Measure Theoretic View of Random Variables

BOGNÁR Miklós

September 18, 2023

σ-algebras, measures and measurable functions

Let T be a set. We will refer to it as the *basis set*, on which we can construct our σ-algebra. σ here refers to *countable*. Let $\mathcal{S} \subseteq \mathcal{P}(T)$ the collection of subsets from T with the property that it is closed under the σ-union, σ-intersection and complement operations, so for any $A_1, A_2, \dots \in \mathcal{S}$ we have

$$\bigcup_i^\infty A_i \in \mathcal{S}$$

$$\bigcap_i^\infty A_i \in \mathcal{S}$$

$$A_i^C \in \mathcal{S}$$

We call $\mathcal{S}$ the *σ-algebra over T*.

We write the structure $(T, \mathcal{S})$ as a tuple, and we call it a *measurable space*. A *measure* is a function that assigns *non-negative real numbers* to subsets of T:

$$\mu : \mathcal{P}(T) \mapsto \mathbb{R}^{+} \cup \{0\}$$

Note that *not every possible subset of T is measurable*, in fact, the ones that are measurable we call *measurable sets*. A measure has to satisfy some additional properties, namely:

$$\mu(\emptyset) = 0$$

$$\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$$

for disjoint sets $E_i \in \mathcal{S}$. We call the latter property σ-additivity.

We call the structure $(T, \mathcal{S}, \mu)$ a *measure space*. Note that in the case of non-disjoint sets, the σ-additivity becomes σ-subadditivity. If there is a set $F \in \mathcal{P}(T)$, then the σ-algebra *generated by* F is the smallest σ-algebra $\sigma(F)$, such that it contains every set that is in F.

Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be both measurable spaces. A function $f : X \mapsto Y$ is said to be *measurable* if for any $E \in \mathcal{T}$, the pre-image of E under f is contained in $\mathcal{S}$, that is:

$$f^{-1}(E) \doteq \{x \in X : f(x) \in E\} \in \mathcal{S}$$

We can write the function mapping in a way that emphasizes the σ-algebras:

$$f : (X, \mathcal{S}) \mapsto (Y, \mathcal{T})$$

The σ-algebra *generated by* f, denoted by $\sigma(f)$, is the set of pre-images

$$\sigma(f) \doteq \{f^{-1}(D) : D \in \mathcal{T}\}$$

Probability space and measure

Let Ω be the set of all possible outcomes of a given experiment. We call Ω the *sample space* or *state space* of the experiment. Let $\mathcal{F} \subseteq \Omega$ be a σ-algebra over Ω such that every *measurable* event is contained inside $\mathcal{F}$. Of course $\emptyset \in \mathcal{F}$, $\Omega \in \mathcal{F}$. Let us introduce the *probability measure* $\mathbb{P}$ as the third and last member of the *probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. In order for $\mathbb{P}$ to be a valid probability measure, it has to fulfill the following requirements:

$$\mathbb{P}(\emptyset) = 0$$

$$\mathbb{P}(\Omega) = 1$$

$$\forall A \in \mathcal{F} : \quad \mathbb{P}(A) \in [0, 1]$$

We can therefore state that the measure $\mathbb{P}$ is a *special type of measure* specifically used in aid of probability theory.

Random variables

A *measureable function* from the measurable space $(\Omega, \mathcal{F})$ to the real numbers equipped with the Borel σ-algebra $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called a *Random variable*, often denoted by the capital X or Y. The σ-algebra $\mathcal{F}$ is the algebra containing *all the events that can have probabilities assigned to them*.

$$X : (\Omega, \mathcal{F}) \mapsto (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

To make the notation easier, sometimes we write

$$X : \Omega \mapsto \mathbb{R}$$

The σ-algebra *generated by a random variable* X is the set

$$\sigma(X) \doteq \{\omega \in \Omega : X(\omega) \in C, \quad \forall C \in \mathcal{B}(\mathbb{R})\}$$

The distribution measure and function

A random variable induces a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, often denoted by μ_X, namely:

$$\mu_X(A) \mathrel{\mathring{=}} \mathbb{P}(X \in A), \quad \forall A \in \mathcal{B}(\mathbb{R})$$

For a given random variable X, the *distribution function* is the monotonically increasing function in the form

$$F_X(x) \mathrel{\mathring{=}} \mu_X((-\infty, x])$$

essentially a shorthand notation when working with distribution measures, as it is almost always easier to consider the range of the random variable instead of its domain.

The density function

Let us now consider the distribution measure in the form of the *integral measure*

$$\mu_X(E) = \int_E g_X d\lambda, \quad \forall E \in \mathcal{F}$$

where $g : \mathbb{R} \mapsto \mathbb{R}$ is a measurable function and λ is the Lebesgue-measure on $\mathbb{R}$. We call g_X the *Radon-Nikodym derivative* of μ_X with respect to λ:

$$\frac{d\mu_X}{d\lambda} = g_X$$

The notation here is symbolic, we can think of it as it we *differentiated both sides* of the integral measure formula with respect to the measure λ. In fact, g is exactly the *probability density function* of the random variable X. In the continuous real case, we can think of g_X as being the derivative of the distribution function F_X (if it exists):

$$g_X(x) = \frac{dF_X(x)}{dx}$$

although the measure-theoretical formalization is a lot less restrictive. Note that here the differentiation with respect to x is analogous with the differentiation with respect to λ, as the Lebesgue-measure can measure individual points. From the points mentioned above, it is also clear that $\forall x \in \mathbb{R}$:

$$F_X(x) = \int_{-\infty}^{x} g_X(y)dy$$

Independence of random variables

Let us conside the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let $\mathcal{A}$ and $\mathcal{B}$ be two *sub-σ-algebras* of $\mathcal{F}$ $(\mathcal{A}, \mathcal{B} \subseteq \mathcal{F})$. We say that $\mathcal{A}$ and $\mathcal{B}$ are *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

for $A \in \mathcal{A}, B \in \mathcal{B}$. We say that two random variables are independent, if *The σ-algebras generated by them are independent*.

Expectation and moments of random variables

The expectation (expected value) of a random variable X is the linear operator in the form of an *abstract integral* (*Lebesgue-integral*)

$$\mathbb{E}[X] \circeq \int_\Omega X d\mathbb{P}$$

where we integrate with respect to the probability measure $\mathbb{P}$. If the random variable is in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, meaning

$$\int_\Omega X^2 d\mathbb{P} < \infty$$

then the second moment of the random variable is

$$\mathbb{E}[X^2] \circeq \int_\Omega X^2 d\mathbb{P}$$

In general, the *p-th moment* of X is $\int_\Omega X^p d\mathbb{P}$ ($p > 0$).

Variance of random variables

The *variance* of a random variable is the expression

$$Var[X] \circeq \mathbb{E}[X^2] - \mathbb{E}^2[X]$$

Note that in order for the variance to be finite, we need both the first and second moments to be finite as well.