

STAT 7200

Introduction to Advanced Probability

Lecture 3

Taylor R. Brown

1 Mathematical Background

- Other Issues
 - Equivalence Relations

2 Why Advanced Probability is Difficult?

- What is a “Reasonable” Probability Measure?
- A Limitation of Probability Measure

3 Probability Triple

- Probability Triples with Finite or Countable Sample Spaces
- Probability Triple on Uncountable Sample Space
- Borel σ -algebra

“A First Look at Rigorous Probability Theory” (Jeffrey Rosenthal) A.5, Section 1

Equivalence Relations

- **Equivalence Relations** We can define relations on a set S so that for any $x, y \in S$, either $x \sim y$ (related) or $x \not\sim y$ (not related). A relation is an equivalence relation if (a) it is reflexive, $x \sim x$ for all $x \in S$; (b) it is symmetric, $x \sim y$ whenever $y \sim x$; and (c) it is transitive, $x \sim z$ if $x \sim y$ and $y \sim z$.
- The three conditions of equivalence relations are used to ensure that we can use the equivalence relations to form a partition of set S . That is, the related elements will be assigned to the same subset in the partition, and each element must and can only belongs to a single subset.
- **Example** Let \mathbf{N} be the set of all natural number. Define relation on \mathbf{N} : for $m, n \in \mathbf{N}$, $m \sim n$ if and only if $|m - n|$ is an integer multiple of 2. Show that: 1) This is an equivalent relation. 2) This equivalent relation leads to the partition of \mathbf{N} as the union of the set of even numbers and the set of odd numbers.

Why Advanced Probability is Difficult?: What is a “Reasonable” Probability Measure?

- **Uniform Measure over $[0,1]$** Let $X \sim \text{Unif}([0,1])$, a uniform measure over $[0,1]$. Then one question is, for any subset $A \subset [0,1]$, can we define the probability $P(A) = P(X \in A)$?
- First, as $[0,1]$ contains all outcomes, we expect $P([0,1]) = 1$.
- Second, as the word “uniform” suggests each x stands “equal chance” to be the realization of X . We would then conclude the probability that X equals any particular number is 0. This “uniform” principle also implies that, for any set A of the form $[a, b], (a, b), [a, b), \text{ or } (a, b], P(A) = b - a$.
- Third, if we ask what is the probability that the value of X falls into one or more disjoint intervals, naturally the answer should be the sum of probabilities of each intervals. This leads to the requirement known as “finite additivity”:
$$P(A \cup B) = P(A) + P(B) \text{ when } A \cap B = \emptyset$$

What is a “Reasonable” Probability Measure?: Continued

- Fourth, we would like to extend “finite additivity” to “countable additivity” to account for the problems like:

$$1 = P([0, 1]) = P(0 \cup [1, \frac{1}{2}] \cup (\frac{1}{2}, \frac{1}{4}] \cup (\frac{1}{4}, \frac{1}{8}] \cup \dots)$$

- Fifth, however, we could not expect “uncountable additivity.” If this was true: we would encounter blatant paradoxes:

$$1 = P([0, 1]) = P(\cup_{x \in [0, 1]} \{x\}) = \sum_{x \in [0, 1]} P(\{x\}) = 0.$$

- Sixth, the “uniform” would also suggest that if we “shift” any subset of $[0, 1]$ by a fixed amount, the probability of the shifted subset should equal the probability of the original subset. That is, for $A \subset [0, 1]$, if we define its *r-shift* as

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

$$\text{Then } P(A \oplus r) = P(A)$$

- However, despite that we may make further arguments about on what set we could define a proper probability measure, we could not exhaust all the possibilities. And it turns out that there is always some set on which the notion of probability does not make sense.

A Limitation of Probability Measure

Proposition 1 (1.2.6)

Let $A \subseteq [0, 1]$. It is impossible to define $P(A)$ that satisfies:

- 1) for interval A in the forms of $[a, b]$, (a, b) , $[a, b)$, $(a, b]$, $P(A) = b - a$;
- 2) countable additivity: $P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ if all A_n are disjoint.
- 3) $P(A \oplus r) = P(A)$.

- **Proof** Assuming it is possible to define a probability measure that satisfies all the conditions. Here we define an equivalent relation on $[0,1]$: $x \sim y$ if and only if $x - y$ is a rational number. This equivalent relation would induce a partition on $[0,1]$. Now by the so-called Axiom of Choice, we can form a subset H consisting of exactly one element from each subset in the partition. We assume $0 \notin H$ as we can always replace it with 0.5 ($0 \sim 0.5$).

A Limitation of Probability Measure: continued

- **Proof: continued:** Now for each rational number $r \in [0, 1]$, we denote the r -shift of H as $H \oplus r$. Note that for $r_1 \neq r_2$, $H \oplus r_1$ and $H \oplus r_2$ are disjoint. If not, we can find $x \in [0, 1]$ so that $x \in H \oplus r_1$ and $x \in H \oplus r_2$. Then there are $h_1 \neq h_2 \in H$ so that $x = h_1 \oplus r_1 = h_2 \oplus r_2$. As a result, $x \sim h_1$ and $x \sim h_2$, and we must accept that $h_1 \sim h_2$. However, this would contradict to the construction principle of H : it only contains exactly one element from each subset in the partition. Thus $H \oplus r_1$ and $H \oplus r_2$ are disjoint.
- It is also easy to check that $[0, 1] = \cup_{r \in \mathbf{Q} \cap [0, 1]} (H \oplus r)$ since each element in $[0, 1]$ must be equivalent to an element in H .
- Finally note that the set of rational number is countable, we may then apply the countable additivity:
$$1 = P([0, 1]) = P(\cup_{r \in \mathbf{Q} \cap [0, 1]} (H \oplus r)) = \sum_{r \in \mathbf{Q} \cap [0, 1]} P(H \oplus r).$$

That is: $1 = \sum_{r \in \mathbf{Q} \cap [0, 1]} P(H)$.
However, this is impossible as the sum of countable number of constant must either be 0 or infinity, but never 1.

Probability Triple

- **Probability Triple** All probability models are built on a triple (Ω, \mathcal{F}, P) , the sample space, σ -algebra and probability measure.
- Ω is the sample space, the set of all possible outcomes.
- \mathcal{F} is a collection of events (measurable subsets of Ω) that forms a σ -algebra. For a collection \mathcal{F} to be qualified as a σ -algebra, the following conditions must be met:
 - a) $\emptyset \in \mathcal{F}$; b) If $A \in \mathcal{F}$, $A^c \in \mathcal{F}$.
 - c) If $A_1, A_2, \dots, A_n, \dots \in \mathcal{F}$, the $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.We can prove that a σ -algebra must also close under countable intersection, finite union and finite intersection.
- P is the probability measure, a mapping from \mathcal{F} to $[0,1]$. A probability measure must satisfy the following properties (Axioms of Probability):
 - a) $P(A) \geq 0$ for any $A \in \mathcal{F}$. b) $P(\Omega) = 1$; c) Countably additivity.More properties regarding probability measure can be found in section 2.1.

Probability Triples with Finite or Countable Sample Spaces

- When cardinality of the sample space $\Omega = \{\omega_1, \omega_2, \dots\}$ is finite or countable, we may define probability measure in the following way:
- \mathcal{F} is chosen as the power set, the collection of all possible subsets of Ω .
- For any sequence $\{a_1, a_2, \dots\}$ ($0 \leq a_i \leq 1, \sum_i a_i = 1$) with the same cardinality as Ω , we can define the probability measure as:

$$P(\omega_i) = a_i \text{ for all } \omega_i \in \Omega$$

$$P(A) = \sum_{i: \omega_i \in A} a_i$$

- It is easy to check that the above definition leads to a proper probability triple: \mathcal{F} is a σ -algebra and the probability measure P defined satisfies all the axioms of probability.

Probability Triple on Uncountable Sample Space

- However, when the sample space is uncountable, such as \mathbf{R} or $[0, 1]$, the problem of constructing probability measure is not trivial at all.
- We have already discussed that it is impossible to define probability measure on all the possible subsets. That is, we can not include all the subsets into the σ -algebra. Still, we expect that the σ -algebra should contain certain subsets, such as intervals.
- **Borel σ -algebra** can be defined as the smallest σ -algebra that contains all the open intervals (a, b) . A set that contains in Borel σ -algebra is called a **Borel set**.
- According to the fact that σ -algebra is closed to complement, finite/countable union and intersection, we can show that the Borel σ -algebra must also contains other intervals as well:
 $(\emptyset, \{a\}, [a, b], (a, b], [a, b), [a, \infty), (a, \infty), (-\infty, a), (-\infty, a])$
- For instance, $[a, b]$ must be a Borel Set since
$$[a, b] = (a, b) \cup \bigcap_{n=1}^{\infty} (a - \frac{1}{n}, a + \frac{1}{n}) \cup \bigcap_{n=1}^{\infty} (b - \frac{1}{n}, b + \frac{1}{n})$$

Borel σ -algebra

- It is easy to write much more complicated Borel set, and in fact, it is quite hard to write down a non Borel set. To see how complicated a Borel set can be, let us see the following cases:
- 1) Any open set on \mathbf{R} is a Borel Set.

To see why this is true, recall that if A is an open set is, for any $x \in A$, we may find an open interval containing x and is also a subset of A . Then for any $x \in A$, denote $I_x = (a_x, b_x)$ so that $x \in I_x, I_x \subset A$ and $a_x, b_x \in \mathbf{Q}$. And we can rewrite $A = \cup_{x \in A} I_x$. However, since I_x is an open interval with both ends as rational numbers, there can only be countable many different I_x . Thus, $\cup_{x \in A} I_x$ is essentially a countable union. And open set A is a Borel Set.

- 2) Any closed set on \mathbf{R} is a Borel Set.

A closed set is defined as the complement of a open set, so a close set should be a Borel set.

Borel σ -algebra: continued

- Now let \mathcal{G}_1 consist of all open sets, \mathcal{F}_1 consist of all the closed sets. A union of open sets is always open, but a union of closed sets may not be closed (for instance, $\cup_n [a, b - \frac{1}{n}] = [a, b)$ is not a closed set).
- Let \mathcal{G}_2 consist of all countable unions of closed sets. But the complement of such a set may not yet have been accounted for (i.e. the complement of the union of rational numbers). So we construct a new collection \mathcal{F}_2 that consists of all complements of sets in \mathcal{G}_2 .
- Continuing in this way, we can define Borel sets of ever increasing complexity (by taking unions or complements of sets generated earlier). However, such process does not stop after any finite number of steps (even after infinitely many steps, you would still not exhaust all Borel sets).
- Thus, it is impractical to define probability measure over Borel σ -algebra as we define probability measure over the discrete space. Instead, we will rely on the so-called extension theorem, so that probability measure can be defined over a smaller collection first, and then uniquely extend to a larger set.