Measure, Probability and Stochastic Process

A rigorous but painless introduction

shuqi

October 1, 2023

xiaosq2000 xiaosq2000@gmail.com

Outline i

1. Measure Theory and Lebesgue Integration
   - Measure
   - σ-algebra
   - Measurable space and measure space
   - Measurable function
   - Lebesgue Measure
   - Lebesgue integral
   - Radon–Nikodym theorem

2. Probability Theory
   Probability
   Random variable

3. Stochastic Process
   Stochastic process
   Autocorrelation and autocovariance
   Stationarity
   Spectral analysis

Measure Theory and Lebesgue Integration

Measure

Remark (motivation of measure)

A measure is a generalization and formalization of geometrical measures (length, area, volume) and other common notions, such as magnitude, mass, and probability of events. It is fundamental in many mathematical fields, such as probability and integration.

Definition (measure)

Let X be a set and $\mathcal{F}$ a σ-algebra over X. A function $\mu : \mathcal{F} \mapsto \mathbb{R}^1_\infty$, where $\mathbb{R}^1_\infty$ is the extended real number field, is called a measure if the following three conditions hold:

1. empty is zero: $\mu(\emptyset) = 0$
2. non-negativity: $\forall E \in \mathcal{F} \, (\mu(E) \geq 0)$
3. special countable-additivity: $\mu(\bigcup_{k=1}^{\infty} E_k) = \Sigma_{k=1}^{\infty} \mu(E_k)$, where $\{E_k\}_{k=1}^{\infty}$ is all countable collections of pairwise disjoint sets in $\mathcal{F}$

Remark (σ-algebra)

The "σ-algebra" and "countable" (actually, closely related to σ-algebra) make the rigorous definition of measure (def. 4) peculiar.

You can check out the definition of σ-algebra and motivations about it in measure theory on Wikipedia [1] , which is quite enlightening. In summary,

- Introducing the set algebra to deliver the addition-preserving property of a measure is natural, and σ-algebra is a set algebra with countable-additivity, alias σ-additivity. But why countable? [2] is a good explanation.
- ZFC (precisely, axiom of choice) entails non-measurable set of $\mathbb{R}^n$, i.e., it is actually impossible to assign a length to all subsets of $\mathbb{R}$ in a way that preserves some natural additivity and translation invariance properties. The Vitali set and the Banach–Tarski paradox are famous examples.

Measurable space and measure space i

Definition (measurable space and measure space)

measurable space

$$\overbrace{(X, \mathcal{F})}^{\text{tuple}} \tag{1}$$

a set

a σ-algebra on X

measure space

$$\overbrace{(X, \mathcal{F}, \mu)}^{\text{triple}} \tag{2}$$

a set

a σ-algebra on X

a measure on $(X, \mathcal{F})$

Measurable function

Definition (measurable function)

Let (X, Σ) and (Y, T) be measurable spaces. A function $f : X \mapsto Y$ is measurable if and only if

$$\forall E \in T \left(f^{-1}(E) \in \Sigma \right) \tag{3}$$

Corollary

$$f \textit{ is measurable} \Leftrightarrow \sigma(f) \subset \Sigma, \tag{4}$$

where $\sigma(f)$ is the σ-algebra generated by f.

Remark (motivation)

Lebesgue measure is the formalized way of assigning a volume to subsets of $\mathbb{R}^n$, and coincides with the standard measure of length ($\mathbb{R}$), area ($\mathbb{R}^2$), or volume ($\mathbb{R}^3$).

Definition (length of an interval)

For any interval $\mathrm{I} = [a, b]$ or $(a, b) \subset \mathbb{R}$, let $\ell(\mathrm{I}) = b - a$ denote its length.

Lebesgue measure ii

Definition (Lebesgue outer measure)

For any subset $\mathbf{E} \subseteq \mathbb{R}$, its Lebesgue outer measure $\lambda^* : \mathfrak{p}(\mathbb{R}) \mapsto [0, +\infty]$ is defined as[1]

$$\lambda^*(\mathrm{E}) := \inf \left\{ \sum_{k=1}^{\infty} \mathfrak{l}(\mathrm{I}_k) : \{\mathrm{I}_k\}_{k \in \mathbb{N}} \text{ is a sequence of open intervals with } \mathbf{E} \subset \bigcup_{k=1}^{\infty} \mathrm{I}_k \right\} \tag{5}$$

Definition (Carathéodory's criterion)

$$\lambda^*(\mathrm{E}) = \lambda^*(\mathrm{E} \cap \mathrm{A}) + \lambda^*(\mathrm{E} \cap \mathrm{A}^c), \quad \forall \mathbf{A} \subset \mathbb{R} \tag{6}$$

Definition (Lebesgue measure)

The set of all such $\mathbf{E} \subseteq \mathbb{R}$ that fulfills the Carathéodory criterion forms a σ-algebra, its Lebesgue measure is defined to be its Lebesgue outer measure,

$$\lambda(\mathbf{E}) = \lambda^*(\mathbf{E}) \tag{7}$$

Remark (motivation of an outer measure and a criterion)

There are four intuitive requirements of a Lebesgue measure, but they are incompatible due to *ZFC*. The purpose of constructing an outer measure on all subsets is to pick out a class of subsets to be called measurable, in such a way (the criterion) to satisfy the countable additivity property. Check out [3] for details.

[1] $\mathfrak{p}(\circ)$ denotes the power set of $\circ$.

Remark

Like step functions in Riemann-Darboux's approach, the concept of Lebesgue integration is built on indicator functions.

Definition (indicator function)

A indicator function of a subset S of a set X is denoted by $\mathbf{1}_S : X \mapsto \{0, 1\}$, defined as

$$\mathbf{1}_S(x) := \begin{cases} 1, & x \in S \qquad (8) \\ 0, & x \notin S \qquad (9) \end{cases}$$

Lebesgue integral ii

Definition (Lebesgue integral of an indicator function)

If the subset S is measurable consistent with a given measure μ, the Lebesgue integral of the indicator function is defined as

$$\int_X 1_S \, \mathrm{d}\mu := \int_S 1_S \, \mathrm{d}\mu := \mu(S) \tag{10}$$

Note that the integral may be equal to ∞, unless μ is a finite measure.

If T is another measurable subset of X,

$$\int_T 1_S \, \mathrm{d}\mu := \mu(S \cap T) \tag{11}$$

Definition (simple function)

A simple function is a finite linear combination of indicator functions

$$\Sigma_{k=0}^{N} a_k \mathbf{1}_{S_k} \tag{12}$$

If $a_k \in \mathbb{R}$ and S_k are disjoint measureable sets, it is called a measurable simple function.

Definition (Lebesgue integral of a measurable simple function)

Linearity is a desired property for Lebesgue integrals, i.e.,its Lebesgue outer measure is defined as

$$\int_X \Sigma_{k=1}^{N} a_k \mathbf{1}_{S_k} \, \mathrm{d}\mu := \Sigma_{k=1}^{N} a_k \int_X \mathbf{1}_{S_k} \, \mathrm{d}\mu = \Sigma_{k=1}^{N} a_k \mu(S_k) \,, \quad \text{if } (\forall\, k : \mu_k = \infty \Rightarrow a_k = 0) \tag{13}$$

to avoid $\infty - \infty$

To compute the Riemann integral of f, one partitions the domain $[a, b]$ into subintervals; while in the Lebesgue integral, one is in effect partitioning the range of f.

— *Folland*

Definition (Lebesgue integral of a non-negative measurable function)

$$\int_E f \, \mathrm{d}\mu := \sup \left\{ \int_E s \, \mathrm{d}\mu : 0 < s < f \wedge s \text{ is simple} \right\} \tag{14}$$

Definition (Lebesgue integral of a measurable function)

$$f = f^+ + f^- \tag{15}$$

, where

$$f^+ := \begin{cases} f(x) & , \quad f(x) > 0 \quad (16) \\ 0 & , \quad \text{otherwise} \quad (17) \end{cases} \qquad f^- := \begin{cases} -f(x), & f(x) < 0 \quad (18) \\ 0 & , \quad \text{otherwise} \quad (19) \end{cases}$$

Then,

$$\int f \, \mathrm{d}\mu := \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu, \quad \text{if } \boxed{\min \left(\int f^+ \, \mathrm{d}\mu, \int f^- \, \mathrm{d}\mu \right) < \infty} \tag{20}$$

at least one is finite, to avoid $\infty - \infty$

Definition (absolute continuity of measures)

Definition (Radon–Nikodym theorem)

Probability Theory

Probability

Definition (probability)

Kolmogorov Axioms

$$\overbrace{\left(\underbrace{\Omega}_{\text{sample space}}, \underbrace{\mathcal{F}}_{\text{event space}}, \underbrace{P}_{\text{probability measure}}\right)}^{\text{probability space}} \tag{21}$$

the probability is a measure with two additional properties,

1. finiteness: $\forall E \in \mathcal{F}\,(P(E) \in \mathbb{R})$
2. unitarity: $P(\Omega) = 1$

Random variable

Definition (random variable)

$$\underbrace{X}_{\text{measurable function}} : \underbrace{\Omega}_{\text{sample space}} \mapsto \underbrace{S}_{\text{state space}} \tag{22}$$

Remark (random variable)

It is a function but called "variable" to emphsaize on its codomain(state space), usually subsets of $\mathbb{R}^n$ or $\mathbb{Z}^n$, which is more convenient for manipulation than the abstract sample space. For example, the event $E := \{\omega \in \Omega : u < X(\omega) \leq v\}$ is usually denoted by $u < X \leq v$, since $\omega \in X^{-1}((u, v]) \Leftrightarrow u < X(\omega) \leq v$.

Stochastic Process

Stochastic process

Definition (stochastic process)

A stochastic process is collection of indexed random variables, denoted by

$$\{X(t) : t \in T\}, \tag{23}$$

where T is the index/parameter set.

Remark (index set)

t usually has a physical meaning of time(continuous) or timestamp(discrete).

Autocorrelation and autocovariance

Let $\mathsf{x}(\omega, t) : \Omega \times \mathbb{R} \mapsto \mathbb{R}^n$ be a continuous-time multivariate real-valued stochastic process[2],

Definition (autocorrelation)

$$\mathsf{R}_{\mathsf{xx}}(t_1, t_2) = \mathsf{E}\left(\mathsf{x}(t_1)\mathsf{x}(t_2)^{\mathrm{T}}\right) \tag{24}$$

Definition (autocovariance)

$$\mathsf{K}_{\mathsf{xx}}(t_1, t_2) = \mathsf{Cov}\left(\mathsf{x}(t_1), \mathsf{x}(t_2)\right) = \mathsf{E}\left(\left(\mathsf{x}(t_1) - \mathsf{E}\left(\mathsf{x}(t_1)\right)\right)\left(\mathsf{x}(t_2) - \mathsf{E}\left(\mathsf{x}(t_2)\right)\right)^{\mathrm{T}}\right) \tag{25}$$

[2] The continuous-time multivariate real-valued stochastic process is the most common in the author's background, so if not mentioned, the following definitions, remarks, etc., are based on it.

Definition (strict stationary process)

Let $F_X\left(X_{t_1+\tau}, \cdots, X_{t_n+\tau}\right)$ be the cumulative distribution function(CDF) of the unconditional joint distribution of the stochastic process $\{X_t\}$ at times $t_1+\tau, \cdots, t_n+\tau$.

$\{X_t\}$ is a (strict(ly)/strong(ly)) stationary process, if the unconditional joint CDF does not change when shifted in time, i.e.

$$\left(\forall\, \tau, t_1, \cdots, t_n \in \mathbb{R}\right)\left(\forall\, n \in \mathbb{N}_+\right)\left(F_X\left(x_{t_1+\tau}, \cdots, x_{t_n+\tau}\right)\right) \tag{26}$$

Definition (wide stationary process)

A wide/weak stationary process loosens the constraints on CDF(eq. 22) to the following first two conditions, with an additional “finite second-moment” condition (eq. 29).

$$\mathrm{E}\left(\mathrm{x}(t+\tau)\right) = \mathrm{E}\left(\mathrm{x}(t)\right), \qquad \forall\, t, \tau \in \mathbb{R} \tag{27}$$

$$\mathrm{K}_{\mathrm{xx}}\left(t_1, t_2\right) = \mathrm{K}_{\mathrm{xx}}\left(t_1 - t_2, 0\right), \qquad \forall\, t_1, t_2 \in \mathbb{R} \tag{28}$$

$$\mathrm{E}\left(|\mathrm{x}_t|^2\right) < \infty, \qquad \forall\, t \in \mathbb{R} \tag{29}$$

Stationarity iii

Corollary (wide stationary process)

- *The expectation is always a constant.*
- *The autocovariance and autocorrelation are better indexed by one variable (time difference) instead of two (timestamps).*
- *Any strictly stationary process which has a finite mean and a covariance is also a wide-sense stationary process.*

Remark (motivation of wide-sense stationarity, WSS)

The "finite second-moment" condition(eq. 29) may remind you of the Hilbert space, if you are good at associating.

The Wikipedia [4] has a wonderful explanation of its mathematical motivation and the reason why the WSS assumption is widely employed in signal processing algorithms.

Definition (energy)

$$E := \int_{-\infty}^{+\infty} \|\mathrm{x}(t)\|^2 \,\mathrm{d}t \tag{30}$$

Theorem (Parseval's theorem)

$$\int_{-\infty}^{+\infty} \|\mathrm{x}(t)\|^2 \,\mathrm{d}t = \int_{-\infty}^{+\infty} \|\hat{\mathrm{x}}(f)\|^2 \,\mathrm{d}f, \tag{31}$$

where $\hat{\mathrm{x}}(f)$ *is the Fourier transform of* $\mathrm{x}(t)$*, i.e.,*

$$\hat{\mathrm{x}}(f) = \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}2\pi f t}\mathrm{x}(t) \,\mathrm{d}t \tag{32}$$

Spectral analysis ii

Definition (energy spectral density)

$$\bar{S}_{\mathsf{xx}} := \|\hat{\mathsf{x}}(f)\|^2 \tag{33}$$

References

[1] Wikipedia contributors. *σ-algebra — Wikipedia, The Free Encyclopedia.* URL: https://en.wikipedia.org/w/index.php?title=%CE%A3-algebra&oldid=1173641705 (cit. on p. 6).

[2] Carl Mummert. *Why do we want probabilities to be countably additive?* Mathematics Stack Exchange. URL: https://math.stackexchange.com/q/566154 (cit. on p. 6).

[3] Wikipedia contributors. *Outer measure — Wikipedia, The Free Encyclopedia.* 2023. URL: https://en.wikipedia.org/w/index.php?title=Outer_measure&oldid=1171717640 (cit. on p. 11).

[4] Wikipedia contributors. *Stationary process — Wikipedia, The Free Encyclopedia.* motivation of weak-sense stationarity. URL: https://en.wikipedia.org/wiki/Stationary_process#Motivation (cit. on p. 28).