A First Look of Probability Measure

Wei-Ning Chen

August 4, 2016

Outline

In this lecture, we will give the probability theory a rigorous definition

Probability Triple

We say a probablity measure composes a probablity triple :

$$(\Omega, \mathcal{F}, \mathcal{P}), \text{ where}$$

- Ω : the sample space
- $\mathcal{F}$: event space, i.e. a collection of events
  (an event is a subset of Ω, i.e. if $A \in \mathcal{F}$, then $A \subset \Omega$)
- $\mathcal{P}$: a set function, such that

$$\mathcal{P} : \mathcal{F} \to [0, 1]$$

PROBABILITY MEASURE

The probability measure is a set function from $\mathcal{F}$ to $[0, 1]$ satisfies the following properties (probability axioms):

- $P(A) \in \mathbb{R}$, $P(A) \geq 0$, for all $A \in \mathcal{F}$
- $P(\phi) = 0^*$ and $P(\Omega) = 1$
- P is countably addtitive, i.e.

  if $A_1, A_2, ..., A_n, ...$ disjoint, then $P(A_1 \cup A_2 ... \cup A_n \cup ...) = \sum_{n=1}^{\infty} P(A_i)$

σ-FIELD

Now let's dive into the event space.

- What is a valid event ?
- Can we put all events into $\mathcal{F}$, i.e. $\mathcal{F} = 2^{\Omega}$?

The answer is yes, but something wierd will happen.

σ-FIELD

EXAMPLE (UNIFORM DISTRIBUTION)

Consider a uniform distibution X on $[0,1]$.
To reflect the fact that X is "uniform" on the interval $[0,1]$, the probability that X lies in some subset should be unaffected by "shifting" (with wrap-around) the subset by a fixed amount.
That is, if for each subset $A \subseteq [0,1]$, we define the **r-shift** by

$$A \oplus r := \{a + r; a \in A, a + r \leq 1\} \cup \{a + r - 1; a \in A, a + r > 1\}$$

then we have $P(A) = P(A \oplus r)$.

Proposition

There does not exist a definition of uniform probability $P(A)$, defined for all subsets $A \subset [0,1]$, satisfying the probability axioms. That is, there does not exist an uniform probability measure $\mathcal{P}$ defined on $\mathcal{F} = 2^{\Omega}$.

Proof.

Define an equivalence relation on $[0,1]$ by

$$x \sim y \text{ if and only if } y - x \text{ is rational.}$$

Let H be a subset of $[0,1]$ consisting of precisely one element from each equivalence class. For definiteness, assume that $0 \notin H$. Now, since H contains an element of each equivalence class, we see that each point in $(0,1]$ is contained in the union $\bigcup_{r \in [0,1),\ r \text{ rational}} (H \oplus r)$ of shifts of H. □

Proof (continue).

Since H contains just one point from each equivalence class, we see that these sets $H \oplus r$, for rational $r \in [0, 1)$, are all disjoint. But then, by countable additivity, we have

$$1 = P((0, 1]) = \sum_{r \in (0,1]} P(H \oplus r) = \sum_{r \in (0,1]} P(H).$$

Hence contrdiction occurs. □

Now it's time to define what we can add into our σ-field.

We say a collection of subsets of Ω is a σ-field if it satisfies

- $\phi \in \mathcal{F}$ and $\Omega \in \mathcal{F}(*)$
- If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$
- If $A_1, ..., A_n, ... \in \mathcal{F}$, then $A_1 \cup ... \cup A_n \cup ... \in \mathcal{F}$
- If $A_1, ..., A_n, ... \in \mathcal{F}$, then $A_1 \cap ... \cap A_n \cap ... \in \mathcal{F}(*)$

Question: How can we construct a σ-field?

DEFINITION (SEMIALGEBRA)

We say a collection $\mathcal{J}$ is a $semialgrebra$, if it contains 0 and ϕ, and is closed under finite intersection, and the complement of any element of $\mathcal{J}$ is equal to a finite disjoint union of elements of $\mathcal{J}$.

Since $\mathcal{J}$ is only a semialgebra, how can we create a cr-algebra?
As a first try, we might consider

$$B_0 = \{\text{all finite unions of elements of} \mathcal{J})$$

EXERCISE

- *Prove that B_0 is an algebra (or, field) of subsets of Ω, meaning that it contains ϕ, and Ω, and is closed under the formation of complements and of finite unions and intersections.*
- *Prove that B_0 is not a σ-algebra.*

As a second try, we might consider

$$B_1 = \{\text{all finite or countable unions of elements of} \mathcal{J})$$

Unfortunately, B_1 is still not a σ-algebra.

EXERCISE

Prove that B_1 is not a σ-algebra. (Hint: consider Cantor set)

Therefore, we introduce the following powerful theorem...

Constructing Probability Measure

Theorem (The Extension Theorem)

Let $\mathcal{J}$ be a semialgebra of Ω, and let $P : \mathcal{J} \to [0,1]$, with $P(\phi) = 0$, $P(\Omega) = 1$, satisfying

- *Finite superadditivity property that*

$$P\left(\bigcup_{i=1}^{k} A_i\right) \geq \sum_{i=1}^{k} P(A_i), \tag{1}$$

whenever $A_1, ..., A_k \in \mathcal{J}$, and $\bigcup_{i=1}^{k} A_i \in \mathcal{J}$, and $\{A_i\}$ are disjoint.

CONSTRUCTING PROBABILITY MEASURE

THEOREM (THE EXTENSION THEOREM)

- *Countable monotonicity property , such that*

$$P(A) \leq \sum_n P(A_n) \tag{2}$$

for $A, A_1, ..., A_n \in \mathcal{J}$ *, and* $A \subseteq \bigcup_n A_n$

Then there is a σ*-algebra* $\mathcal{M} \supseteq \mathcal{J}$*, and a countably additive probability measure* P^* *on* $\mathcal{M}$*, such that* $P^*(A) = P(A)$ *for all* $A \in \mathcal{J}$*.*
(That is, $(\Omega, \mathcal{M}, P^*)$ *is a valid probability triple, which agrees with our previous probabilities on* $\mathcal{J}$*)*

Uniqueness of Extension

Proposition

The Uniqueness of Extensions The extended probability P^* is unique, i.e. if $(\Omega, \mathcal{F}, P)$ and $(\Omega, \mathcal{M}, Q)$ are two probability triples and $\mathcal{F} \subset \mathcal{M}$, $P(\mathcal{A}) = Q(\mathcal{A})$ for all $\mathcal{A} \in \mathcal{J}$, then $P(\mathcal{A}) = Q(\mathcal{A})$ for all $\mathcal{A} \in \mathcal{F}$

In fact, the uniqueness allows us to define cumulative distribution function (CDF) of $(\Omega, \mathcal{F}, P)$.

It will be clear after we introduce Borel Sets on $\mathbb{R}$.

Now, we already define a probability function P on a σ-field $\mathcal{F}$.

DEFINITION (MEASURABLE SET)

We say a subset $\mathcal{A}$ is measurable with respect to $(\Omega, \mathcal{F}, P)$, if $\mathcal{A} \in \mathcal{F}$.

EXERCISE

Construct a non-measurable set over [0,1], with uniform distrobution.

DEFINITION (MEASURABLE FUNCTION)

Let $(\Omega, \mathcal{F}, P)$ and $(X, \mathcal{F}_x, P_x)$ be two probability spaces. We say a function $f : \Omega \to X$ is measurable, if $f^{-1}(\mathcal{A}) \in \mathcal{F}$, for all $\mathcal{A} \in \mathcal{F}_x$

Borel Set

To move on to the next topic **random variable**, we first need to consider measurable sets on $\mathbb{R}$.

Definition

Let $\mathcal{A}$ be a collection of subsets of Ω. We say a σ-field is generated by $\mathcal{A}$, denoted as $\sigma(\mathcal{A})$, if it is the smallest σ-field containing $\mathcal{A}$.
In fact, $\sigma(\mathcal{A}) = \bigcap \{\mathcal{F}\}$, for all $\mathcal{F}$ contains $\mathcal{A}$.

Now, let $\mathcal{J} = \{$ all intervals in $\mathbb{R}\}$

We say a Borel σ- field $\mathcal{B}$ is $\sigma(\mathcal{J})$, and elements in $\mathcal{B}$ is called Borel sets.

Exercise

Show that $\sigma(\mathcal{A}) = \sigma(\{(-\infty, x]\})$.

Since a probability on $\{(-\infty, x]\}$ uniquely determines a probablility on $\mathcal{J}$, and hence by the uniqueness of extension, we have the following corollary:

Corollary (Cumulative Distribution)

We can define the culmulative distribution function as $F(x) = P((-\infty, x])$, and it uniquely determines the probability function on $(\mathbb{R}, \mathcal{F}, P^*)$.

Definition

Given a probability triple $(\Omega, \mathcal{F}, P)$, a random variable is a real function $X : \Omega \to \mathbb{R}$, such that $X(\omega)$ is measurable.
Alternatively, we can write as $\forall x \in \mathbb{R}$, $\{\omega \in \Omega; X(\omega) \leq\} \in \mathcal{F}$, or $X^{-1}((-\infty, x]) \in \mathcal{F}$.

Remark:
Complements, unions and intersections are preserved under inverse image, i.e.

$$f^{-1}(D^c) = (f^{-1}(D))^c$$

$$f^{-1}(D_1 \cup D_2) = f^{-1}(D_1) \cup f^{-1}(D_2)$$

$$f^{-1}(D_1 \cap D_2) = f^{-1}(D_1) \cap f^{-1}(D_2)$$

Hence $X^{-1}(B) \in \mathcal{F}$, $\forall B$ Borel.

Random Variable

Proposition (Properties of Random Variable)

- $X = \mathbb{1}_A$ *is a random variable* $\forall A \in \mathcal{F}$
- *If* X, Y *are random variables, then* $X + c, cX, X^2, X + Y, XY$ *are all random variables*
- *If* $Z_1, Z_2, \ldots$ *are random variables and* $\lim_{n\to\infty} Z_n(\omega)$ *exists for all* $\omega \in \Omega$*, then* $Z(\omega) = \lim_{n\to\infty} Z_n(\omega)$ *is a random variable.*

The proof is left as exercises.

Random Variable

Quesion: Why do we need random variables?

Random Variable

- Random variable simplies the (probably) complicated probablity space while preserving the structure of the orighinal probability function
- Random variable induces another probablity triple on $\mathbb{R}$, that is

$$(\Omega, \mathcal{F}, P) \to (\mathcal{X}, \mathcal{B}, P_x)$$

- We can do more analysis on real number