556: MATHEMATICAL STATISTICS I
INTRODUCTION TO MEASURE AND INTEGRATION

1. PROBABILITY AND MEASURE

In formal probability theory, a probability specification has three components:

- **The Sample Space :** a set Ω with elements ω
- **A Sigma-Algebra :** a collection of subsets of Ω, denoted $\mathcal{E}$, say, that obeys the following properties
  - I $\Omega \in \mathcal{E}$
  - II *Closure under countable union:*
  $$E_1, E_2, \ldots \in \mathcal{E} \Longrightarrow \bigcup_{k=1}^{\infty} E_k \in \mathcal{E}$$
  - III *Closure under complementation:* $E \in \mathcal{E} \Longrightarrow E' \in \mathcal{E}$
- **A Probability Measure :** a real-valued set function $\mathbb{P}$ that obeys the general properties of a **measure** with one additional requirement. A measure, denoted μ, is a real-valued set function such that for arbitrary sets E and $E_1, E_2, \ldots$
  - I *Non-negativity:* $\mu(E) \geq 0$.
  - II *Sub-additivity:*
  $$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} \mu(E_k)$$
  - III *Preservation under Limits:* If $E_1 \subset E_2 \subset \ldots$ is an increasing sequence of sets, we use the notation
  $$\lim_{n \longrightarrow \infty} E_n \equiv \bigcup_{i=1}^{\infty} E_i.$$
  Then
  $$\mu\left(\lim_{n \longrightarrow \infty} E_n\right) = \lim_{n \longrightarrow \infty} \mu(E_n).$$
  Similarly, if $E_1 \supset E_2 \supset \ldots$ is a decreasing sequence of sets, we use the notation
  $$\lim_{n \longrightarrow \infty} E_n \equiv \bigcap_{i=1}^{\infty} E_i.$$
  and again
  $$\mu\left(\lim_{n \longrightarrow \infty} E_n\right) = \lim_{n \longrightarrow \infty} \mu(E_n).$$

  Examples of Measures: For sample space Ω, and $A \subseteq \Omega$,
  - *Counting Measure :* $\mu(A) = |A|$ if A is a finite subset, $\mu(A) = \infty$ if A is an infinite subset.
  - *Lebesgue Measure :* If $\Omega \equiv \mathbb{R}$, then, for $a < b$,
  $$\mu((a, b)) = \mu((a, b]) = \mu([a, b)) = \mu([a, b]) = b - a.$$

  Probability measures have the additional property that $\mathbb{P}(\Omega) = 1$.

We use the terminology

- **Measurable space** to describe the pair $(\Omega, \mathcal{E})$
- **Measure space** to describe the triple $(\Omega, \mathcal{E}, \mu)$
- **Probability space** to describe the triple $(\Omega, \mathcal{E}, \mathbb{P})$

2. MEASURABLE FUNCTIONS

DEFINITION Borel σ-algebra
Let $\Omega \equiv \mathbb{R}$, and $\mathcal{C}$ be the collection of all finite open intervals of $\mathbb{R}$, that is

$$\mathcal{C} \equiv \{(a,b) : a < b \in \mathbb{R}.\}$$

Then $\mathcal{B} \equiv \sigma(\mathcal{C})$ is the **Borel σ-algebra**, and $B \in \mathcal{B}$ are the **Borel sets**, which are of the form

$$(a,b), (a,b], [a,b), [a,b] \qquad -\infty \leq a \leq b \leq \infty.$$

DEFINITION Measurability

The real-valued function f defined with domain $E \subset \Omega$, for measurable space $(\Omega, \mathcal{E})$, is **Borel measurable** with respect to $\mathcal{E}$ if the inverse image of set B, defined as

$$f^{-1}(B) \equiv \{\omega \in E : f(\omega) \in B\}$$

is an element of σ-algebra $\mathcal{E}$, for all Borel sets B of $\mathbb{R}$ (strictly, of the *extended* real number system $\mathbb{R}^*$, including $\pm\infty$ as elements). The following conditions are each necessary and sufficient for f to be measurable

(a) $f^{-1}(A) \in \mathcal{E}$ for all open sets $A \subset \mathbb{R}^*$,

(b) $f^{-1}([-\infty, x)) \in \mathcal{E}$ for all $x \in \mathbb{R}^*$,

(c) $f^{-1}([-\infty, x]) \in \mathcal{E}$ for all $x \in \mathbb{R}^*$,

(d) $f^{-1}([x, \infty]) \in \mathcal{E}$ for all $x \in \mathbb{R}^*$,

(e) $f^{-1}((x, \infty]) \in \mathcal{E}$ for all $x \in \mathbb{R}^*$.

NOTES:

(i) The **Borel σ-algebra** in $\mathbb{R}$, $\mathcal{B}$, is the smallest (or **minimal**) σ-algebra containing all **open sets**.

(ii) It is possible to extend this definition to a general **topological space** Ω equipped with a **topology**, that is, a collection, $\mathcal{T}$, of sets in Ω that (I) $\mathcal{T}$ contains $\emptyset$ and Ω, (II) $\mathcal{T}$ is closed under finite intersection, and (III) if $\mathcal{A}$ is a sub-collection of $\mathcal{T}$, $\mathcal{A} \subset \mathcal{T}$, and $A_1, A_2, A_3, ... \in \mathcal{A}$, then

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{T}.$$

In this context, it is possible to define a general Borel σ-algebra on Ω; the **open sets** are the elements $T_1, T_2, T_3, ...$ of the topology $\mathcal{T}$, and the Borel sets are the elements of the smallest σ-algebra generated by $\mathcal{T}$, $\sigma(\mathcal{T})$. However, we will not be studying general toplogical spaces; we shall restrict attention to $\mathbb{R}$, and thus refer to **the** Borel sets and **the** Borel σ-algebra, meaning the Borel sets/σ-algebra defined on $\mathbb{R}$.

(iii) Strictly, a function f is a **Borel function** if, for $B \in \mathcal{B}$, $f^{-1}(B) \in \sigma(\mathcal{T})$; however, we will generally consider measure spaces $(\Omega, \mathcal{E})$ and say that f is a **Borel function** if it is Borel measurable, as defined in the first paragraph above.

The measurability of functions is preserved under the following operations: if g_1 and g_2 are measurable functions defined on $E \subset \Omega$ into $\mathbb{R}^*$, and c is any real number, then all of the following composite and other related functions are measurable

$$g_1 + g_2, g_1 + c, g_1 g_2, c g_1, g_1/g_2, |g_1|^c, g_1 \vee g_2, g_1 \wedge g_2, g_1^+, g_1^-.$$

where

- $g_1 \vee g_2(x) = \max\{g_1(x), g_2(x)\}$
- $g_1 \wedge g_2(x) = \min\{g_1(x), g_2(x)\}$
- $f^+(x) = f(x) \vee 0 = \max\{f(x), 0\}$
- $f^-(x) = -f(x) \vee 0 = \max\{-f(x), 0\}$

so that

$$f(x) = f^+(x) - f^-(x) \qquad\qquad |f(x)| = f^+(x) + f^-(x).$$

Furthermore, if $\{g_n\}$ is a sequence of measurable functions, then the functions defined by

$$\overline{g}(x) = \sup_n g_n(x) \qquad\qquad \underline{g}(x) = \inf_n g_n(x)$$

are also measurable. Finally, the functions $\limsup_n g_n(x)$ and $\liminf_n g_n(x)$ are also measurable.

3. INTEGRATION

Let $(\Omega, \mathcal{E}, \mu)$ be a measure space, and ψ be a **non-negative** simple function, $\psi : \Omega \longrightarrow \mathbb{R}^\star$, that is, for $\omega \in \Omega$,

$$\psi(\omega) = \sum_{i=1}^{k} a_i I_{A_i}(\omega)$$

for real constants $a_1, ..., a_k \geq 0$ and measurable sets $A_1, ..., A_k \in \mathcal{E}$, for some $k = 1, 2, 3, ...$, where $I_A(\omega)$ is the indicator function for set A.

(I) The **integral of ψ with respect to** μ is denoted and defined by

$$\int_\Omega \psi \, d\mu = \sum_{i=1}^{k} a_i \mu(A_i).$$

(II) Now suppose that f is a **non-negative** (Borel) measurable function, and let $\mathcal{S}_f$ be the set of all non-negative **simple** functions defined by

$$\mathcal{S}_f \equiv \{\psi : \psi(\omega) \leq f(\omega), \forall \omega \in \Omega\}.$$

Then the integral of f with respect to μ is defined by

$$\int_\Omega f \, d\mu = \sup_{\psi \in \mathcal{S}_f} \int_\Omega \psi \, d\mu$$

that is, the **supremum** (least upper bound) over all possible choices of k, $a_1, ..., a_k \in \mathbb{R}^+$ and $A_1, ..., A_k \in \mathcal{E}$ such that, for all $\omega \in \Omega$,

$$\psi(\omega) = \sum_{i=1}^{k} a_i I_{A_i}(\omega) \leq f(\omega)$$

We refer to this as the **Supremum Definition**.

(III) Finally, suppose that f is an **arbitrary** measurable function defined on Ω. Then, using the max/min functions

$$f^+(\omega) = \max\{f(\omega), 0\} \qquad f^-(\omega) = \max\{-f(\omega), 0\} \qquad \therefore \qquad f(\omega) = f^+(\omega) - f^-(\omega),$$

we define the integral of f with respect to μ by

$$\int_\Omega f \, d\mu = \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu.$$

where the two integrals on the right hand side are integrals of non-negative functions, and thus given by the supremum definition above.

NOTES

(i) In (III) above, it might be that at least one of the two integrals

$$\int_\Omega f^+ \, d\mu \qquad \int_\Omega f^- \, d\mu.$$

is not finite. If precisely one is finite, we say that

$$\int_\Omega f^+ \, d\mu = \infty.$$

and that the integral of f *exists*. If both are finite, we say that the integral of f *exists and is finite*, and f is *integrable with respect to* μ. If neither is finite, then we say that the integral of f *does not exist*, and f is *not-integrable*.

(ii) For $E \subset \Omega$, we can also define

$$\int_E f \, d\mu = \int_E I_E f \, d\mu$$

(iii) All of the following pieces of notation are equivalent and used in the literature:

$$\int f \, d\mu \qquad \int_\Omega f \, d\mu \qquad \int f(\omega) \, d\mu \qquad \int f(\omega) \, d\mu(\omega) \qquad \int f(\omega) \, \mu(d\omega)$$

(iv) Previous results show that measurable functions have representations as limits of sequences of simple functions. Other results show that measurability is preserved under composition, and also under limit behaviour. Consider a non-negative measurable function f. Then

$$f = \lim_{n \longrightarrow \infty} \psi_n$$

for a sequence of non-negative simple functions $\psi_1, \psi_2, \ldots$ with $0 \leq \psi_n(\omega) \leq f(\omega)$, for all n and for all $\omega \in \Omega$. Then it can be shown

$$\lim_{n \longrightarrow \infty} \int \psi_n \, d\mu = \lim_{n \longrightarrow \infty} \sum_{i=1}^{k_n} a_{n,i} I_{A_{n,i}} = \sum_{i=1}^{k} a_i I_{A_i},$$

say, where

$$\lim_{n \longrightarrow \infty} k_n = k \qquad \lim_{n \longrightarrow \infty} a_{n,i} = a_i \qquad \lim_{n \longrightarrow \infty} I_{A_{n,i}} = I_{A_i}.$$

Thus

$$\lim_{n \longrightarrow \infty} \int \psi_n \, d\mu = \int \lim_{n \longrightarrow \infty} \psi_n \, d\mu = \int f \, d\mu$$

and the integral is preserved under the limit operation.

$$\lim_{n \longrightarrow \infty} \int \psi_n \, d\mu = \int \lim_{n \longrightarrow \infty} \psi_n \, d\mu = \int f \, d\mu$$