# Searching for Universal Truths

# Measure Theory

**Sunghee Yun**

sunghee.yun@gmail.com

# Navigating Mathematical and Statistical Territories

# Notations

- sets of numbers
  - $\mathbf{N}$ - set of natural numbers
  - $\mathbf{Z}$ - set of integers
  - $\mathbf{Z}_+$ - set of nonnegative integers
  - $\mathbf{Q}$ - set of rational numbers
  - $\mathbf{R}$ - set of real numbers
  - $\mathbf{R}_+$ - set of nonnegative real numbers
  - $\mathbf{R}_{++}$ - set of positive real numbers
  - $\mathbf{C}$ - set of complex numbers
- sequences $\langle x_i \rangle$ and the like
  - finite $\langle x_i \rangle_{i=1}^n$, infinite $\langle x_i \rangle_{i=1}^\infty$ - use $\langle x_i \rangle$ whenever unambiguously understood
  - similarly for other operations, *e.g.*, $\sum x_i$, $\prod x_i$, $\cup A_i$, $\cap A_i$, $\bigtimes A_i$
  - similarly for integrals, *e.g.*, $\int f$ for $\int_{-\infty}^{\infty} f$
- sets
  - $\tilde{A}$ - complement of $A$

- $A \sim B$ - $A \cap \tilde{B}$
- $A \Delta B$ - $(A \cap \tilde{B}) \cup (\tilde{A} \cap B)$
- $\mathcal{P}(A)$ - set of all subsets of $A$

- sets in metric vector spaces
  - $\overline{A}$ - closure of set $A$
  - $A^\circ$ - interior of set $A$
  - $\mathbf{relint}\, A$ - relative interior of set $A$
  - $\mathbf{bd}\, A$ - boundary of set $A$
- set algebra
  - $\sigma(\mathcal{A})$ - $\sigma$-algebra generated by $\mathcal{A}$, $i.e.$, smallest $\sigma$-algebra containing $\mathcal{A}$
- norms in $\mathbf{R}^n$
  - $\|x\|_p$ $(p \geq 1)$ - $p$-norm of $x \in \mathbf{R}^n$, $i.e.$, $(|x_1|^p + \cdots + |x_n|^p)^{1/p}$
  - $e.g.$, $\|x\|_2$ - Euclidean norm
- matrices and vectors
  - $a_i$ - $i$-th entry of vector $a$
  - $A_{ij}$ - entry of matrix $A$ at position $(i, j)$, $i.e.$, entry in $i$-th row and $j$-th column
  - $\mathbf{Tr}(A)$ - trace of $A \in \mathbf{R}^{n \times n}$, $i.e.$, $A_{1,1} + \cdots + A_{n,n}$

- symmetric, positive definite, and positive semi-definite matrices
  - $\mathbf{S}^n \subset \mathbf{R}^{n\times n}$ - set of symmetric matrices
  - $\mathbf{S}^n_+ \subset \mathbf{S}^n$ - set of positive semi-definite matrices; $A \succeq 0 \Leftrightarrow A \in \mathbf{S}^n_+$
  - $\mathbf{S}^n_{++} \subset \mathbf{S}^n$ - set of positive definite matrices; $A \succ 0 \Leftrightarrow A \in \mathbf{S}^n_{++}$
- sometimes, use Python script-like notations (with serious abuse of mathematical notations)
  - use $f : \mathbf{R} \to \mathbf{R}$ as if it were $f : \mathbf{R}^n \to \mathbf{R}^n$, *e.g.*,

$$\exp(x) = (\exp(x_1), \ldots, \exp(x_n)) \quad \text{for } x \in \mathbf{R}^n$$

and

$$\log(x) = (\log(x_1), \ldots, \log(x_n)) \quad \text{for } x \in \mathbf{R}^n_{++}$$

which corresponds to Python code `numpy.exp(x)` or `numpy.log(x)` where `x` is instance of `numpy.ndarray`, *i.e.*, `numpy` array
  - use $\sum x$ to mean $\mathbf{1}^T x$ for $x \in \mathbf{R}^n$, *i.e.*

$$\sum x = x_1 + \cdots + x_n$$

which corresponds to Python code `x.sum()` where `x` is `numpy` array

– use $x/y$ for $x, y \in \mathbf{R}^n$ to mean

$$\begin{bmatrix} x_1/y_1 & \cdots & x_n/y_n \end{bmatrix}^T$$

which corresponds to Python code `x / y` where `x` and `y` are 1-d `numpy` arrays

– use $X/Y$ for $X, Y \in \mathbf{R}^{m \times n}$ to mean

$$\begin{bmatrix} X_{1,1}/Y_{1,1} & X_{1,2}/Y_{1,2} & \cdots & X_{1,n}/Y_{1,n} \\ X_{2,1}/Y_{2,1} & X_{2,2}/Y_{2,2} & \cdots & X_{2,n}/Y_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m,1}/Y_{m,1} & X_{m,2}/Y_{m,2} & \cdots & X_{m,n}/Y_{m,n} \end{bmatrix}$$

which corresponds to Python code `X / Y` where `X` and `Y` are 2-d `numpy` arrays

## Some definitions

**Definition 1. [infinitely often - i.o.]** *statement* $P_n$*, said to happen* infinitely often *or* i.o. *if*

$$(\forall N \in \mathbf{N})\,(\exists n > N)\,(P_n)$$

**Definition 2. [almost everywhere - a.e.]** *statement* $P(x)$*, said to happen* almost everywhere *or* a.e. *or* almost surely *or* a.s. *(depending on context) associated with measure space* $(X, \mathscr{B}, \mu)$ *if*

$$\mu\{x|P(x)\} = 1$$

*or equivalently*

$$\mu\{x| \sim P(x)\} = 0$$

# Some conventions

- (for some subjects) use following conventions
  - $0 \cdot \infty = \infty \cdot 0 = 0$
  - $(\forall x \in \mathbf{R}_{++})(x \cdot \infty = \infty \cdot x = \infty)$
  - $\infty \cdot \infty = \infty$

# Real Analysis

# Set Theory

# Some principles

**Principle 1. [principle of mathematical induction]**

$$P(1)\&[P(n \Rightarrow P(n+1)] \Rightarrow (\forall n \in \mathbf{N})P(n)$$

**Principle 2. [well ordering principle]** *each nonempty subset of* $\mathbf{N}$ *has a smallest element*

**Principle 3. [principle of recursive definition]** *for* $f : X \to X$ *and* $a \in X$*, exists unique infinite sequence* $\langle x_n \rangle_{n=1}^{\infty} \subset X$ *such that*

$$x_1 = a$$

*and*

$$(\forall n \in \mathbf{N})\,(x_{n+1} = f(x_n))$$

- note that Principle 1 $\Leftrightarrow$ Principle 2 $\Rightarrow$ Principle 3

# Some definitions for functions

**Definition 3. [functions]** *for* $f : X \to Y$

- *terms,* map *and* function*, exterchangeably used*
- $X$ *and* $Y$*, called* domain of $f$ *and* codomain of $f$ *respectively*
- $\{f(x)|x \in X\}$*, called* range of $f$
- *for* $Z \subset Y$*,* $f^{-1}(Z) = \{x \in X|f(x) \in Z\} \subset X$*, called* preimage *or* inverse image of $Z$ under $f$
- *for* $y \in Y$*,* $f^{-1}(\{y\})$*, called* fiber of $f$ over $y$
- $f$*, called* injective *or* injection *or* one-to-one *if* $(\forall x \neq v \in X) (f(x) \neq f(v))$
- $f$*, called* surjective *or* surjection *or* onto *if* $(\forall x \in X) (\exists yinY) (y = f(x))$
- $f$*, called* bijective *or* bijection *if* $f$ *is both injective and surjective, in which case,* $X$ *and* $Y$*, said to be* one-to-one correspondece *or* bijective correspondece
- $g : Y \to X$*, called* left inverse *if* $g \circ f$ *is identity function*
- $h : Y \to X$*, called* right inverse *if* $f \circ h$ *is identity function*

## Some properties of functions

**Lemma 1. [functions]** *for* $f : X \to Y$

- $f$ *is injective* if and only if $f$ *has left inverse*

- $f$ *is surjective* if and only if $f$ *has right inverse*

- *hence,* $f$ *is bijective* if and only if $f$ *has both left and right inverse because if* $g$ *and* $h$ *are left and right inverses respectively,* $g = g \circ (f \circ h) = (g \circ f) \circ h = h$

- *if* $|X| = |Y| < \infty$, $f$ *is injective* if and only if $f$ *is surjective* if and only if $f$ *is bijective*

# Countability of sets

- set $A$ is countable if range of some function whose domain is **N**

- **N**, **Z**, **Q**: countable

- **R**: *not* countable

# Limit sets

- for sequence, $\langle A_n \rangle$, of subsets of $X$
  - *limit superior or limsup of $\langle A_n \rangle$*, defined by

$$\limsup \langle A_n \rangle = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

  - *limit inferior or liminf of $\langle A_n \rangle$*, defined by

$$\liminf \langle A_n \rangle = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

- always

$$\liminf \langle A_n \rangle \subset \limsup \langle A_n \rangle$$

- when $\liminf \langle A_n \rangle = \limsup \langle A_n \rangle$, sequence, $\langle A_n \rangle$, said to *converge to it*, denote

$$\lim \langle A_n \rangle = \liminf \langle A_n \rangle = \limsup \langle A_n \rangle = A$$

# Algebras of sets

- collection $\mathscr{A}$ of subsets of $X$ called *algebra* or *Boolean algebra* if

$$(\forall A, B \in \mathscr{A})(A \cup B \in \mathscr{A}) \text{ and } (\forall A \in \mathscr{A})(\tilde{A} \in \mathscr{A})$$

  - $(\forall A_1, \ldots, A_n \in \mathscr{A})(\cup_{i=1}^n A_i \in \mathscr{A})$
  - $(\forall A_1, \ldots, A_n \in \mathscr{A})(\cap_{i=1}^n A_i \in \mathscr{A})$
- algebra $\mathscr{A}$ called *$\sigma$-algebra* or *Borel field* if
  - every union of a countable collection of sets in $\mathscr{A}$ is in $\mathscr{A}$, *i.e.*,

$$(\forall \langle A_i \rangle)(\cup_{i=1}^\infty A_i \in \mathscr{A})$$

- given sequence of sets in algebra $\mathscr{A}$, $\langle A_i \rangle$, exists disjoint sequence, $\langle B_i \rangle$ such that

$$B_i \subset A_i \text{ and } \bigcup_{i=1}^\infty B_i = \bigcup_{i=1}^\infty A_i$$

# Algebras generated by subsets

- *algebra generated by* collection of subsets of $X$, $\mathcal{C}$, can be found by

$$\mathscr{A} = \bigcap \{\mathscr{B} | \mathscr{B} \in \mathcal{F}\}$$

  where $\mathcal{F}$ is family of all algebras containing $\mathcal{C}$
  – smallest algebra $\mathscr{A}$ containing $\mathcal{C}$, $i.e.$,

$$(\forall \mathscr{B} \in \mathcal{F})(\mathscr{A} \subset \mathscr{B})$$

- *$\sigma$-algebra generated by* collection of subsets of $X$, $\mathcal{C}$, can be found by

$$\mathscr{A} = \bigcap \{\mathscr{B} | \mathscr{B} \in \mathcal{G}\}$$

  where $\mathcal{G}$ is family of all $\sigma$-algebras containing $\mathcal{C}$
  – smallest $\sigma$-algebra $\mathscr{A}$ containing $\mathcal{C}$, $i.e.$,

$$(\forall \mathscr{B} \in \mathcal{G})(\mathscr{A} \subset \mathscr{B})$$

# Relation

- $x$ said to *stand in relation* $\mathbf{R}$ to $y$, denoted by $x \mathbf{\ R\ } y$
- $\mathbf{R}$ said to *be relation on* $X$ if $x \mathbf{\ R\ } y \Rightarrow x \in X$ and $y \in X$
- $\mathbf{R}$ is
  - transitive if $x \mathbf{\ R\ } y$ and $y \mathbf{\ R\ } z \Rightarrow x \mathbf{\ R\ } z$
  - symmetric if $x \mathbf{\ R\ } y = y \mathbf{\ R\ } x$
  - reflexive if $x \mathbf{\ R\ } x$
  - antisymmetric if $x \mathbf{\ R\ } y$ and $y \mathbf{\ R\ } x \Rightarrow x = y$
- $\mathbf{R}$ is
  - *equivalence relation* if transitive, symmetric, and reflexive, $e.g.$, modulo
  - *partial ordering* if transitive and antisymmetric, $e.g.$, "$\subset$"
  - *linear (or simple) ordering* if transitive, antisymmetric, and $x \mathbf{\ R\ } y$ or $y \mathbf{\ R\ } x$ for all $x, y \in X$
    - $e.g.$, "$\geq$" linearly orders **R** while "$\subset$" does not $\mathcal{P}(X)$

# Ordering

- given partial order, $\prec$, $a$ is
  - a first/smallest/least element if $x \neq a \Rightarrow a \prec x$
  - a last/largest/greatest element if $x \neq a \Rightarrow x \prec a$
  - a minimal element if $x \neq a \Rightarrow x \not\prec a$
  - a maximal element if $x \neq a \Rightarrow a \not\prec x$
- partial ordering $\prec$ is
  - strict partial ordering if $x \not\prec x$
  - reflexive partial ordering if $x \prec x$
- strict linear ordering $<$ is
  - *well ordering* for $X$ if every nonempty set contains a first element

## Axiom of choice and equivalent principles

**Axiom 1. [axiom of choice]** *given a collection of nonempty sets, $\mathcal{C}$, there exists $f : \mathcal{C} \to \cup_{A\in\mathcal{C}} A$ such that*

$$(\forall A \in \mathcal{C}\ )\,(f(A) \in A)$$

- also called *multiplicative axiom* - preferred to be called to axiom of choice by Bertrand Russell for reason writte on page 20
- no problem when $\mathcal{C}$ is finite
- need axiom of choice when $\mathcal{C}$ is not finite

**Principle 4. [Hausdorff maximal principle]** *for particial ordering $\prec$ on $X$, exists a maximal linearly ordered subset $S \subset X$, $i.e.$, $S$ is linearity ordered by $\prec$ and if $S \subset T \subset X$ and $T$ is linearly ordered by $\prec$, $S = T$*

**Principle 5. [well-ordering principle]** *every set $X$ can be well ordered, $i.e.$, there is a relation $<$ that well orders $X$*

- note that Axiom 1 $\Leftrightarrow$ Principle 4 $\Leftrightarrow$ Principle 5

# Infinite direct product

**Definition 4. [direct product]** *for collection of sets, $\langle X_\lambda \rangle$, with index set, $\Lambda$,*

$$\bigtimes_{\lambda \in \Lambda} X_\lambda$$

*called* direct product

- *for $z = \langle x_\lambda \rangle \in \bigtimes X_\lambda$, $x_\lambda$ called* $\lambda$-th coordinate of $z$

- if one of $X_\lambda$ is empty, $\bigtimes X_\lambda$ is empty
- *axiom of choice* is equivalent to converse, *i.e.*, if none of $X_\lambda$ is empty, $\bigtimes X_\lambda$ is not empty
  if one of $X_\lambda$ is empty, $\bigtimes X_\lambda$ is empty
- this is why Bertrand Russell prefers *multiplicative axiom* to *axiom of choice* for name of axiom (Axiom 1)

# Real Number System

# Field axioms

- field axioms - for every $x, y, z \in \mathbf{F}$
  - $(x + y) + z = x + (y + z)$ - additive associativity
  - $(\exists 0 \in \mathbf{F})(\forall x \in \mathbf{F})(x + 0 = x)$ - additive identity
  - $(\forall x \in \mathbf{F})(\exists w \in \mathbf{F})(x + w = 0)$ - additive inverse
  - $x + y = y + x$ - additive commutativity
  - $(xy)z = x(yz)$ - multiplicative associativity
  - $(\exists 1 \neq 0 \in \mathbf{F})(\forall x \in \mathbf{F})(x \cdot 1 = x)$ - multiplicative identity
  - $(\forall x \neq 0 \in \mathbf{F})(\exists w \in \mathbf{F})(xw = 1)$ - multiplicative inverse
  - $x(y + z) = xy + xz$ - distributivity
  - $xy = yx$ - multiplicative commutativity
- system (set with $+$ and $\cdot$) satisfying axiom of field called *field*
  - $e.g.$, field of module $p$ where $p$ is prime, $\mathbf{F}_p$

# Axioms of order

- axioms of order - subset, $\mathbf{F}_{++} \subset \mathbf{F}$, of positive (real) numbers satisfies
  - $x, y \in \mathbf{F}_{++} \Rightarrow x + y \in \mathbf{F}_{++}$
  - $x, y \in \mathbf{F}_{++} \Rightarrow xy \in \mathbf{F}_{++}$
  - $x \in \mathbf{F}_{++} \Rightarrow -x \not\in \mathbf{F}_{++}$
  - $x \in \mathbf{F} \Rightarrow x = 0 \vee x \in \mathbf{F}_{++} \vee -x \in \mathbf{F}_{++}$
- system satisfying field axioms & axioms of order called *ordered field*
  - $e.g.$, set of real numbers ($\mathbf{R}$), set of rational numbers ($\mathbf{Q}$)

# Axiom of completeness

- completeness axiom
  - every nonempty set $S$ of real numbers which has an upper bound has a least upper bound, *i.e.*,
$$\{l|(\forall x \in S)(l \leq x)\}$$
has least element.
  - use $\inf S$ and $\sup S$ for least and greatest element (when exist)

- ordered field that is complete is *complete ordered field*
  - *e.g.*, **R** (with $+$ and $\cdot$)

$\Rightarrow$ axiom of Archimedes
  - given any $x \in \mathbf{R}$, there is an integer $n$ such that $x < n$

$\Rightarrow$ corollary
  - given any $x < y \in \mathbf{R}$, exists $r \in \mathbf{Q}$ such tat $x < r < y$

# Sequences of R

- sequence of **R** denoted by $\langle x_i \rangle_{i=1}^{\infty}$ or $\langle x_i \rangle$
  - mapping from **N** to **R**
- limit of $\langle x_n \rangle$ denoted by $\lim_{n\to\infty} x_n$ or $\lim x_n$ - defined by $a \in$ **R** such that

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(n \geq N \Rightarrow |x_n - a| < \epsilon)$$

  - $\lim x_n$ unique if exists
- $\langle x_n \rangle$ called Cauchy sequence if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(n, m \geq N \Rightarrow |x_n - x_m| < \epsilon)$$

- Cauchy criterion - characterizing complete metric space (including **R**)
  - sequence converges *if and only if* Cauchy sequence

# Other limits

- cluster point of $\langle x_n \rangle$ - defined by $c \in \mathbf{R}$

$$(\forall \epsilon > 0, N \in \mathbf{N})(\exists n > N)(|x_n - c| < \epsilon)$$

- limit superior or limsup of $\langle x_n \rangle$

$$\limsup x_n = \inf_n \sup_{k>n} x_k$$

- limit inferior or liminf of $\langle x_n \rangle$

$$\liminf x_n = \sup_n \inf_{k>n} x_k$$

- $\liminf x_n \leq \limsup x_n$
- $\langle x_n \rangle$ converges *if and only if* $\liminf x_n = \limsup x_n$ $(=\lim x_n)$

# Open and closed sets

- $O$ called *open* if

$$(\forall x \in O)(\exists \delta > 0)(y \in \mathbf{R})(|y - x| < \delta \Rightarrow y \in O)$$

  - intersection of finite collection of open sets is open
  - union of any collection of open sets is open

- $\overline{E}$ called *closure* of $E$ if

$$(\forall x \in \overline{E} \;\&\; \delta > 0)(\exists y \in E)(|x - y| < \delta)$$

- $F$ called *closed* if

$$F = \overline{F}$$

  - union of finite collection of closed sets is closed
  - intersection of any collection of closed sets is closed

# Open and closed sets - facts

- *every open set is union of countable collection of disjoint open intervals*

- (Lindelöf) any collection $\mathcal{C}$ of open sets has a countable subcollection $\langle O_i \rangle$ such that

$$\bigcup_{O \in \mathcal{C}} O = \bigcup_i O_i$$

  – equivalently, any collection $\mathcal{F}$ of closed sets has a countable subcollection $\langle F_i \rangle$ such that

$$\bigcap_{O \in \mathcal{F}} F = \bigcap_i F_i$$

# Covering and Heine-Borel theorem

- collection $\mathcal{C}$ of sets called *covering* of $A$ if

$$A \subset \bigcup_{O \in \mathcal{C}} O$$

  - $\mathcal{C}$ said to *cover* $A$
  - $\mathcal{C}$ called *open covering* if every $O \in \mathcal{C}$ is open
  - $\mathcal{C}$ called *finite covering* if $\mathcal{C}$ is finite

- *Heine-Borel theorem* - for any closed and bounded set, every open covering has finite subcovering

- corollary
  - any collection $\mathcal{C}$ of closed sets including at least one bounded set every finite subcollection of which has nonempty intersection has nonempty intersection.

# Continuous functions

- $f$ (with domain $D$) called *continuous at* $x$ if

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall y \in D)(|y - x| < \delta \Rightarrow |f(y) - f(x)| < \epsilon)$$

- $f$ called *continuous on* $A \subset D$ if $f$ is continuous at every point in $A$

- $f$ called *uniformly continuous on* $A \subset D$ if

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x, y \in D)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$$

# Continuous functions - facts

- $f$ is continuous *if and only if* for every open set $O$ (in co-domain), $f^{-1}(O)$ is open

- $f$ continuous on closed and bounded set is uniformly continuous

- *extreme value theorem* - $f$ continuous on closed and bounded set, $F$, is *bounded on $F$ and assumes its maximum and minimum on $F$*

$$(\exists x_1, x_2 \in F)(\forall x \in F)(f(x_1) \leq f(x) \leq f(x_2))$$

- *intermediate value theorem* - for $f$ continuous on $[a, b]$ with $f(a) \leq f(b)$,

$$(\forall d)(f(a) \leq d \leq f(b))(\exists c \in [a, b])(f(c) = d)$$

# Borel sets and Borel $\sigma$-algebra

- *Borel set*
  - any set that can be formed from open sets (or, equivalently, from closed sets) through the operations of countable union, countable intersection, and relative complement

- *Borel algebra* or *Borel $\sigma$-algebra*
  - *smallest $\sigma$-algebra containing all open sets*
  - also
    - smallest $\sigma$-algebra containing all closed sets
    - smallest $\sigma$-algebra containing all open intervals (due to statement on page 28)

# Various Borel sets

- countable union of closed sets (in **R**), called *an* $F_\sigma$ ($F$ for closed & $\sigma$ for sum)
  - thus, every countable set, every closed set, every open interval, every open sets, is an $F_\sigma$ (note $(a,b) = \bigcup_{n=1}^\infty [a+1/n, b-1/n]$)
  - countable union of sets in $F_\sigma$ again is an $F_\sigma$
- countable intersection of open sets called *a* $G_\delta$ ($G$ for open & $\delta$ for durchschnitt - average in German)
  - complement of $F_\sigma$ is a $G_\delta$ and vice versa
- $F_\sigma$ and $G_\delta$ are simple types of Borel sets
- countable intersection of $F_\sigma$'s is $F_{\sigma\delta}$, countable union of $F_{\sigma\delta}$'s is $F_{\sigma\delta\sigma}$, countable intersection of $F_{\sigma\delta\sigma}$'s is $F_{\sigma\delta\sigma\delta}$, $etc.$, & likewise for $G_{\delta\sigma\ldots}$
- below are all classes of Borel sets, but not every Borel set belongs to one of these classes

$$F_\sigma, F_{\sigma\delta}, F_{\sigma\delta\sigma}, F_{\sigma\delta\sigma\delta}, \ldots, G_\delta, G_{\delta\sigma}, G_{\delta\sigma\delta}, G_{\delta\sigma\delta\sigma}, \ldots,$$

# Lebesgue Measure

# Riemann integral

- Riemann integral
  - partition induced by sequence $\langle x_i \rangle_{i=1}^n$ with $a = x_1 < \cdots < x_n = b$
  - lower and upper sums
    * $L(f, \langle x_i \rangle) = \sum_{i=1}^{n-1} \inf_{x \in [x_i, x_{i+1}]} f(x)(x_{i+1} - x_i)$
    * $U(f, \langle x_i \rangle) = \sum_{i=1}^{n-1} \sup_{x \in [x_i, x_{i+1}]} f(x)(x_{i+1} - x_i)$
  - always holds: $L(f, \langle x_i \rangle) \leq U(f, \langle y_i \rangle)$, hence

$$\sup_{\langle x_i \rangle} L(f, \langle x_i \rangle) \leq \inf_{\langle x_i \rangle} U(f, \langle x_i \rangle)$$

  - Riemann integrable if

$$\sup_{\langle x_i \rangle} L(f, \langle x_i \rangle) = \inf_{\langle x_i \rangle} U(f, \langle x_i \rangle)$$

- every continuous function is Riemann integrable

# Motivation - want measure better than Riemann integrable

- consider indicator (or characteristic) function $\chi_{\mathbf{Q}} : [0,1] \to [0,1]$

$$\chi_{\mathbf{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbf{Q} \\ 0 & \text{if } x \not\in \mathbf{Q} \end{cases}$$

- *not* Riemann integrable: $\sup_{\langle x_i \rangle} L(f, \langle x_i \rangle) = 0 \neq 1 = \inf_{\langle x_i \rangle} U(f, \langle x_i \rangle)$

- however, irrational numbers infinitely more than rational numbers, hence
  - *want to* have some integral $\int$ such that, $e.g.$,

$$\int_{[0,1]} \chi_{\mathbf{Q}}(x) dx = 0 \text{ and } \int_{[0,1]} (1 - \chi_{\mathbf{Q}}(x)) dx = 1$$

## Properties of desirable measure

- want some measure $\mu : \mathcal{M} \to \mathbf{R}_+ = \{x \in \mathbf{R} | x \geq 0\}$
  - defined for every subset of **R**, $i.e.$, $\mathcal{M} = \mathcal{P}(\mathbf{R})$
  - equals to length for open interval

$$\mu[b, a] = b - a$$

  - countable additivity: for disjoint $\langle E_i \rangle_{i=1}^{\infty}$

$$\mu(\cup E_i) = \sum \mu(E_i)$$

  - translation invariant

$$\mu(E + x) = \mu(E) \text{ for } x \in \mathbf{R}$$

- *no* such measure exists
- *not* known whether measure with first three properties exists
- want to find translation invariant *countably additive measure*
  - hence, *give up on first property*

# Race won by Henri Lebesgue in 1902!

- mathematicians in 19th century struggle to solve this problem

- race won by French mathematician, *Henri Léon Lebesgue in 1902!*

- Lebesgue integral covers much wider range of functions
  - indeed, $\chi_{\mathbf{Q}}$ is Lebesgue integrable

$$\int_{[0,1]} \chi_{\mathbf{Q}}(x)dx = 0 \text{ and } \int_{[0,1]} (1 - \chi_{\mathbf{Q}}(x))dx = 1$$

# Outer measure

- for $E \subset \mathbf{R}$, define outer measure $\mu^* : \mathcal{P}(\mathbf{R}) \to \mathbf{R}_+$

$$\mu^* E = \inf_{\langle I_i \rangle} \left\{ \sum l(I_i) \middle| E \subset \cup I_i \right\}$$

  where $I_i = (a_i, b_i)$ and $l(I_i) = b_i - a_i$
- outer measure of open interval is length

$$\mu^*(a_i, b_i) = b_i - a_i$$

- countable subadditivity

$$\mu^* \left( \cup E_i \right) \leq \sum \mu^* E_i$$

- corollaries
  - $\mu^* E = 0$ if $E$ is countable
  - $[0, 1]$ not countable

# Measurable sets

- $E \subset \mathbf{R}$ called measurable if for every $A \subset \mathbf{R}$

$$\mu^* A = \mu^*(E \cup A) + \mu^*(\tilde{E} \cup A)$$

- $\mu^* E = 0$, then $E$ measurable
- every open interval $(a, b)$ with $a \geq -\infty$ and $b \leq \infty$ is measurable
- disjoint countable union of measurable sets is measurable, $i.e.$, $\cup E_i$ is measurable
- collection of measurable sets is $\sigma$-algebra

# Borel algebra is measurable

- note
  - every open set is disjoint countable union of open intervals (page 28)
  - disjoint countable union of measurable sets is measurable (page 40)
  - open intervals are measurable (page 40)
- hence, every open set is measurable
- also
  - collection of measurable sets is $\sigma$-algebra (page 40)
  - every open set is Borel set and Borel sets are $\sigma$-algebra (page 32)
- hence, *Borel sets are measurable*
- specifically, *Borel algebra (smallest $\sigma$-algebra containing all open sets) is measurable*

# Lebesgue measure

- restriction of $\mu^*$ in collection $\mathcal{M}$ of measurable sets called *Lebesgue measure*

$$\mu : \mathcal{M} \to \mathbf{R}_+$$

- countable subadditivity - for $\langle E_n \rangle$

$$\mu(\cup E_n) \leq \sum \mu E_n$$

- *countable additivity* - for disjoint $\langle E_n \rangle$

$$\mu(\cup E_n) = \sum \mu E_n$$

- for dcreasing sequence of measurable sets, $\langle E_n \rangle$, *i.e.*, $(\forall n \in \mathbf{N})(E_{n+1} \subset E_n)$

$$\mu\left(\bigcap E_n\right) = \lim \mu E_n$$

## (Lebesgue) measurable sets are nice ones!

- following statements are equivalent
  - $E$ is measurable
  - $(\forall \epsilon > 0)(\exists \text{ open } O \supset E)(\mu^*(O \sim E) < \epsilon)$
  - $(\forall \epsilon > 0)(\exists \text{ closed } F \subset E)(\mu^*(E \sim F) < \epsilon)$
  - $(\exists G_\delta)(G_\delta \supset E)(\mu^*(G_\delta \sim E) < \epsilon)$
  - $(\exists F_\sigma)(F_\sigma \subset E)(\mu^*(E \sim F_\sigma) < \epsilon)$
- if $\mu^* E$ is finite, above statements are equivalent to

$$(\forall \epsilon > 0) \left( \exists U = \bigcup_{i=1}^{n} (a_i, b_i) \right) (\mu^*(U \Delta E) < \epsilon)$$

# Lebesgue measure resolves problem in movitation

- let

$$E_1 = \{x \in [0,1] | x \in \mathbf{Q}\},\ E_2 = \{x \in [0,1] | x \not\in \mathbf{Q}\}$$

- $\mu^* E_1 = 0$ because $E_1$ is countable, hence measurable and

$$\mu E_1 = \mu^* E_1 = 0$$

- algebra implies $E_2 = [0,1] \cap \tilde{E}_1$ is measurable

- countable additivity implies $\mu E_1 + \mu E_2 = \mu[0,1] = 1$, hence

$$\mu E_1 = 1$$

# Lebesgue Measurable Functions

## Lebesgue measurable functions

- for $f : X \to \mathbf{R} \cup \{-\infty, \infty\}$, $i.e.$, extended real-valued function, the followings are equivalent
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) < a\}$ is measurable
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) \leq a\}$ is measurable
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) > a\}$ is measurable
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) \geq a\}$ is measurable

- if so,
  - for every $a \in \mathbf{R} \cup \{-\infty, \infty\}$, $\{x \in X | f(x) = a\}$ is measurable

- extended real-valued function, $f$, called *(Lebesgue) measurable function* if
  - domain is measurable
  - any one of above four statements holds

(refer to page **??** for abstract counterpart)

# Properties of Lebesgue measurable functions

- for real-valued measurable functions, $f$ and $g$, and $c \in \mathbf{R}$
  - $f + c$, $cf$, $f + g$, $fg$ are measurable

- for every extended real-valued measurable function sequence, $\langle f_n \rangle$
  - $\sup f_n$, $\limsup f_n$ are measurable
  - hence, $\inf f_n$, $\liminf f_n$ are measurable
  - thus, if $\lim f_n$ exists, it is measurable

(refer to page **??** for abstract counterpart)

# Almost everywhere - a.e.

- statement, $P(x)$, called *almost everywhere* or *a.e.* if

$$\mu\{x| \sim P(x)\} = 0$$

  - $e.g.$, $f$ said to be equal to $g$ a.e. if $\mu\{x|f(x) \neq g(x)\} = 0$
  - $e.g.$, $\langle f_n \rangle$ said to converge to $f$ a.e. if

$$(\exists E \text{ with } \mu E = 0)(\forall x \not\in E)(\lim f_n(x) = f(x))$$

- facts
  - if $f$ is measurable and $f = g$ i.e., then $g$ is measurable
  - if measurable extended real-valued $f$ defined on $[a, b]$ with $f(x) \in$ **R** a.e., then for every $\epsilon > 0$, exist step function $g$ and continuous function $h$ such that

$$\mu\{x||f - g| \geq \epsilon\} < \epsilon, \ \mu\{x||f - h| \geq \epsilon\} < \epsilon$$

# Characteristic & simple functions

- for any $A \subset \mathbf{R}$, $\chi_A$ called *characteristic function* if

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \not\in A \end{cases}$$

  - $\chi_A$ is measurable *if and only if* $A$ is measurable

- measurable $\varphi$ called *simple* if for some distinct $\langle a_i \rangle_{i=1}^n$

$$\varphi(x) = \sum_{i=1}^n a_i \chi_{A_i}(x)$$

  where $A_i = \{x | x = a_i\}$

(refer to page **??** for abstract counterpart)

# Littlewood's three principles

let $M(E)$ with measurable set, $E$, denote set of measurable functions defined on $E$

- *every (measurable) set is nearly finite union of intervals*, $e.g.$,
  - $E$ is measurable *if and only if*

$$(\forall \epsilon > 0)(\exists \{I_i : \text{open interval}\}_{i=1}^{n})(\mu^*(E \Delta (\cup I_n)) < \epsilon)$$

- *every (measurable) function is nearly continuous*, $e.g.$,
  - (Lusin's theorem)

$$(\forall f \in M[a,b])(\forall \epsilon > 0)(\exists g \in C[a,b])(\mu\{x | f(x) \neq g(x)\} < \epsilon)$$

- *every convergent (measurable) function sequence is nearly uniformly convergent*, $e.g.$,

$$(\forall \text{ measurable } \langle f_n \rangle \text{ converging to } f \text{ a.e. on } E \text{ with } \mu E < \infty)$$

$$(\forall \epsilon > 0 \text{ and } \delta > 0)(\exists A \subset E \text{ with } \mu(A) < \delta \text{ and } N \in \mathbf{N})$$

$$(\forall n > N, x \in E \sim A)(|f_n(x) - f(x)| < \epsilon)$$

# Egoroff's theorem

- *Egoroff theorem* - provides stronger version of third statement on page 50

$$(\forall \text{ measurable } \langle f_n \rangle \text{ converging to } f \text{ a.e. on } E \text{ with } \mu E < \infty)$$

$$(\exists A \subset E \text{ with } \mu(A) < \epsilon)(f_n \text{ uniformly converges to } f \text{ on } E \sim A)$$

# Lebesgue Integral

# Integral of simple functions

- *canonical representation* of simple function

$$\varphi(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x)$$

where $a_i$ are *distinct* $A_i = \{x | \varphi(x) = a_i\}$ - note $A_i$ are *disjoint*

- when $\mu\{x|\varphi(x) \neq 0\} < \infty$ and $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$ is canonical representation, define *integral of $\varphi$* by

$$\int \varphi = \int \varphi(x) dx = \sum_{i=1}^{n} a_i \mu A_i$$

- when $E$ is measurable, define

$$\int_E \varphi = \int \varphi \chi_E$$

(refer to page **??** for abstract counterpart)

# Properties of integral of simple functions

- for simple functions $\varphi$ and $\psi$ that vanish out of finite measure set, *i.e.*, $\mu\{x|\varphi(x) \neq 0\} < \infty$, $\mu\{x|\psi(x) \neq 0\} < \infty$, and for every $a, b \in \mathbf{R}$

$$\int (a\varphi + b\psi) = a \int \varphi + b \int \psi$$

(refer to page **??** for abstract counterpart)

- thus, even for simple function, $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$ that vanishes out of finite measure set, not necessarily in canonical representation,

$$\int \varphi = \sum_{i=1}^{n} a_i \mu A_i$$

- if $\varphi \geq \psi$ a.e.

$$\int \varphi \geq \int \psi$$

## Lebesgue integral of bounded functions

- for bounded function, $f$, and finite measurable set, $E$,

$$\sup_{\varphi\text{: simple, } \varphi \leq f} \int_E \varphi \leq \inf_{\psi\text{: simple, } f \leq \psi} \int_E \psi$$

  - if $f$ is defined on $E$, $f$ is measurable function *if and only if*

$$\sup_{\varphi\text{: simple, } \varphi \leq f} \int_E \varphi = \inf_{\psi\text{: simple, } f \leq \psi} \int_E \psi$$

- for bounded measurable function, $f$, defined on measurable set, $E$, with $\mu E < \infty$, define *(Lebesgue) integral of $f$ over $E$*

$$\int_E f(x)dx = \sup_{\varphi\text{: simple, } \varphi \leq f} \int_E \varphi = \inf_{\psi\text{: simple, } f \leq \psi} \int_E \psi$$

(refer to page **??** for abstract counterpart)

# Properties of Lebesgue integral of bounded functions

- for bounded measurable functions, $f$ and $g$, defined on $E$ with finite measure
  - for every $a, b \in \mathbf{R}$
$$\int_E (af + bg) = a \int_E f + b \int_E g$$
  - if $f \leq g$ a.e.
$$\int_E f \leq \int_E g$$
  - for disjoint measurable sets, $A, B \subset E$,
$$\int_{A \cup B} f = \int_A f + \int_B f$$

(refer to page **??** for abstract counterpart)

- hence,
$$\left| \int_E f \right| \leq \int_E |f| \ \& \ f = g \text{ a.e. } \Rightarrow \int_E f = \int_E g$$

# Lebesgue integral of bounded functions over finite interval

- if bounded function, $f$, defined on $[a,b]$ is Riemann integrable, then $f$ is measurable and
$$\int_{[a,b]} f = R\int_a^b f(x)dx$$
where $R\int$ denotes Riemann integral

- bounded function, $f$, defined on $[a,b]$ is Riemann integrable *if and only if* set of points where $f$ is discontinuous has measure zero

- for sequence of measurable functions, $\langle f_n \rangle$, defined on measurable $E$ with finite measure, and $M > 0$, if $|f_n| < M$ for every $n$ and $f(x) = \lim f_n(x)$ for every $x \in E$
$$\int_E f = \lim \int_E f_n$$

# Lebesgue integral of nonnegative functions

- for nonnegative measurable function, $f$, defined on measurable set, $E$, define

$$\int_E f = \sup_{h\text{: bounded measurable function, } \mu\{x|h(x)\neq 0\}<\infty,\ h\leq f} \int_E h$$

(refer to page **??** for abstract counterpart)

- for nonnegative measurable functions, $f$ and $g$
  - for every $a, b \geq 0$
$$\int_E (af + bg) = a \int_E f + b \int_E g$$
  - if $f \geq g$ a.e.
$$\int_E f \leq \int_E g$$
- thus,
  - for every $c > 0$
$$\int_E cf = a \int_E f$$

# Fatou's lemma and monotone convergence theorem for Lebesgue integral

- *Fatou's lemma* - for nonnegative measurable function sequence, $\langle f_n \rangle$, with $\lim f_n = f$ a.e. on measurable set, $E$

$$\int_E f \leq \liminf \int_E f_n$$

  – note $\lim f_n$ is measurable (page 47), hence $f$ is measurable (page 48)

- *monotone convergence theorem* - for nonnegative increasing measurable function sequence, $\langle f_n \rangle$, with $\lim f_n = f$ a.e. on measurable set, $E$

$$\int_E f = \lim \int_E f_n$$

(refer to page **??** for abstract counterpart)

- for nonnegative measure function, $f$, and sequence of disjoint measurable sets, $\langle E_i \rangle$,

$$\int_{\cup E_i} f = \sum \int_{E_i} f$$

# Lebesgue integrability of nonnegative functions

- nonnegative measurable function, $f$, said to be *integrable* over measurable set, $E$, if

$$\int_E f < \infty$$

(refer to page **??** for abstract counterpart)

- for nonnegative measurable functions, $f$ and $g$, if $f$ is integrable on measurable set, $E$, and $g \leq f$ a.e. on $E$, then $g$ is integrable and

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

- for nonnegative integrable function, $f$, defined on measurable set, $E$, and every $\epsilon$, exists $\delta > 0$ such that for every measurable set $A \subset E$ with $\mu A < \epsilon$ (then $f$ is integrable on $A$, of course),

$$\int_A f < \epsilon$$

# Lebesgue integral

- for (any) function, $f$, define $f^+$ and $f^-$ such that for every $x$

$$
\begin{aligned}
f^+(x) &= \max\{f(x), 0\} \\
f^-(x) &= \max\{-f(x), 0\}
\end{aligned}
$$

- note $f = f^+ - f^-$, $|f| = f^+ + f^-$, $f^- = (-f)^+$
- measurable function, $f$, said to be *(Lebesgue) integrable* over measurable set, $E$, if (nonnegative measurable) functions, $f^+$ and $f^-$, are integrable

$$
\int_E f = \int_E f^+ - \int_E f^-
$$

(refer to page **??** for Lebesgue counterpart)

# Properties of Lebesgue integral

- for $f$ and $g$ integrable on measure set, $E$, and $a, b \in \mathbf{R}$
  - $af + bg$ is integral and

$$\int_E (af + bg) = a \int_E f + b \int_E g$$

  - if $f \geq g$ a.e. on $E$,

$$\int_E f \geq \int_E g$$

  - for disjoint measurable sets, $A, B \subset E$

$$\int_{A \cup B} f = \int_A f + \int_B g$$

(refer to page **??** for abstract counterpart)

# Lebesgue convergence theorem (for Lebesgue integral)

- *Lebesgue convergence theorem -* for measurable $g$ integrable on measurable set, $E$, and measurable sequence $\langle f_n \rangle$ converging to $f$ with $|f_n| < g$ a.e. on $E$, ($f$ is measurable (page 47), every $f_n$ is integrable (page 60)) and

$$\int_E f = \lim \int_E f_n$$

(refer to page **??** for abstract counterpart)

# Generalization of Lebesgue convergence theorem (for Lebesgue integral)

- *generalization of Lebesgue convergence theorem* - for sequence of functions, $\langle g_n \rangle$, integrable on measurable set, $E$, converging to integrable $g$ a.e. on $E$, and sequence of measurable functions, $\langle f_n \rangle$, converging to $f$ a.e. on $E$ with $|f_n| < g_n$ a.e. on $E$, if

$$\int_E g = \lim \int_E g_n$$

then ($f$ is measurable (page 47), every $f_n$ is integrable (page 60)) and

$$\int_E f = \lim \int_E f_n$$

# Comments on convergence theorems

- Fatou's lemma (page 59), monotone convergence theorem (page 59), Lebesgue convergence theorem (page 63), *all* state that under suitable conditions, we say something about
$$\int \lim f_n$$
in terms of
$$\lim \int f_n$$
- Fatou's lemma requires weaker condition than Lebesgue convergence theorem, $i.e.$, only requires "bounded below" whereas Lebesgue converges theorem also requires "bounded above"
$$\int \lim f_n \leq \liminf \int f_n$$
- monotone convergence theorem is somewhat between the two;
  - advantage - applicable even when $f$ not integrable
  - Fatou's lemma and monotone converges theorem very clsoe in sense that can be derived from each other using only facts of positivity and linearity of integral

## Convergence in measure

- $\langle f_n \rangle$ of measurable functions said to *converge $f$ in measure* if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(\forall n > N)(\mu\{x||f_n - f| > \epsilon\} < \epsilon)$$

- thus, third statement on page 50 implies

$$(\forall \langle f_n \rangle \text{ converging to } f \text{ a.e. on } E \text{ with } \mu E < \infty)(f_n \text{ converge in measure to } f)$$

- however, the converse is *not* true, *i.e.*, exists $\langle f_n \rangle$ converging in measure to $f$ that does not converge to $f$ a.e.
  - *e.g.*, XXX

- Fatou's lemma (page 59), monotone convergence theorem (page 59), Lebesgue convergence theorem (page 63) *remain valid!* even when "convergence a.e." replaced by "convergence in measure"

# Conditions for convergence in measure

**Proposition 1. [necessary condition for converging in measure]**

$$(\forall \langle f_n \rangle \text{ converging in measure to } f)\left(\exists \text{ subsequence } \left\langle f_{n_k} \right\rangle \text{ converging a.e. to } f\right)$$

**Corollary 1. [necessary and sufficient condition for converging in measure]** *for sequence* $\langle f_n \rangle$ *measurable on* $E$ *with* $\mu E < \infty$

$$\begin{aligned}&\langle f_n \rangle \text{ converging in measure to } f\\ &\quad\Leftrightarrow\quad (\forall \text{ subsequence } \left\langle f_{n_k} \right\rangle)\left(\exists \text{ its subsequence } \left\langle f_{n_{k_l}} \right\rangle \text{ converging a.e. to } f\right)\end{aligned}$$

# References

# References

[Roy88] H.L. Royden. *Real Analysis*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632, USA, 3rd edition, 1988.

# Index