STAT 7200

Introduction to Advanced Probability
Lecture 1

Taylor R. Brown

"A First Look at Rigorous Probability Theory" (Jeffrey Rosenthal)
Sections A.1 and A.2

Set: Definition and Basic Operations

- **Set**: A set is a collection of elements.

$$i.e, A = \{0, 1, 2, 3, 4\}$$

Element: $a \in A$ or $a \notin A$
Subset: $A \subseteq B$ if every element of A is also an element of B.
Empty Set and Universal Set: $\emptyset$ and Ω.

- **Basic Operations: Union, intersection and complement**:

$$A \cup B = \{\omega : \omega \in A \text{ or } \omega \in B\}$$
$$A \cap B = \{\omega : \omega \in A \text{ and } \omega \in B\}$$
$$A^C = \{\omega \in \Omega : \omega \notin A\}$$

Other Operations

- **Difference**:

$$A \backslash B = \{\omega : \omega \in A \text{ and } \omega \notin B\} = A \cap B^C$$

- **Product**:

$$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

Example:

1) $\Omega_1 = \{0, 1, 4\}, \Omega_2 = \{3, 7\}$

$$\Omega_1 \times \Omega_2 = \{(0, 3), (0, 7), (1, 3), (1, 7), (4, 3), (4, 7)\}$$

2) $R = \{\text{All real number}\}$
$R^2 = R \times R = \{(x, y) : x \in R, y \in R\}$

Operations Involving Collection of Sets

- In probability theory, we often have to deal with the collection of a large number of sets, which can be represented as:

  $\{A_\alpha\}_{\alpha \in I}$, where I can be any indicator set.

  And we can expend the definition of union and intersection to any collection of sets. For instance:

  $$\bigcup_\alpha A_\alpha = \{\omega : \omega \in A_\alpha \text{ for some } \alpha \in I\}$$

- **Example:** For $x_0 \in R$, prove:

  $$\bigcap_{n=1}^{\infty} \{x : x \leq x_0 + \frac{1}{n}\} = \{x : x \leq x_0\}, \quad \bigcup_{x_0 \text{ rational}} \{x : |x - x_0| < 0.1\} = R$$

- **de Morgan's Law** See A.1 of the textbook.

Partition

- **Partition**: A collection $\{A_\alpha\}_{\alpha \in I}$ of subsets of Ω forms a partition of Ω if:

  1) $\{A_\alpha\}_{\alpha \in I}$ are disjoint: $A_{\alpha_1} \cap A_{\alpha_2} = \emptyset$ for any pair $\alpha_1 \neq \alpha_2$.

  2) and $\bigcup_\alpha A_\alpha = \Omega$.

- **Example:** Define $\{A_n\}_{n \in \mathbb{N}}$ so that $A_1 = [1, \infty)$, $A_k = [1/k, 1/(k-1))$ for $k \geq 2$. Then $\{A_n\}_{n \in \mathbb{N}}$ forms a partition of $(0, \infty)$.

- For collection $\{A_n\}_{n \in \mathbb{N}}$ (may not be disjoint) that $\bigcup_1^\infty A_n = \Omega$.
  Define:
  $B_1 = A_1$, $B_2 = A_2 \backslash A_1$, $B_3 = A_3 \backslash (A_2 \cup A_1) \cdots$,
  $B_n = A_n \backslash \bigcup_1^{n-1} A_k, \cdots$,
  Then $\{B_n\}_{n \in \mathbb{N}}$ forms a partition of Ω.

Set and Event

- In probability theory, we call universal set Ω as sample space, set A as event, and element ω as outcome. Furthermore, we can interpret the union of a collection of events as the event in which one or more events in the collection occur, and intersection as the event in which all the events in the collection occur. One key to understand rigorous probability is to think in term of sets/events.

  Example: Suppose that X and Y are two random variables, ε is a constant. Prove the following probability inequality:

  $$P(X + Y < \varepsilon) \leq P(X < \varepsilon/2) + P(Y < \varepsilon/2)$$

  Hint: Show that, $\{X + Y < \varepsilon\} \subseteq \{X < \varepsilon/2\} \cup \{Y < \varepsilon/2\}$.

Function: Definition

- **Function** is defined as a mapping from one set to another ($f : \Omega_1 \to \Omega_2$). In probability, the most commonly used function is the random variable, which refers to a function from the sample space to the real space.

- For a function f from set Ω_1 to Ω_2, we say f maps Ω_1 **onto** Ω_2 if for every $b \in \Omega_2$, there is at least one $a \in \Omega_1$ so that $f(a) = b$. f is **one to one** if $f(a_1) = f(a_2)$ would imply $a_1 = a_2$. These are also known as surjective and injective functions, respectively. We call f a **bijection** from Ω_1 to Ω_2 if f is both one to one and onto.

  Example: Considering functions from R to R.
  $f(x) = x^2$ is not onto nor one to one.
  $f(x) = e^x$ is one to one but not onto.
  $f(x) = x$ is a bijection.

Image and Inverse Image

- **Function** For $f : \Omega_1 \to \Omega_2$, if $S_1 \subseteq \Omega_1$ and $S_2 \subseteq \Omega_2$, then the **image** of S_1 under f is defined as
  $f(S_1) = \{\omega_2 \in \Omega_2 : \omega_2 = f(\omega_1) \text{ for some } \omega_1 \in S_1\}$, and the **inverse image** of S_2 is defined as: $f^{-1}(S_2) = \{\omega_1 \in \Omega_1 : f(\omega_1) \in S_2\}$.

  Example:
  $f(x) = x^2 : R \to R$.
  The image of $(1, 2)$ is $f(1, 2) = (1, 4)$, and the inverse image of $(-1, 2)$ is $f^{-1}((-1, 2)) = (-\sqrt{2}, \sqrt{2})$.

Inverse Image and Set Operations

Theorem 1

The inverse images preserve the usual set operations:
$f^{-1}(S_1 \cup S_2) = f^{-1}(S_1) \cup f^{-1}(S_2)$
$f^{-1}(S_1 \cap S_2) = f^{-1}(S_1) \cap f^{-1}(S_2)$
$f^{-1}(S_1^C) = (f^{-1}(S_1))^C$

Proof: Inverse images preserve union operation.
For any $\omega \in f^{-1}(S_1 \cup S_2)$, by definition of inverse image,
$f(\omega) \in S_1 \cup S_2$. Thus, we have $f(\omega) \in S_1$ or $f(\omega) \in S_2$. We can then
deduce that either $\omega \in f^{-1}(S_1)$ or $\omega \in f^{-1}(S_2)$, which leads to
$\omega \in f^{-1}(S_1) \cup f^{-1}(S_2)$. Therefore, $f^{-1}(S_1 \cup S_2) \subseteq f^{-1}(S_1) \cup f^{-1}(S_2)$.
Similarly, $\omega \in f^{-1}(S_1) \cup f^{-1}(S_2)$ implies $\omega \in f^{-1}(S_1 \cup S_2)$, which
implies $f^{-1}(S_1) \cup f^{-1}(S_2) \subseteq f^{-1}(S_1 \cup S_2)$.

Combing the above results, we have
$f^{-1}(S_1 \cup S_2) = f^{-1}(S_1) \cup f^{-1}(S_2)$.

Cardinality: Definition

- **Cardinality** is a concept that tries to measure the size of a set through counting the number of elements.
- **Question:** How should we define cardinality for a set with an infinite number of elements? For instance, which of the following two sets has the greater cardinality: the set of even number, or the set of natural number?
- Two sets Ω_1 and Ω_2 have the same **cardinality** if there is a *bijection* from Ω_1 to Ω_2. Similarly, we say that the cardinality of Ω_1 is less than the cardinality of Ω_2 if there is a one to one function from Ω_1 to Ω_2.
- For the purpose of this class, the infinity sets have two types of infinity: **countable (or countably infinite** and **uncountable (or uncountably infinite)**.

Countable Set

- **Countable Set** A set Ω is countable if it has the same cardinality as the set of natural number $N = \{1, 2, 3, \cdots, \}$. Or equivalently, Ω is countable if the all the elements in Ω can be listed as $\omega_1, \omega_2, \cdots,$

- **Question:** Show that, the set of all integer number Z is countable.

Proposition 2

Any infinite subset of N (or any countable set) is countable.

- For the set of rational number Q, note that we can represent each rational number as: $q = m/n$, where $m \in Z$ and $n \in N$. In this regard, Q is equivalent to a subset of $Z \times N$. Is $Z \times N$ countable?

Cardinality of the Product of Countable Sets

Proposition 3

The product of two countable sets is countable. That is, if Ω_1 and Ω_2 are countable, then $\Omega_1 \times \Omega_2$ is also countable.

- **Proof** Use the so-called zig-zag method. Since Ω_1 and Ω_2 are countable, we can write $\Omega_1 = \{a_1, a_2, \cdots\}$, and $\Omega_2 = \{b_1, b_2, \cdots\}$. Then each element in $\Omega_1 \times \Omega_2$ can be represented as (a_m, b_n). Now let us construct a grid with countably many rows and columns, and put (a_m, b_n) on the intersection of the mth row and nth column. It is easy to see that this grid contains all elements in $\Omega_1 \times \Omega_2$. Furthermore, this grid allows us to list the elements in $\Omega_1 \times \Omega_2$ as a sequence by following a zig-zag path starting from (a_1, b_1).

Corollary 4

The set of rational number $\mathbb{Q}$ is countable.

Cardinality of the Union of Countable Sets

Proposition 5

The union of a finite number, or a countable number of countable sets is countable. That is, given a sequence of countable set $\Omega_1, \Omega_2, \ldots$, the union $\bigcup_i \Omega_i$ is also countable.

- **Proof** Similar to the previous proof. The only difference is we will place the mth element of set Ω_n on the intersection of the mth row and nth column.
- **Remark** The **product** of a countable number of sets is not countable. We will discuss this in a future lecture.

Uncountable Set

Theorem 6

The set of real numbers R is not countable.

- **Uncountable** The cardinality of R is uncountable.
- **Proof** We only need to prove that the interval $[0, 1)$ is not countable.

  1) Since $[0, 1)$ is clearly infinity, we assume it is countable. Then we can list all numbers in $[0, 1)$ as a sequence $x_1, x_2, \cdots ,$. For each x_i, we assume its (unique) decimal expansion as $x_i = 0.d_{i1}d_{i2}d_{i3} \cdots$.

  2) Define a new number $f = 0.f_1 f_2 f_3 \cdots$ so that $f_k = d_{kk} + 1$ (If $d_{kk} = 9$, then $f_k = 0$).

  3) It is clearly that $f \in [0, 1)$ and $f \neq x_k$ for any k (since the kth digit of f differs from the kth digit of x_k). In this way, we construct a number that belongs to $[0, 1)$ but is not included in the presumably exhaustive list.

  4) From this contradiction we deduce that $[0, 1)$ is not countable.

Cardinality of the Support of a Discrete Random Variable

Theorem 7

The support of any discrete random variable is at most countable.

- **Note** If X is a discrete random variable, then the support of X is defined as: $S = \{x : P(X = x) > 0\}$.
- **Proof** We firstly note that, for discrete r.v. X, $\sum_{x \in S} P_x = 1$.

  1) Firstly we create a countable partition of the support: $S = \bigcup_1^\infty A_k$, $A_k = \{x : \frac{1}{k+1} < P_x \leq \frac{1}{k}\}$.

  2) Now we argue that, the number of points in each subset A_k must be finite. If not, since $P_x > 1/(k+1)$ for any $x \in A_k$, $\sum_{x \in A_k} P_x$ would equal infinity if the number of points in A_k is infinity.

  3) Therefore, S can be represented as a countable partition, and the number of points in each subset is finite. Consequently, we can list all the points in S as a finite or countable sequence. Thus the support S of a discrete random variable must be at most countable.