Measure, Probability and Stochastic Process

A rigorous but painless introduction

shuqi

September 29, 2023

xiaosq2000 xiaosq2000@gmail.com

Outline i

Measure Theory

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

σ-algebra

Remark (σ-algebra)

The "σ-algebra" and "countable" (actually, closely related to σ-algebra) make the rigorous definition of measure (def. 3) 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.

Other basic concepts 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 $(X, \mathcal{F})$

Other basic concepts ii

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

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

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

Probability Theory

Probability

Definition (probability)

Kolmogorov Axioms

$$\overbrace{\left(\underset{\text{sample space}}{\Omega}, \underset{\text{event space}}{\mathcal{F}}, \underset{\text{probability measure}}{P}\right)}^{\text{probability space}} \tag{5}$$

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)

$$\overbrace{X}^{\text{measurable function}} : \underbrace{\Omega}_{\text{sample space}} \mapsto \underbrace{S}_{\text{state space}} \tag{6}$$

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

$$\{X(t) : t \in T\}\,, \tag{7}$$

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[1],

Definition (autocorrelation)

$$\mathsf{R}_{\mathsf{xx}}(t_1, t_2) = \mathsf{E}\left(\mathsf{x}(t_1)\mathsf{x}(t_2)^{\mathrm{T}}\right) \tag{8}$$

Definition (autocovariance)

$$\mathsf{K}_{\mathsf{xx}}(t_1, t_2) = \mathsf{Cov}\left(\mathsf{x}(t_1), \mathsf{x}(t_2)\right) = \mathsf{E}\left(\left(\mathsf{x}(t_1) - \mathsf{E}\left(\mathsf{x}(t_1)\right)\right)\left(\mathsf{x}(t_2) - \mathsf{E}\left(\mathsf{x}(t_2)\right)\right)^{\mathrm{T}}\right) \tag{9}$$

[1] 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.

Stationarity i

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.

$$(\forall\,\tau,t_1,\cdots,t_n\in\mathbb{R})\,(\forall\,n\in\mathbb{N}_+)\,\left(F_X\left(x_{t_1+\tau},\cdots,x_{t_n+\tau}\right)\right)\tag{10}$$

Definition (wide stationary process)

A wide/weak stationary process loosens the constraints on CDF(eq. 10) to the following first two conditions, with an additional "finite second-moment" condition (eq. 13).

$$\mathrm{E}\left(\mathrm{x}(t+\tau)\right) = \mathrm{E}\left(\mathrm{x}(t)\right), \qquad \forall\, t, \tau \in \mathbb{R} \tag{11}$$

$$\mathrm{K}_{\mathrm{xx}}\left(t_1, t_2\right) = \mathrm{K}_{\mathrm{xx}}\left(t_1 - t_2, 0\right), \qquad \forall\, t_1, t_2 \in \mathbb{R} \tag{12}$$

$$\mathrm{E}\left(|\mathrm{x}_t|^2\right) < \infty, \qquad \forall\, t \in \mathbb{R} \tag{13}$$

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.*

Stationarity iv

Remark (motivation of wide-sense stationarity, WSS)

The “finite second-moment” condition(eq. 13) may remind you of the Hilbert space.

The Wikipedia [3] has a wonderful explanation of its mathematical motivation and the reason why the WSS assumption is widely employed in signal processing algorithms.

Definition (energy)

$$E := \int_{-\infty}^{+\infty} \|\mathsf{x}(t)\|^2 \,\mathrm{d}t \tag{14}$$

Theorem (Parseval's theorem)

$$\int_{-\infty}^{+\infty} \|\mathsf{x}(t)\|^2 \,\mathrm{d}t = \int_{-\infty}^{+\infty} \|\hat{\mathsf{x}}(f)\|^2 \,\mathrm{d}f, \tag{15}$$

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

$$\hat{\mathsf{x}}(f) = \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}2\pi f t} \mathsf{x}(t) \,\mathrm{d}t \tag{16}$$

Spectral analysis ii

Definition (energy spectral density)

$$\bar{S}_{\mathrm{xx}} := \|\hat{\mathbf{x}}(f)\|^2 \tag{17}$$

References

References i

[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. 5).

[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. 5).

[3] 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. 17).