_n) =$$
$$\lim_{n\to\infty} \int f\Big(\frac{X_1 + X_2 + \cdots + X_n}{n}\Big) dP$$

Since $P(\Omega) = 1$ and f is bounded, by the Bounded convergence theorem we have

$$\lim_{n\to\infty} \int f\Big(\frac{X_1 + X_2 + \cdots + X_n}{n}\Big) dP = \int \lim_{n\to\infty} f\Big(\frac{X_1 + X_2 + \cdots + X_n}{n}\Big) dP$$

And by the Strong Law of Large Numbers we have

$$\lim_{n\to\infty} \Big(\frac{X_1 + X_2 + \cdots + X_n}{n}\Big) \to EX_1 = \frac{1}{2} \quad \text{a.s.(P)}$$

Hence,

$$\lim_{n\to\infty} \int f\Big(\frac{X_1 + X_2 + \cdots + X_n}{n}\Big) dP = \int_0^1 f\Big(\frac{1}{2}\Big) dP = f\Big(\frac{1}{2}\Big)$$

CLXI. **Example involving SLLN.** Let $\{X_n\}_{n\geq 1}$ be a sequence of iid random variables with $E|X_1|^p = \infty$ for some $p \in (0,2)$. Then $P(\limsup_{n\to\infty} |n^{-1/p}\sum_{i=1}^n X_i| = \infty) = 1$.

Proof: Suppose $E|X_1|^p = \infty$ for some $p \in (0,2)$ such that $\frac{S_n - nc}{n^{1/p}}$ does not converge for any $c \in \mathbb{R}$. We know $\limsup_{n\to\infty} |n^{-1/p}\sum_{i=1}^n X_i|$ is a tail event and by Kolomogorov's 0-1 Law, it is degenerate. Consequently, $\text{P}(\limsup_{n\to\infty} |n^{-1/p}\sum_{i=1}^n X_i| = c) = 0$ or 1 for $c > 0$ including ∞. Assume $\text{P}(\limsup_{n\to\infty} |n^{-1/p}\sum_{i=1}^n X_i| = c) = 1$ for $c < \infty$. Then we see,

$$\begin{aligned}
\frac{|X_n|}{n^{1/p}} &= \frac{S_n - S_{n-1}}{n^{1/p}} \\
&\leq \frac{S_n}{n^{1/p}} + \frac{(n-1)^{1/p}S_{n-1}}{(n-1)^{1/p}n^{1/p}} \\
\limsup_{n\to\infty}\frac{|X_n|}{n^{1/p}} &\leq \limsup_{n\to\infty}\frac{S_n}{n^{1/p}} + \limsup_{n\to\infty}\frac{(n-1)^{1/p}S_{n-1}}{(n-1)^{1/p}n^{1/p}}
\end{aligned}$$

Hence,

$$P\Big(\limsup_{n\to\infty}\frac{|X_n|}{n^{1/p}} \leq 2c < 2c+1\Big) = 1$$

And By the Borel Cantelli Lemma

$$P\Big(\limsup_{n\to\infty}\frac{|X_n|}{n^{1/p}} > 2c+1\Big) = 0$$

Implies

$$\sum_{n=1}^{\infty} P\Big(\limsup_{n\to\infty}\frac{|X_n|}{n^{1/p}} > 2c+1\Big) < \infty$$

If and only if $E|X_1|^p < \infty$ which is a contradiction. Thus since $E|X_1|^p = \infty$, c must equal ∞ and

$$P\Big(\limsup_{n\to\infty}\Big|n^{-1/p}\sum_{i=1}^n X_i\Big| = \infty\Big) = 1 \quad \blacksquare$$

CLXII. **Empirical CDF.**

Definition: The **empirical cumulative distribution function** of random variables $X_i, \ldots X_n$ is the random cdf:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(X_i \leq x), \quad x \in \mathbb{R}$$

Notes:

(1) With X_i's defined on a psp $(\Omega, \mathcal{F}, P)$, then for each $\omega \in \Omega$, "the value of $F_n(x)$" at

$$\omega \equiv F_n(x,\omega) = \frac{1}{n}\sum_{i=1}^n \mathbb{I}(X_i(\omega) \leq x)$$

(2) $F_n(x)$ is a right continuous, non decreasing function of $x \in \mathbb{R}$.

(3) For any $x \in \mathbb{R}, F_n(x)$ is a random variable, i.e. is $\langle \mathcal{F}, \mathcal{B}(\mathbb{R})\rangle$-measurable.

CLXIII. **Glivenko Cantelli Theorem.** Let $\{X_n\}_{n\geq 1}$ be iid random variables on a psp $(\Omega, \mathcal{F}, P)$ with common cdf $F(\cdot)$. Let $F_n(\cdot)$ be the empirical cdf based on $X_i, \ldots X_n$ and let

$$D_n = \sup_{x\in\mathbb{R}} \left|F_n(x) - F(x)\right|$$

Then, (i) D_n is a random variable, for any $n \geq 1$, and (ii) $D_n \to 0$ a.s. (P) as $n \to \infty$.

CLXIV. **Example.** Let $\{X_i\}_{i\geq 1}$ be iid random variables with cdf $F(\cdot)$. Let $F_n(x) \equiv \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(X_i \leq x)$ be the empirical cdf. Suppose $x_n \to x_0$ and $F(\cdot)$ is continuous at x_0. Show that $F_n(x_n) \to F(x_0)$ w.p. 1.

Proof: By the Glivenko Cantelli theorem, $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \equiv D_n \to 0$ a.s.(P). Then

$$\begin{aligned} |F_n(x_n) - F(x_0)| &\leq |F_n(x_n) - F(x_n)| + |F(x_n) - F(x_0)| \\ &\leq D_n + |F(x_n) - F(x_0)| \to 0 \quad \text{a.s.(P)} \end{aligned}$$

Since $x_n \to x_0 \Rightarrow |F(x_n) - F(x_0)| \to 0$. ■

CLXV. **Convergence in Distribution.** Let $\mu_n, n \geq 0$ be probability measures on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)), 1 \leq k < \infty$.

(1) For $\mathbf{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k$,

$$F_n(x) = \mu_n\Big((-\infty, x_i] \times \cdots \times (\infty, x_k]\Big)$$

is called the **cumulative distribution function** (cdf) of μ_n.

If a random vector X_n on some probability space $(\Omega_n, \mathcal{F}_n, P_n)$ has probability distribution μ_n [meaning $P_n(X_n \in A) = \mu_n(A), A \in \mathcal{B}(\mathbb{R}^k)$], then F_n is also called the cdf of X_n.

(2) We say that a sequence of probability measures μ_n (or cdfs F_n) **converges weakly** to μ_0 (to F_0), denoted as $\mu_n \stackrel{d}{\Rightarrow} \mu_0$ if

$$\lim_{n\to\infty} F_n(x) \to F_0(x) \quad \forall x \in C(F_0)$$

Where $C(F_0) = \{\mathbf{x} = (x_1, \ldots, x_k) \in \mathbb{R}^k : F_0 \text{ is continuous at } \mathbf{x}\}$

(3) **Convergence in Distribution** is the same thing as **Convergence in Law** is the same thing as **Converges weakly**.

Note! If $X_n \stackrel{d}{\to} X_0$, then it is **not** necessarily true that "$X_n \stackrel{p}{\to} X_0$" or that $X_n \to X_0$ a.s.(P).

In other words, $X_n \stackrel{p}{\to} X_0$ and $X_n \to X_0$ a.s.(P) only make sense on a common psp $(\Omega, \mathcal{F}, P)$. But there is one exception, $X_n \stackrel{p}{\to} c$, where c is a constant. However, if we have a common psp $(\Omega, \mathcal{F}, P)$, $X_n \stackrel{d}{\to} X_0$ doesn't imply $X_n \stackrel{p}{\to} X_0$. For instance, let $X_0 \sim N(0,1)$ and let $X_n = -X_0$ for all $n \geq 1$. Then $X_n \stackrel{d}{\to} X_0$, but $X_n \stackrel{p}{\not\to} X_0$ because

$$\lim_{n\to\infty} P(|X_n - X_0| > \epsilon) = \lim_{n\to\infty} P(|2X_0| > \epsilon) \not\to 0$$

CLXVI. **Convergence in probability implies convergence in Distribution.** Let $\{X_n\}_{n\geq 0}$ be random variables on $(\Omega, \mathcal{F}, P)$. If $X_n \xrightarrow{p} X_0$, then $X_n \xrightarrow{d} X_0$ as $n \to \infty$.
This also holds for random vectors.

CLXVII. **Special Case: Distribution implies probability.** Suppose $\{X_n\}_{n\geq 0}$, are random variables with each X_n defined on some probability space $(\Omega_n, \mathcal{F}_n, P_n)$, and $X_n \xrightarrow{d} X_0$ where $P_0(X_0 = c) = 1$ for some $c \in \mathbb{R}$. Then, $X_n \xrightarrow{p} c$, meaning $P_n(|X_n - c| > \epsilon) \to 0$ as $n \to \infty$ for any $\epsilon > 0$.
Note: On a common $(\Omega, \mathcal{F}, P)$, if $X_n \xrightarrow{d} X_0$ with $P(X_0 = c) = 1$, then $X_n \xrightarrow{p} c$.

CLXVIII. **Skorohod's Embedding Thoerem.** Suppose $\mu_n, n \geq 0$ are probability measures on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)), 1 \leq k < \infty$, such that $\mu_n \overset{d}{\Rightarrow} \mu_0$. Then there exists random vectors $Y_n, n \geq 0$ on a common psp $(\Omega, \mathcal{F}, P)$ such that Y_n has probability distribution $\mu_n, n \geq 0$ and $Y_n \to Y_0$ a.s.(P). That is, $P(Y_n \in A) = P(Y_n^{-1}(A)) = \mu_n(A), A \in \mathcal{B}(\mathbb{R}^k), Y_n^{-1}(A) \in \mathcal{F}$.

CLXIX. **Continuous Mapping Theorems.**

a.) Let $\mu_n, n \geq 0$, be probability measures on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)), 1 \leq k < \infty$, and let $h : \mathbb{R}^k \to \mathbb{R}^m (1 \leq m < \infty)$ be a $\langle \mathbb{R}^k, \mathcal{B}(\mathbb{R}^k) \rangle$-measurable function such that $\mu_0(D_h) = 0$, where $D_h \in \mathcal{B}(\mathbb{R}^k)$ denotes the set of all points of discontinuity of the function h. If $\mu_n \overset{d}{\Rightarrow} \mu_0$ then

$$\mu_n h^{-1} \overset{d}{\Rightarrow} \mu_0 h^{-1}$$

b.) Let $X_n, n \geq 0$ be $\mathbb{R}^k$-valued random vectors (each on some $(\Omega_n, \mathcal{F}_n, P_n)$) and measurable $h : \mathbb{R}^k \to \mathbb{R}^m (1 \leq m < \infty)$ be such that $P_0(X_0 \in D_h) = 0$, where D_h is as above. If $X_n \xrightarrow{d} X_0$, then

$$h(X_n) \xrightarrow{d} h(X_0)$$

Proof(a): By Skorohod's Theorem, $\mu_n \overset{d}{\Rightarrow} \mu_0$. This implies there exists some psp $(\Omega, \mathcal{F}, P)$ and random variables $Y_n, n \geq 0$, such that $Y_n \to Y_0$ a.s.(P) and $\mu_{Y_n} \equiv \mu_n$. There exists $A \in \mathcal{F}, P(A) = 1$, and $Y_n(\omega) \to Y_0(\omega)$ for all $\omega \in A$. Let $B = A^c \cup \{\omega : Y_0(\omega) \in D_h\}$. Then $P(B) \leq P(A) + P(\{\omega : Y_0(\omega) \in D_h\}) = \mu_0(D_h) = 0$. So $P(B^c) = 1$ and for $\omega \in B^c, Y_n(\omega) \to Y_0(\omega)$ and h is continuous at $Y_0(\omega)$.

$$\begin{aligned} &\Rightarrow \quad h(Y_n(\omega)) \to h(Y_0(\omega)), \quad \omega \in B^c \\ &\Rightarrow \quad h(Y_n(\omega)) \to h(Y_0(\omega)) \quad a.s.(P) \\ &\Rightarrow \quad h(Y_n(\omega)) \xrightarrow{d} h(Y_0(\omega)) \\ &\Rightarrow \quad \mu_{Y_n} h^{-1} \overset{d}{\Rightarrow} \mu_{Y_0} h^{-1} \end{aligned}$$

CLXX. **Corollary.** Let X_n and $Y_n, n \geq 0$ be random variables such that (X_n, Y_n) are defined on a probability space $(\Omega_n, \mathcal{F}_n, P_n)$ for each $n \geq 0$ and $(X_n, Y_n) \xrightarrow{d} (X_0, Y_0)$, then

i.) $X_n + Y_n \xrightarrow{d} X_0 + Y_0$
Because $h(x, y) = x + y$ is continuous in $\mathbb{R}^2$.

ii.) $X_nY_n \xrightarrow{d} X_0Y_0$
Because $h(x,y) = xy$ is continuous in $\mathbb{R}^2$.

iii.) $X_n/Y_n \xrightarrow{d} X_0/Y_0$ if $P_0(Y_0 = 0) = 0$
Because $h(x,y) = x/y$ has discontinuity points $D_h = \{(x,y) \in \mathbb{R}^2 : y = 0\}$ but $P((X_0, Y_0) \in D_h) = P(Y_0 = 0) = 0$. so (iii.) holds by CMT.

CLXXI. **Slutsky's Theorem.** Let X_n and $Y_n, n \geq 1$ be random variables such that (X_n, Y_n) are defined on a probability space $(\Omega_n, \mathcal{F}_n, P_n)$ for each $n \geq 1$. If $X_n \xrightarrow{p} X$ and $Y_n \xrightarrow{p} a$ for some $a \in \mathbb{R}$ then

i.) $X_n + Y_n \xrightarrow{p} X + a$

ii.) $X_nY_n \xrightarrow{p} aX$

iii.) $X_n/Y_n \xrightarrow{p} X/a$ provide $a \neq 0$

Proof: Make a probability space $(\Omega_0, \mathcal{F}_0, P_0)$ and a random variable X_0 which has the same distribution as X. And on $(\Omega_0, \mathcal{F}_0, P_0)$, let $Y_n(\omega) = a$ for all $\omega \in \Omega_0$. Then (X_0, Y_0) has cdf

$$P(X_0 \leq x, Y_0 \leq y) = \begin{cases} P(X_0 \leq x) & \text{if } y \geq a, x \in \mathbb{R} \\ 0 & \text{if } y < a, x \in \mathbb{R} \end{cases}$$

Which is continuous at (x,y) where $y \neq a$ and $P_0(X_0 < x) = P_)(X_0 \leq x)$. Now check $(X_n, Y_n) \xrightarrow{d} (X_0, Y_0)$ and apply previous corollary.

CLXXII. **Example.** Let $X_n \xrightarrow{d} X$ and $x_n \to x$ as $n \to \infty$. If $P(X = x) = 0$, then show that $P(X_n \leq x_n) \to P(X \leq x)$.

Proof: Since $X_n \xrightarrow{d} X_0$ and $x_n \to x$, we can define a new random variable $Z_n = (x - x_n) \xrightarrow{p} 0$. Let $Y_n = X_n - (x_n - x)$. Then $Y_n \xrightarrow{d} X$ by Slutsky's theorem. Since $P(X = x) = 0$, x is a continuity point of the distribution of X. Therefore, for a fixed x

$$P(Y_n \leq x) \to P(X \leq x)$$

Which means

$$P(X_n + x - x_n \leq x) \to P(X \leq x)$$

This means that

$$P(X_n \leq x_n) \to P(X \leq x)$$

as was to be shown. ■

CLXXIII. **μ-continuity set.** Let μ be a probability measure on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$. Then, a set $A \in \mathcal{B}(\mathbb{R}^k)$ is called a **μ-continuity set** if $\mu(\partial A) = 0$, where $\partial A =$ the boundary of A (i.e. $\overline{A} \cap [A^o]^c$ where $\overline{A}$ is the closure of A and A^o is the interior of A.)

CLXXIV. **Helly Bray Theorem**: Let $\mu_n, n \geq 0$ be probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then,

i.) $\mu_n \overset{d}{\Rightarrow} \mu_0$ if and only if $\mu_n(A) \to \mu_0(A)$, for all $A \in \mathcal{B}(\mathbb{R})$ with $\mu_0(\partial(A) = 0$.

ii.) $\mu_n \overset{d}{\Rightarrow} \mu_0$ if and only if

$$\int f\, d\mu_n \to \int f\, d\mu_0$$

for all bounded continuous functions $f : \mathbb{R} \to \mathbb{R}$.

Note: $\mu_n \overset{d}{\Rightarrow} \mu_0$ does not guarantee that $\mu_n(A) \to \mu_0(A)$ for any $A \in \mathcal{B}(\mathbb{R}^k)$.

CLXXV. **Helly Bray-esque Lemma.** If $\mu_n \overset{d}{\Rightarrow} \mu_0$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $f : \mathbb{R} \to \mathbb{R}$ is a bounded continuous measurable function with $\mu_0(D_f) = 0$ where $D_f \in \mathcal{B}(\mathbb{R})$ is the set of discontinuity points of f, then

$$\int f\, d\mu_n \to \int f\, d\mu_0$$

Proof: By Skorohod's Theorem, there exists a probability space $(\Omega, \mathcal{F}, P)$ and random variables Y_n where $Y_n \to Y_0$ a.s.(P) amd $P(Y_n \in A) = P(Y_n^{-1}(A)) = \mu_n(A)$ for all $n \geq 0$ and $A \in \mathcal{B}(\mathbb{R})$. Since $P(Y_0 \in D_f) = \mu_0(D_f) = 0$ we have $f(Y_n) \to f(Y_0)$ a.s.(P) by CMT implying that $|f(Y_n(\omega)| \leq M$ for some $M \in \mathbb{R}$, for all $\omega \in \Omega, n \geq 0$. By the bounded convergence theorem,

$$\int_\Omega f(Y_n) dP \to \int_\Omega f(Y_0) dP \qquad \text{as } n \to \infty$$

Then by a change of variables,

$$\int_\Omega f(Y_n) dP \to \int_{\mathbb{R}} f\, d(pY_n^{-1}) = \int_{\mathbb{R}} f\, d\mu_n \quad \text{for all } n \geq 0 \quad \blacksquare$$

CLXXVI. **Tight or Stochastically Bounded.** A sequence of probability measures $\mu_n, n \geq 1$, on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called **tight** or **stochastically bound** if for any $\epsilon > 0$ there exists $M \equiv M_\epsilon > 0$ such that

$$\sup_{n \geq 1} \mu_n\big([-M, M]^c\big) < \epsilon$$

As sequence of random variables $X_n, n \geq 1$ (each defined on a psp $(\Omega_n, \mathcal{F}_n, P_n)$) is called **tight** or **stochastically bound** if the sequence $\{\mu_n\}_{n\geq 1}$ of probability distributions of $\{X_n\}_{n \geq 1}$ is tight. That is, for any $\epsilon > 0$ there exists $M \equiv M_\epsilon > 0$ such that

$$\sup_{n \geq 1} P_n\big(|X_n| > M\big) < \epsilon \quad = \quad \sup_{n\geq 1} \mu_n\big([-M, M]^c\big) < \epsilon$$

where μ_n denotes the probability measure of X_n on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. i.e. $\mu_n(A) = P_n(X_n \in A), A \in \mathcal{B}(\mathbb{R})$.

Example: $X_n \sim N\big(\frac{1}{n}, 1\big), n \geq 1$ implies $\{X_n\}_{n \geq 1}$ is tight. For any $\epsilon > 0$, find $M \in \mathbb{R}$ such that $\sup_{n \geq 1} P_n\big(|X_n| > M\big) < \epsilon$.

CLXXVII. **Uniform Integrability $\Rightarrow$ Tightness.** A sequence of random variables $\{X_n\}_{n\geq 1}$ (each defined on a psp $(\Omega_n, \mathcal{F}_n, P_n)$) is called **uniformly integrable** (UI) if for any $\epsilon > 0$, there exists $t \equiv t_\epsilon > 0$ such that

$$\sup_{n\geq 1} E_n\big|X_n\big|\,\mathbb{I}\big(|X_n| > t\big) = \sup_{n\geq 1}\int_{|X_n|>t} |X_n| dP_n = \sup_{n\geq 1}\int_{|X|>t} |X| d\mu_n < \epsilon$$

where μ_n denotes the probability measure of X_n on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Proof:

$$\sup_{n\geq 1} P_n\big(|X_n| > M\big) \leq \sup_{n\geq 1}\int_{|X_n|>M} \frac{|X_n|}{M} dP_n = \frac{1}{M}\sup_{n\geq 1}\int_{|X_n|>M} |X_n| dP_n < \epsilon$$

CLXXVIII. **Tightness Propositions.** Suppose $X_n, n \geq 0$ are random variables with each X_n defined on some probability space $(\Omega_n, \mathcal{F}_n, P_n)$.

i.) If $X_n \xrightarrow{d} X_0$ then $\{X_n\}_{n\geq 1}$ is tight.

ii.) If $\{X_n\}_{n\geq 1}$ is tight and $Y_n \xrightarrow{p} 0$ (X_n, Y_n defined on $(\Omega_n, \mathcal{F}_n, P_n)$), then $X_n Y_n \xrightarrow{p} 0$.

Proof(i): Let $F_n(x) = P_n(X_n \leq x)$, for all $x \in \mathbb{R}$. Let $\epsilon > 0$. Let $M_0 > 0$ be such that $P_0\big(|X_0| \geq M_0\big) < \epsilon/4$ and $\pm M_0 \in C(F_0)$, then $F_n(\pm M_0) \to F_0(\pm M_0)$. This means there exists $N \in \mathbb{N}$ such that whenever $n \geq \mathbb{N}$, $|F_n(\pm M_0) - F_0(\pm M_0)| < \epsilon/4$. So for $n \geq N$,

$$\begin{aligned} P_n\big(|X_n| > M_0\big) &\leq F_n(-M_0) + \big(1 - F_n(M_0)\big) \\ &\leq F_0(-M_0) + \frac{\epsilon}{4} + \big(1 - F_0(M_0) + \frac{\epsilon}{4}\big) \\ &\leq P_0\big(|X_0| \geq M_0\big) + \frac{\epsilon}{2} \\ &< \epsilon \end{aligned}$$

Then for $i = 1, \ldots, n$ pick M_i where $P_i(|X_i| > M_i) < \epsilon$. Now let $M = \max\{M_0, \ldots, M_n\}$. This means $\sup_{n\geq 1} P_n(|X_n| > M) < \epsilon$.

Proof(ii): Pick $M > 0$ such that $\sup_{n\geq 1} P_n(|X_n| > M) < \delta$, for some $\delta > 0$. Fix $\epsilon > 0$, then

$$\begin{aligned} P_n\big(|X_nY_n| > \epsilon\big) &= P_n\big(|X_nY_n| > \epsilon, X_n > M\big) + P_n\big(|X_nY_n| > \epsilon, X_n \leq M\big) \\ &\leq P_n\Big(|Y_n| > \frac{\epsilon}{M}\Big) + P_n\big(|X_n| > t\big) \\ &\leq P_n\Big(|Y_n| > \frac{\epsilon}{M}\Big) + \delta \end{aligned}$$

Since $P_n\big(|Y_n| > \frac{\epsilon}{M}\big) \to 0$ as $n \to \infty$, $\overline{\lim}_{n\to\infty} P_n\big(|X_nY_n| > \epsilon\big) \leq \delta$. Since $\delta > 0$ was arbitrary, $\overline{\lim}_{n\to\infty} P_n\big(|X_nY_n| > \epsilon\big) = 0$. or $\lim_{n\to\infty} P_n\big(|X_nY_n| > \epsilon\big)$. Therefore $X_nY_n \xrightarrow{p} 0$.

CLXXIX. **Sub-subsequence convergence in distribution implies tightness.** A sequence of random variables $\{X_n\}_{n\geq 1}$ is tight if and only if, for any subsequence X_{n_k} of X_n, there exists a further subsequence $X_{n_{k_i}}$ of X_n and a random variable X_0 such that $X_{n_{k_i}} \xrightarrow{d} X_0$. This also holds true for sequences of probability measures $\{\mu_n\}_{n\geq 1}$.

Note: X_0 depends on the particular subsequence of X_{n_k}. For example, suppose $X_n \sim \text{Unif}(0, 2 + a_n)$, $n \geq 1$ where $a_n = (-1)^n$. (note $\sup_{n\geq 1} P_n\big(|X_n| > 3\big) = 0 \Rightarrow \{X_n\}$ is tight). But,

$$\begin{aligned} Y_n &\equiv X_{2n} \xrightarrow{d} Y_0 \sim \text{Unif}(0,3) \\ Z_n &\equiv X_{2n-1} \xrightarrow{d} Z_0 \sim \text{Unif}(0,1) \end{aligned}$$

This shows tightness does not imply convergence in distribution.

CLXXX. **Corollary for tightness and convergence. in distribution.** If a sequence $\{X_n\}_{n\geq 1}$ of random variables is tight and all its convergent subsequences converge in distribution to the same random variable X_0, then $X_n \xrightarrow{d} X_0$.

Proof: Let $F_n(x) = P_n(X_n \leq x)$ for all $x \in \mathbb{R}$. If possible suppose that $X_n \stackrel{d}{\not\to} X_0$. Then, there exists $x_0 \in C(F_0)$ such that $F_n(x_0) \not\to F_0(x_0)$. So there exists a subsequence of $\{n_k\}$ and some $\epsilon > 0$ where $|F_{n_k}(x_0) - F_0(x_0)| > \epsilon$ for all $k \geq 1$. Since $\{X_n\}$ is tight, there exists a subsequence $X_{n_{k_i}}$ of X_{n_k} which converges in distribution and by assumption, the only possibility is to X_0. So $|F_{n_{k_o}}(x_0) - F_0(x_0)| \to 0$ contradicting our assumption. ■

CLXXXI. **Theorem.** Suppose $\{X_n\}_{n\geq 0}$ are random variables with each X_n defined on some probability space $(\Omega_n, \mathcal{F}_n, P_n)$. If the sequence $\{X_n\}_{n\geq 1}$ is U.I. and $X_n \xrightarrow{d} X_0$, then $E_0|X_0| < \infty$ and $E_n|X_n| \to E_0|X_0|$.

Proof: By Skorohod's theorem there exists a common probability space $(\Omega, \mathcal{F}, P)$ and random variables $Y_n, n \geq 0$ such that Y_n have the same distribution as $X_n, n \geq 0$ and $Y_n \to Y_0$ a.s.(P). Since $\{X_n\}$ are U.I. $\{Y_n\}$ are U.I. Therefore, by the Uniform integrability convergence theorem, $\int_\Omega |Y_0| dP < \infty$ and

$$\int_\Omega |Y_n| dP \to \int_\Omega |Y_0| dP$$

and $E_p Y_n = E_n Y_n$ and $E_p Y_0 = E_0 Y_0$. Therefore, $E_n X_n \to E_0 X_0$. note: since Y_n has the same distribution as X_n, they have the same moments.

CLXXXII. **Convergence of Moments.** Suppose $\{X_n\}_{n\geq 0}$ are random variables with each X_n defined on some probability space $(\Omega_n, \mathcal{F}_n, P_n)$. If $\sup_{n\geq 1} E_n|X_n|^{r+\delta} < \infty$ for some $r \geq 1$ and $\delta > 0$ and $X_n \xrightarrow{d} X_0$, then $E_0|X_0|^r < \infty$.

Proof:

$$\begin{aligned} \sup_{n\geq 1} E_n|X_n|^{r+\delta} < \infty \quad &\Rightarrow \quad \{|X_n|^r\} \text{are UI} \\ X_n \xrightarrow{d} X_0 \quad &\Rightarrow \quad X_n^r \xrightarrow{d} X_0^r \quad \text{by CMT} \end{aligned}$$

By the previous theorem, $E_0|X_0|^r < \infty$ and $E_n|X_n|^r \to E_0|X_0|^r$

CLXXXIII. **Method of Moments Theorem.** *(Frechet-Shohat Theorem)* Suppose $\{X_n\}, n \geq 0$ are random variables with each X_n defined on some probability space $(\Omega_n, \mathcal{F}_n, P_n)$. If $\lim_{n\to\infty} E_n X_n^r = \beta_r \in \mathbb{R}$ for all $r \geq 1$ and if $\{\beta_r : r \geq 1\}$ are the moments of a unique random variable X_0, then $X_n \xrightarrow{d} X_0$.

Proof: Since $E_n X_n^2 \to \beta_2 < \infty$, then $\{X_n\}$ are tight. note:

$$P\big(|X_n| > M\big) \le E_n\Big(\frac{|X_n|^2}{M^2}\Big) \le \frac{1}{M^2} E_n X_n^2$$

So given a subsequence X_{n_k} of X_n, there exists a further subsequence $X_{n_{k_i}}$ and a random variable $\tilde{X_0}$ such that $X_{n_{k_i}} \xrightarrow{d} \tilde{X_0}$. Now for any integer $r \ge 1$, $E_n X_n^{2r} \to \beta_{2r} \in \mathbb{R}$ which means $\sup_{n \ge N} E_n |X_n|^{2r} < \infty$. From an earlier corollary we get $E_0 |\tilde{X_0}| < \infty$ and $E_{n_{k_i}} |X_{n_{k_i}}^r| \to E_{\tilde{X_0}} |\tilde{X_0}|^r$. But $E_n X_n^r \to \beta_r$ by assumption which implies $E_{n_{k_i}} |X_{n_{k_i}}^r| \to \beta_r$. Since $\{\beta_r\}_{r \ge 1}$ uniquely determines the distribution of X_0, X_0 and $\tilde{X_0}$ have the same distribution. So the distribution of $\tilde{X_0}$ does not depend on the subsequence X_{n_k} of X_n and by Corollary $X_n \xrightarrow{d} X_0$.

Summarizing ideas (Nordman's Final Exam Review)

Types of "convergence" of Random Variables

(1) $X_n \to X_0$ a.s.(P). This means that there is a common $(\Omega, \mathcal{F}, P)$ where $X_n(\omega) : \Omega \to \mathbb{R}$, $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-measurable and there exists $A \in \mathcal{F}$ where $P(A) = 1$ such that $X_n(\omega) \to X_0(\omega)$, for all $\omega \in A$.

(2) $X_n \xrightarrow{p} X_0$. There's a common $(\Omega, \mathcal{F}, P)$ and $P\big(|X_n - X_0| > \epsilon\big) = P\big(\{\omega : |X_n(\omega) - X_0(\omega)| > \epsilon\}\big) \to 0$ as $n \to \infty$ for all $\epsilon > 0$.

But, $X_n \xrightarrow{p} a$, $a \in \mathbb{R}$, can be meaningful and well defined even if X_n is defined on $(\Omega_n, \mathcal{F}_n, P_n), n \ge 1$. $P_n(|X_n - a| > \epsilon) \to 0$ as $n \to \infty$ for all $\epsilon > 0$. note: $\{|X_n - a| > \epsilon\} \in \mathcal{F}_n$

Also, $X_n \to X_0$ a.s.(P) implies $X_n \xrightarrow{p} X_0$

$X_n \to X_0$ a.s.(P) if and only if $\sup_{m \ge n} |X_m - X_n| \xrightarrow{p} 0$ if and only if $\sup_{m \ge n} |X_m - X_n| \to 0$ a.s.(P).

$X_n \xrightarrow{p} X_0$ if and only if $\sup_{m \ge n} P\big(|X_m - X_n| > \epsilon\big) \to 0$ as $n \to \infty$ if and only if for any subsequence X_{n_k} of X_n, there's a further subsequence $X_{n_{k_i}}$, where $X_{n_{k_i}} \to X_0$ a.s.(P).

(3) $X_n \xrightarrow{d} X_0$:

Special case: if $X_n, n \ge 1$ are defined on the same $(\Omega, \mathcal{F}, P)$, then $X_n \to X_0$ a.s.(P) $\Rightarrow X_n \xrightarrow{p} X_0 \Rightarrow X_n \xrightarrow{d} X_0$.

Each random variable could be defined on a different psp $(\Omega_n, \mathcal{F}_n, P_n)$. Then each random variable X_n has a probability measure μ_n on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and given $A \in \mathcal{B}(\mathbb{R})$,
$\mu_n(A) = P_n(X_n \in A) = P(X_n^{-1}(A)), n \ge 0.$ and has cdf $F_n(x) = \mu_n\big((-\infty, x]\big) = P_n(X_n \le x)$.

We say $X_n \xrightarrow{d} X_0$ (or $\mu_n \overset{d}{\Rightarrow} \mu_0$) if $\lim_{x \uparrow x_0} F(x) = P(X < x_0)$ and $F_n(x) \to F_0(x)$ as $n \to \infty$ for all $x \in \mathbb{R}$ where $x \in C(F_0)$. That is $\mu_0(\{x\}) = 0$ or $P_0(X_0 = x) = 0$.

Characterizations of $X_n \xrightarrow{d} X_0$

$X_n \xrightarrow{d} X_0$ if and only if $\mu_n(A) \to \mu_0(A)$, for all $A \in \mathcal{B}(\mathbb{R})$ where $\mu_0(\partial A) = 0$ if and only if $\big($note: $A = (-\infty, x], \partial A = \{x\}\big)$.

$$\int_{\mathbb{R}} f(x) d\mu_n(x) \to \int_{\mathbb{R}} f(x) d\mu_0(x)$$

for all bounded, continuous $f : \mathbb{R} \to \mathbb{R}$.

Also, if $\{X_n\}$ is tight, then given $\epsilon > 0$, there exist $M > 0$ where

$$\sup_{n \geq 1} P_n\big(|X_n| > M\big) = \sup_{n \geq 1} \mu_n\big([-M, M]^c\big) < \epsilon$$

and if all convergent subsequences of X_n converge in distribution to the same X_0, then $X_n \xrightarrow{d} X_0$.

Method of Moments

$E_n X_n^r \to E_0 X_0^r$, for all integers $r \geq 1$ and X_0 is uniquely determined by its moments. (e.g. if the m.g.f. of X_0, $M_{X_0}(t) = E_0 e^{tX_0} < \infty$ for all $|t| < \delta$ then $X_n \xrightarrow{d} X_0$.

Showing a function is Measurable

$(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2), f : \Omega_1 \to \Omega_2$. by definition, f is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable if, for any $A \in \mathcal{F}_2$, $f^{-1}(A) = \{\omega \in \Omega_1 : f(\omega) \in A\} \in \mathcal{F}_1$. It suffices to show that $f^{-1}(A) \in \mathcal{F}_1$ for all $A \in \mathcal{C}$ where $\mathcal{C}$ is some collection of subsets in $\mathcal{F}_2$ where $\sigma\langle \mathcal{C} \rangle = \mathcal{F}_2$. That is $\mathcal{C}$ is the smallest collection of subsets in $\mathcal{F}_2$ such that $\sigma\langle \mathcal{C} \rangle = \mathcal{F}_2$.

To show $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is measurable, we need $X^{-1}(A) \in \mathcal{F}$ for all $A \in \mathcal{B}(\mathbb{R})$. But it suffices to show that $X^{-1}\big((-\infty, \alpha]\big) \in \mathcal{F}$ for all $\alpha \in \mathbb{R}$ since $\sigma\langle \{(-\infty, \alpha] : \alpha \in \mathbb{R}\}\big) = \mathcal{B}(\mathbb{R})$.

To show $f : (\Omega, \mathcal{F}) \to (\Omega_1 \times \Omega_2,\ \mathcal{F}_1 \times \mathcal{F}_2)$ is measurable, it suffices to check $f^{-1}(A_1 \times A_2) \in \mathcal{F}$ for all measurable rectangles $A_1 \times A_2 \in \mathcal{F}_1 \times \mathcal{F}_2$ since $\sigma\langle \{A_1 \times A_2\} \rangle = \mathcal{F}_1 \times \mathcal{F}_2$.